LNAI 4869

Francisco Botana
Tomas Recio (Eds.)

Automated Deduction in Geometry

6th International Workshop, ADG 2006
Pontevedra, Spain, August\September 2006
Revised Papers

Springer

Francisco Botana Tomas Recio (Eds.)

Automated Deduction in Geometry

6th International Workshop, ADG 2006
Pontevedra, Spain, August 31-September 2, 2006
Revised Papers

Springer

Series Editors

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

Volume Editors

Francisco Botana
EUET Forestal
Universidad de Vigo
Campus A Xunqueira
36005 Pontevedra, Spain
E-mail: fbotana@uvigo.es

Tomas Recio
Universidad de Cantabria
Departamento de Matemáticas, Estadística y Computación
Avenida de los Castros, s/n, 39071 Santander, Spain
E-mail: tomas.recio@unican.es

Library of Congress Control Number: 2007941260

CR Subject Classification (1998): I.2.3, I.3.5, F.4.1, I.5, G.2

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN	0302-9743
ISBN-10	3-540-77355-X Springer Berlin Heidelberg New York
ISBN-13	978-3-540-77355-9 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: 12206292 06/3180 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 4869
Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Preface

After five successful editions (Toulouse, 1996; Beijing, 1998; Zurich, 2000; Linz, 2002, Gainesville, Fl., 2004), the series of international workshops on Automated Deduction in Geometry (ADG) has consolidated its fundamental role concerning the scientific community working on the interaction between geometry and automated deduction. From August 31 to September 2, 2006, a new ADG meeting took place at the Pontevedra (Galicia, Spain) campus of the University of Vigo, as a satellite event of the International Congress of Mathematicians (Madrid, August 22–30, 2006). We acknowledge the financial support for ADG 2006, provided by the University of Vigo and the Spanish Ministerio de Educación y Ciencia under grant MTM2005-24580-E.

It was a fruitful meeting – made possible by the work of the Organizing Committee (see next page) – for exchanging ideas and for the presentation of original results and software novelties – 21 contributions in total – under the scientific guidance of the Program Committee (listed on the next page). Moreover, it was a privilege to receive the lectures of our distinguished guest speakers, Thomas Hales (U. Pittsburgh) and Martin Peternell (T.U. Wien), dealing with the so-called Flyspeck project, i.e., the automatization of Hales' solution to Kepler's conjecture, and with rational offset surfaces and related issues in CAGD, respectively.

Shortly after the meeting, a call for papers – within the scope of ADG, but with content not necessarily related to a presentation at ADG 2006 – was launched. After a long and detailed process of peer review and revision, we – the editors – have selected the 13 papers of this volume, as a testimony of the current state of the art concerning automated deduction in geometry.

This volume includes a paper by X. Chen and D. Wang proposing a system in the form of a textbook – an electronic geometry textbook, to be more precise – for managing geometric knowledge dynamically, effectively, and interactively. The contribution by T. Hales, in the context of the "Flyspeck" project, describes an algorithm that decides whether a region in three dimensions, described by quadratic constraints, is equidecomposable with a collection of primitive regions and, when a decomposition exists, finds the volume of the given region. P. Janičić and P. Quaresma present an application of automatic theorem proving in the verification of constructions made with dynamic geometry software. The paper by P. Lebmeir and J. Richter-Gebert proposes an algorithm for automated recognition of computationally constructed curves and discusses several aspects of the recognition problem. R. H. Lewis and E. Coutsias deal with polynomial systems, flexibility of three-dimensional objects, computational chemistry, and computer algebra. D. Lichtblau's contribution on computational real enumerative geometry discusses the number and reality of the cylinders generically determined by five points in R^3. D. Michelucci and S. Foufou address the detection of dependences in geometric constraints solving, and propose to use the recently published witness method. The paper by A. Montes and T. Recio merges two techniques (automatic discovery and minimal canonical comprehensive Gröbner systems) to discover missing hypotheses in generally false statements. J. Narboux describes the mechanization of the proofs of the first eight chapters of the classic book "Metamathematische Methoden in der Geometrie" by Schwabäuser, Szmielew and Tarski. The paper by P. Pech deals with the problem, posed by Chou long ago, of finding a natural geometry problem where hypotheses are not described by a radical ideal, such as the existence of regular polygons (pentagons, heptagons) of even dimension. E. Roanes–Macías and E. Roanes–Lozano present a Maple package, on the interaction of computer algebra and dynamic geometry, for investigating problems about configuration theorems in 3D geometry and performing mechanical theorem proving and discovery. P. Todd presents an interactive symbolic geometry package, "Geometry Expressions," generating algebraic formulas from geometry in an interactive style which is convenient not only for high school students, but also for mechanical engineers, graphics programmers, architects, surveyors, machinists, and many more. Finally, the paper by L. Yang and Z. Zeng, employing a method of distance geometry, achieves the symbolic solution to the following problem: express the edge-lengths of a tetrahedron in terms of its heights and widths.

We, the editors, would like to thank the efforts of so many anonymous referees involved in the process of selection and improvement of the submitted papers. We think that, as a consequence of their work, this collection of papers, although necessarily incomplete, shows the lively variety of topics and methods and the current applicability of ADG to different branches of mathematics and to other sciences and technologies.

October 2007

Francisco Botana
Tomas Recio

Organization

Invited Speakers

Thomas Hales (University of Pittsburgh, USA)
Martin Peternell (Vienna University of Technology, Austria)

Organizing Committee

Francisco Botana (Pontevedra, Spain), Chair
Miguel Abánades (Madrid, Spain)
Jesús Escribano (Madrid, Spain)
Eugenio Roanes–Lozano (Madrid, Spain)
José L. Valcarce (Santiago, Spain)

Program Committee

Tomas Recio (Santander, Spain), Chair
Hirokazu Anai (Kawasaki, Japan)
Giuseppa Carrà Ferro (Catania, Italy)
Shang-Ching Chou (Wichita, USA)
Arthur Chtcherba (New York, USA)
Luis Fariñas del Cerro (Toulouse, France)
Jacques D. Fleuriot (Edinburgh, UK)
Xiao-Shan Gao (Beijing, China)
Laureano González-Vega (Santander, Spain)
Hoon Hong (Raleigh, USA)
Deepak Kapur (Albuquerque, USA)
Hongbo Li (Beijing, China)
Manfred Minimair (South Orange, USA)
Jürgen Richter-Gebert (Munich, Germany)
Meera Sitharam (Gainesville, USA)
Thomas Sturm (Passau, Germany)
Quoc-Nam Tran (Beaumont, USA)
Dongming Wang (Beijing,China/Paris, France)
Neil White (Gainesville, USA)
Franz Winkler (Linz, Austria)
Lu Yang (Chengdu, China)
Zhenbing Zeng (Shanghai, China)

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 4874: J. Neves, M.F. Santos, J.M. Machado (Eds.), Progress in Artificial Intelligence. XVIII, 704 pages. 2007.

Vol. 4869: F. Botana, T. Recio (Eds.), Automated Deduction in Geometry. X, 213 pages. 2007.

Vol. 4850: M. Lungarella, F. Iida, J. Bongard, R. Pfeifer (Eds.), 50 Years of Artificial Intelligence. X, 399 pages. 2007.

Vol. 4845: N. Zhong, J. Liu, Y. Yao, J. Wu, S. Lu, K. Li (Eds.), Web Intelligence Meets Brain Informatics. XI, 516 pages. 2007.

Vol. 4830: M.A. Orgun, J. Thornton (Eds.), AI 2007: Advances in Artificial Intelligence. XIX, 841 pages. 2007.

Vol. 4828: M. Randall, H.A. Abbass, J. Wiles (Eds.), Progress in Artificial Life. XII, 402 pages. 2007.

Vol. 4827: A. Gelbukh, Á.F. Kuri Morales (Eds.), MICAI 2007: Advances in Artificial Intelligence. XXIV, 1234 pages. 2007.

Vol. 4826: P. Perner, O. Salvetti (Eds.), Advances in Mass Data Analysis of Signals and Images in Medicine, Biotechnology and Chemistry. X, 183 pages. 2007.

Vol. 4819: T. Washio, Z.-H. Zhou, J.Z. Huang, X. Hu, J. Li, C. Xie, J. He, D. Zou, K.-C. Li, M.M. Freire (Eds.), Emerging Technologies in Knowledge Discovery and Data Mining. XIV, 675 pages. 2007.

Vol. 4798: Z. Zhang, J.H. Siekmann (Eds.), Knowledge Science and Engineering and Management. XVI, 669 pages. 2007.

Vol. 4795: F. Schilder, G. Katz, J. Pustejovsky (Eds.), Annotating, Extracting and Reasoning about Time and Events. VII, 141 pages. 2007.

Vol. 4790: N. Dershowitz, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIII, 562 pages. 2007.

Vol. 4788: D. Borrajo, L. Castillo, J.M. Corchado (Eds.), Current Topics in Artificial Intelligence. XI, 280 pages. 2007.

Vol. 4775: A. Esposito, M. Faundez-Zanuy, E. Keller, M. Marinaro (Eds.), Verbal and Nonverbal Communication Behaviours. XII, 325 pages. 2007.

Vol. 4772: H. Prade, V.S. Subrahmanian (Eds.), Scalable Uncertainty Management. X, 277 pages. 2007.

Vol. 4766: N. Maudet, S. Parsons, I. Rahwan (Eds.), Argumentation in Multi-Agent Systems. XII, 211 pages. 2007.

Vol. 4755: V. Corruble, M. Takeda, E. Suzuki (Eds.), Discovery Science. XI, 298 pages. 2007.

Vol. 4754: M. Hutter, R.A. Servedio, E. Takimoto (Eds.), Algorithmic Learning Theory. XI, 403 pages. 2007.

Vol. 4737: B. Berendt, A. Hotho, D. Mladenic, G. Semeraro (Eds.), From Web to Social Web: Discovering and Deploying User and Content Profiles. XI, 161 pages. 2007.

Vol. 4733: R. Basili, M.T. Pazienza (Eds.), AI*IA 2007: Artificial Intelligence and Human-Oriented Computing. XVII, 858 pages. 2007.

Vol. 4724: K. Mellouli (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty. XV, 914 pages. 2007.

Vol. 4722: C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (Eds.), Intelligent Virtual Agents. XV, 425 pages. 2007.

Vol. 4720: B. Konev, F. Wolter (Eds.), Frontiers of Combining Systems. X, 283 pages. 2007.

Vol. 4702: J.N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Knowledge Discovery in Databases: PKDD 2007. XXIV, 640 pages. 2007.

Vol. 4701: J.N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Machine Learning: ECML 2007. XXII, 809 pages. 2007.

Vol. 4696: H.-D. Burkhard, G. Lindemann, R. Verbrugge, L.Z. Varga (Eds.), Multi-Agent Systems and Applications V. XIII, 350 pages. 2007.

Vol. 4694: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. XXIX, 1126 pages. 2007.

Vol. 4693: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXII, 1380 pages. 2007.

Vol. 4692: B. Apolloni, R.J. Howlett, L. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LV, 882 pages. 2007.

Vol. 4687: P. Petta, J.P. Müller, M. Klusch, M. Georgeff (Eds.), Multiagent System Technologies. X, 207 pages. 2007.

Vol. 4682: D.-S. Huang, L. Heutte, M. Loog (Eds.), Advanced Intelligent Computing Theories and Applications. XXVII, 1373 pages. 2007.

Vol. 4676: M. Klusch, K.V. Hindriks, M.P. Papazoglou, L. Sterling (Eds.), Cooperative Information Agents XI. XI, 361 pages. 2007.

Vol. 4667: J. Hertzberg, M. Beetz, R. Englert (Eds.), KI 2007: Advances in Artificial Intelligence. IX, 516 pages. 2007.

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

Vol. 4659: V. Mařík, V. Vyatkin, A.W. Colombo (Eds.), Holonic and Multi-Agent Systems for Manufacturing. VIII, 456 pages. 2007.

Vol. 4651: F. Azevedo, P. Barahona, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. VIII, 185 pages. 2007.

Vol. 4648: F. Almeida e Costa, L.M. Rocha, E. Costa, I. Harvey, A. Coutinho (Eds.), Advances in Artificial Life. XVIII, 1215 pages. 2007.

Vol. 4635: B. Kokinov, D.C. Richardson, T.R. Roth-Berghofer, L. Vieu (Eds.), Modeling and Using Context. XIV, 574 pages. 2007.

Vol. 4632: R. Alhajj, H. Gao, X. Li, J. Li, O.R. Zaïane (Eds.), Advanced Data Mining and Applications. XV, 634 pages. 2007.

Vol. 4629: V. Matoušek, P. Mautner (Eds.), Text, Speech and Dialogue. XVII, 663 pages. 2007.

Vol. 4626: R.O. Weber, M.M. Richter (Eds.), Case-Based Reasoning Research and Development. XIII, 534 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. 4520: M.V. Butz, O. Sigaud, G. Pezzulo, G. Baldassarre (Eds.), Anticipatory Behavior in Adaptive Learning Systems. X, 379 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. 4460: S. Aguzzoli, A. Ciabattoni, B. Gerla, C. Manara, V. Marra (Eds.), Algebraic and Proof-theoretic Aspects of Non-classical Logics. VIII, 309 pages. 2007.

Vol. 4457: G.M.P. O'Hare, A. Ricci, M.J. O'Grady, O. Dikenelli (Eds.), Engineering Societies in the Agents World VII. XI, 401 pages. 2007.

Vol. 4456: Y. Wang, Y.-m. Cheung, H. Liu (Eds.), Computational Intelligence and Security. XXIII, 1118 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. 4442: L. Antunes, K. Takadama (Eds.), Multi-Agent-Based Simulation VII. X, 189 pages. 2007.

Vol. 4441: C. Müller (Ed.), Speaker Classification II. X, 309 pages. 2007.

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

Vol. 4434: G. Lakemeyer, E. Sklar, D.G. Sorrenti, T. Takahashi (Eds.), RoboCup 2006: Robot Soccer World Cup X. XIII, 566 pages. 2007.

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

Vol. 4428: S. Edelkamp, A. Lomuscio (Eds.), Model Checking and Artificial Intelligence. IX, 185 pages. 2007.

¥530.00元

Table of Contents

Towards an Electronic Geometry Textbook

Xiaoyu Chen[1] and Dongming Wang[1,2]

[1] LMIB – School of Science, Beihang University, Beijing 100083, China
[2] Laboratoire d'Informatique de Paris 6, Université Pierre et Marie Curie – CNRS
104 avenue du Président Kennedy, F-75016 Paris, France

Abstract. This paper proposes a system in the form of a textbook for managing geometric knowledge dynamically, effectively, and interactively. Such a system, called an *Electronic Geometry Textbook*, can be viewed or printed as a traditional textbook and run as dynamic software on computer. The knowledge in the textbook is being formalized by using standard formal languages and may be processed by software modules developed for geometric computing and reasoning, diagram generation, and visualization. The textbook can be generated automatically by organizing and presenting the textbook data according to some specifications. The system allows the user to manipulate (query, modify, restructure, etc.) the textbook with automated consistency checking. We present the main ideas on the design of the electronic geometry textbook, explain the features of the system, propose five phases of creating and managing the geometric knowledge in the textbook, discuss the involved tasks and some of the fundamental research problems in each phase, and report our progress and experiments on a preliminary implementation of the system.

1 Motivation and Introduction

Elementary geometry has been developed for more than 2000 years and an enormous amount of geometric knowledge has been accumulated. It is now desirable to digitalize, organize, and process various kinds of geometric knowledge including concepts, objects, axioms, theorems, diagrams, texts, and computing and reasoning mechanisms using advanced computer technology to make them more easily accessible, manageable, and usable.

In recent decades, remarkable progress has been made on computer-aided geometric problem solving. Most of the theorems in elementary geometry can now be proved or even rediscovered automatically on computer. Software systems are available for generating algebraic and readable proofs and drawing dynamic diagrams automatically. However, research in this direction is by no means exhausted and there is an urgent need to exchange data among different dynamic geometry software systems (such as Cinderella [5], GCLC [7], and GEOTHER [15]) and to integrate them together with other mathematical software like computer algebra systems and automated theorem provers to extend their power and functionality. Little work has been done on the management of geometric knowledge

F. Botana and T. Recio (Eds.): ADG 2006, LNAI 4869, pp. 1–23, 2007.

and there is no standard geometric knowledge base that can be accessed and used by different geometric software systems. Although GeoThms [13] integrates GCLC with a repository of theorem statements, their proofs (generated automatically), and the corresponding illustrations in a web interface, relations among the geometric problems in the database have not been considered.

When speaking about managing knowledge, we may first think about textbooks where knowledge is represented systematically and hierarchically according to certain logical structures, e.g., from the simplest to the most complicated and from the basic to the advanced. Due to the well-organized structure of the domain knowledge, textbooks play an important role in education and research to store, manage, and impart knowledge for new learners. Therefore, constructing dynamic, interactive, and machine-processable textbooks (instead of traditional static textbooks) is an interesting project of research and development. The Electronic Geometry Textbook (EGT) aims at providing such a tool by integrating elementary plane geometric knowledge and software modules into a single computing environment to support geometry education, research, and application. The idea of developing such an integrated software system in the form of a textbook for systematical and efficient management of geometric knowledge originates from the second author who has been working on automated geometric reasoning in the last two decades. The first author has been stimulated to elaborate the idea and to carry out the actual implementation. The start of this project has also been motivated considerably by the work of Zeilberger [17] and his invited talk at ADG 2004.

The purpose of this paper is to propose a new style of knowledge management and to present the design and preliminary implementation of a software system in the form of a dynamic textbook for managing geometric knowledge interactively: the system can be viewed or printed as a traditional textbook and can run as dynamic software on computer. The knowledge in the textbook, being formalized by using standard formal languages, can be processed by software modules developed for geometric computing and reasoning, diagram generation, and visualization and can be enriched systematically by "self-learning". The textbook may be generated automatically by organizing and presenting the textbook data according to the user's specifications. The system allows the user to manipulate (search, modify, restructure, etc.) the textbook with automated consistency checking (of its logical structure). We will present the main ideas on the design of the electronic geometry textbook, explain the features of the system, propose five phases of creating and managing the geometric knowledge in the textbook, discuss the involved tasks and some of the fundamental research problems in each phase, and report our progress and experiments on a preliminary implementation of the system. As the implementation is still at an early stage, our emphasis here is placed mainly on the design methodology.

While plane Euclidean geometry is the target of our current investigation, the basic principles and ideas discussed in this paper apply to any geometry or even any subject of mathematics.

2 Objective and Design Methodology

In this section, we provide a short review on the state of the art of mathematical knowledge management, describe the main objective of our EGT project, and discuss our methodology for the EGT database and system design.

2.1 Mathematical Knowledge Management

Currently mathematical knowledge is archived and stored mainly in printed and/or electronic documents, like books and articles. It is presented in a static way and searching for items in such documents can be done only syntactically. The structure of the documents cannot be changed and the knowledge therein cannot be processed by problem solvers, theorem provers, or symbolic calculators.

As Internet has become a major channel for information service, various efforts have been devoted to making mathematical knowledge available on the Internet and exchangeable among different software programs. Some markup languages such as MathML [16] have been developed, making it possible to display and communicate mathematical formulas on the web. As an application of XML, MathML benefits from the tools existing for XML file manipulation. Although it does offer some semantics for symbols in mathematical formulas, the set of symbols supported, when compared to those used by working mathematicians, is very restricted. To ameliorate this situation, projects like OpenMath [4] and OMDoc [8] have emerged. The OpenMath standard focuses on describing the semantic meanings of mathematical objects instead of their appearance by using *content dictionaries*, in which mathematical symbols are defined syntactically and semantically and thus allowed to be exchanged between computer programs. Content dictionaries can be stored in databases or published on the world-wide web. OMDoc is an extension to the OpenMath standard by markup for the structure and the theory level of mathematical documents, adding capabilities of describing the mathematical context of the used OpenMath objects. These languages make mathematical knowledge not only machine-readable but also machine-understandable and provide a foundation for developing, communicating, and publishing mathematics on the web.

Mathematical software systems for symbolic and numeric computation, formal reasoning, proof checking, algorithm verification, etc. need domain knowledge to support relevant (automated) activities. Some systems, such as Mizar [14] whose objective focuses on writing and checking formal mathematics written in the Mizar language, provide mathematical knowledge bases (that contain a large amount of domain knowledge in different mathematical theories) and facilities to browse formal mathematical documents. Other systems like Theorema [12] provide environments for building mathematical knowledge bases in a systematical way and such bases can be browsed, extended, and used for teaching, learning, and mathematical discovering. Although mathematical contents embedded in such documents can be processed inside the systems, the documents lack the characteristics of traditional textbooks: mathematical contents are not presented in natural style, the structures of the mathematical documents are static and passive, the logical relations of knowledge are not considered as in traditional textbooks, and no tool is provided to construct, manipulate, and present these documents effectively.

There are few facilities for data exchange among the systems mentioned above and it is very difficult to share, reuse, and interact on domain knowledge resources. Repeated developments are waste of efforts, time, and energy. To integrate system resources, some mathematical knowledge engineering projects, including MBase [9], HELM [3], the Formal Digital Library [2], and the NIST Digital Library of Mathematical Functions [11], have emerged, aiming at building general mathematical knowledge bases for retrieving, representing, acquiring, and reusing various kinds of mathematical domain knowledge on the web. As the knowledge in the bases is represented at the sematic level, it is possible to make the knowledge bases serve for different levels of need, e.g., automated theorem proving, automated programming, and mathematical education. Some projects make use of web-based semantic knowledge bases to develop mathematical intelligent education environments such as ActiveMath [1] which is designed for students to proceed with e-learning. The courseware (textbook) is produced automatically by weighing the student's preferences. Different preferences may lead to different textbook configurations, but manipulating (such as reconstructing or modifying) the textbook is not allowed.

EGT will synthesize the functionality of document creation and manipulation together with automated processing of mathematical contents in the document. Similar to ActiveMath, EGT is also based on a formalized knowledge base in which mathematical contents can be easily converted to the internal representations of other application software. However, EGT is designed mainly as a tool for human authors to construct their own textbooks with automated verification of structural consistency. The process of producing the textbook is mostly human-driven and manipulations on the textbook are allowed and may lead to new, modified, or improved versions of the textbook.

2.2 Objective and System Description

As one of the most fundamental and oldest subjects of mathematics, geometry is founded on graphical objects abstracted from the real visual world. Geometric computation, reasoning, and visualization require the support and integration of logical deduction mechanisms, effective algebraic methods, and graphical drawing tools, involving both abstract quantities and intuitive figures. Moreover, there is no other mathematical subject than geometry in which there are so many highly interesting and fascinating theorems and such theorems may be proved automatically on computer. In fact, automated deduction is much more mature and successful in geometry than in any other domain of mathematics. When mathematical knowledge is organized in a textbook, it is important to track the logical clues of the knowledge, such as how a concept is introduced and how a theorem is used in the proof of another theorem. The availability of automated reasoning devices is a prerequisite for knowledge organization.

Therefore, we consider geometry a unique and rich subject of mathematics that should be chosen for study in the context of knowledge management. In this study, the full power of computer for symbolic, numeric, and graphical computing and data processing may be used and our idea and methodology may be effectively tested.

The objective of our textbook project is to provide a tool for human authors to construct dynamic, interactive textbooks (see Fig. 1).

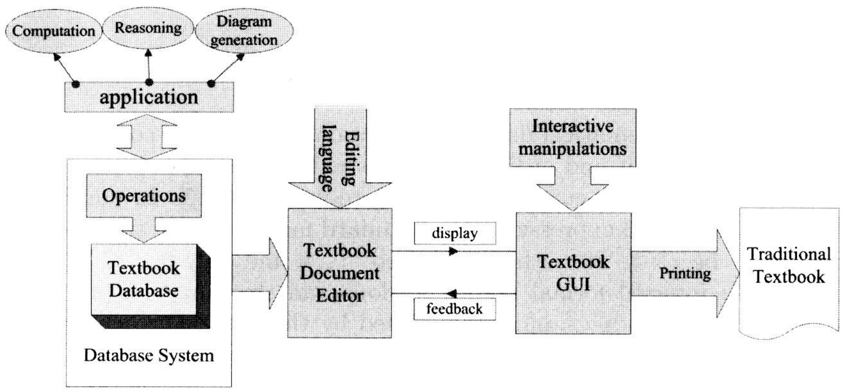

Fig. 1. System Framework

- The formulation and representation of geometric knowledge and the structure of the textbook should be *standardized*, so that the knowledge, when digitalized, can be processed with little extra effort by different software modules existing or being developed for geometric computing and reasoning, interactive or automated dynamic diagram generation, and visualization of geometric objects and the results produced by those modules can be easily integrated into and displayed in the textbook.
- Under this standardization, a *textbook database* will be built, collecting textbook data which refer to as formalized geometric knowledge and related contents (such as literature information, background of remarkable theorems, and explanation texts), and new knowledge may be incorporated properly and automatically into the textbook database by implementing a self-learning and extracting mechanism. The purpose of building such a textbook database is to provide a dataset for reusing and sharing textbook data and semantic information for various applications.
- The user will be able to *edit* his/her own *textbook documents* in an editing language of "mixture" style.
- A *dynamic* textbook graphical user interface (GUI) can be created automatically to display the textbook document by translating formalized geometric statements into natural languages in traditional style automatically.

– The user is permitted to *manipulate* (such as search, modify, and restructure) the textbook or parts thereof and process the knowledge (such as proving theorems, performing computations, and drawing diagrams) in a visual environment using available software modules.
– The system will *verify* some geometric statements by employing other available software modules and *check* some structural consistency of the constructed document in the sense of textbook.

2.3 Design Methodology

The creation and management of the textbook can be divided into the following five phases.

Classification. Mathematical documents contain two kinds of contents, mathematical expressions and ordinary texts, which are mixed and not separated in commonly used document processing tools like LaTeX. In our EGT system, mathematical expressions and standard mathematical statements such as "let ... be ...", "if ... then ...", and "... if and only if ..." are formalized (with internal formal representations and thus may be processed by other software modules and manipulated by the user). Ordinary (explanation) texts are clearly identified. We standardize the structure of textbooks by classifying textbook contents with hierarchy and provide the user with standard *templates* for uploading the contents. The classification of geometric objects, concepts, and knowledge "segments" is also a prerequisite for the formalization of geometric knowledge. We shall discuss the classification of textbook contents and geometric knowledge in Section 3.1.

Organization. After textbook contents are classified into parts, our problem is to assemble the chosen "parts" in a reasonable "order" for a concrete textbook. Although our EGT system is human-driven (i.e., how to organize the textbook is decided by the user), the system is capable of verifying the suitability and consistency of the logical structure of the textbook according to some characteristic rules on the structure of standard textbooks and providing the user with suggestions when such rules are violated. The reader will find more details in Section 3.2.

Representation. The above two phases are mainly concerned about the macro-structural aspects of the textbook. As for the contents of each "template", two types of information need be provided: the information for displaying the contents in natural style and the underlying formal representation for interacting with other modules. How to define the two types of information, in particular the semantic formal representation of knowledge and the information needed for various applications? Several data definition models will be described in Section 3.3.

Manipulation. Based on the classification and representation, a language for editing textbook documents has to be developed. How to present formal data in natural style and how a user interacts with the GUI? We propose a kind of language in a mixture style to edit textbook documents with possible