A Database Schema Integration Methodology for Data Warehouse

陳明毅、黃士銘

E-mail: 9121492@mail.dyu.edu.tw

ABSTRACT

Designing a data warehouse system needs to integrate distributed/heterogeneous databases into a reconciled data platform for strategic use. To establish the reconciled data platform, a global schema or global data views which contain higher semantic and integrity constraints are required during the development phase. In this dissertation, we describe a novel methodology to integrate independent local logical database schemas into a global semantic database schema. These schemas are converted from the local databases to Extended Entity Relationship (EER) models. Equivalency of domain, attribute, entity, and relationship are identified between two of these EER models. While a discovery and resolution mechanism by using these equivalent definitions is applied to deal with the naming conflict and structural conflict between two EER models, a reconciled semantic view between the two EER models is established. A semantic merge mechanism by using data analysis technique is also applied to extract more semantics during the merge phase. Two local database schemas are then integrated. These steps are repeated until the schemas of all databases to be integrated have been consolidated into a single global schema.

Keywords : data warehouse ; database schema ; scheam integration

Table of Contents

REFERENCES

1] Berson, A., and Smith, S. J., “ Data Warehousing, Data Mining, and OLAP, ” McGraw-Hill, pp. 14-21, 1997.

[2] S. B. Navathe and S. G. Gadgil, “ A methodology for view integration in logical data base design, ” in Proc. 8th Int. Conf. Very Large Data Bases, Mexico City, Sept. 1982 pp. 142-152.

[3] C. Batini, M. Lenzerini, and S. B. Navathe, “ A Comparative Analysis of Methodologies for Database Schema Integration ”, ACM Computing Surveys, VOL. 18, NO. 4, pp.323-364, Dec. 1986.

[4] Tse-Min Hung, “ A Novel Data Warehouse Architecture: A Database Proxy Server Approach, ” Master Thesis, Department of Computer Science and Engineering, Tatung Institute of Technology, June 1999.
[5] Philippe-Pierre Dornier, Ricardo Ernst, Michel fender, and Panos Kouvelis, “ Global Operations and Logostics, ” John Wiley & Sons, Inc., ISBN0-471-12036-7, pp. 1-5, 1998.
[6] C.Palvia Prashant, “ Research Issues in Global Information Technology Management, ” Information Resource Management Journal, Vol. 11, No. 2, 1998.
[7] Larson, J. A, Navathe, S. B., and Elmasri, R., “ A Theory of Attribute Equivalence in Databases with Application to Schema Integration, ” IEEE Transaction on Software Engineering, VOL. 15, NO. 4, pp. 449-463, Apr. 1989.
[8] Gengo Suzuki, Masashi Yamamuro, “ Schema Integration Methodology Including Structural Conflict Resolution and Checking Conceptual Similarity, ” Database Reengineering and Interoperability, Plenum Press, New York, pp. 247-260, 1996.
[9] SHOVAL, P. and ZOHN, A., “ Binary-relationship integration methodology ”, Data & Knowledge Engineering, VOL. 6, pp.225- 250, 1996.
[10] Ernst, R., Fender, M. and Kouvelis, P., “ Global Operations and Logistics Text and Cases, ” John Wiley & Sons, pp. 1-5, INC. 1998.
[11] Stefano Saccapietra, Christine Parent, and Yann Dupont, “ Model Independent Assertions for Integration of Heterogeneous Schemas ”, The VLDB Journal, Vol. 1, No 1, pp. 81-126, July 1992.
[12] Omran A. Bukhres and Ahmed K. Elmagarmid, “ Object- Oriented Multidatabase Systems A Solution for Advanced Applications, ” Prentice Hall, pp. 105-202, 1996.
[13] Giuseppe Santucci, “ Semantic schema refinements for multilevel schema integration, ” Data & Knowledge Engineering 25, pp. 301-326, 1998.
[14] Li S. H., Huang S. M., and Chen H. H., “ Discover Missing Semantic from Existing Relational Databases, ” 8th International Database Workshop, Hong Kong, pp. 275-286, 1997.
[15] Shing-Han Li, “ Translating A Relational Database Schema into An Extended Entity-Relationship Database Schema: A Data Mining and Data Dictionary Approach, ” Master Thesis, Department of Computer Science and Engineering, Tatung Institute of Technology, Jun. 1997.
[16] Joseph Fong and Shi-Ming Huang, “ Architecture of a Universal Database: A Frame Model Approach, ” The International Journal of Cooperative Information System, World Scientific Pub., Jun. 1999.
[17] Joseph Fong and Shi-Ming Huang, “ Information Systems Reegineering, ” Springer-Verlag Singapore Pte. Ltd., 1997.