Ana Iglesias, Paloma Martínez, Dolores Cuadra, Elena Castro, Fernando Fernández, Learning to teach database desing by trial and error. En: José Braz, Mario Piattini, Joaquim Filipe (eds.) ICEIS 2002 Proceedings: Fourth International Conference on Enterprise Information Systems: Ciudad Real, Spain, April 3-6, 2002. Volume 1. Portugal: ICEIS Press, 2002, pp. 500-505. ISBN 972-98050-6-7

International Conference on Enterprise Information Systems

Ciudad Real / 2002

Spain

ICEIS 2002

FOURTH INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS

Proceedings

Volume 1

CIUDAD REAL, SPAIN · APRIL 3-6, 2002

HOSTED BY
THE SCHOOL OF INFORMATICS OF THE
UNIVERSITY OF CASTILLA-LA MANCHA

UCLM
UNIVERSIDAD DE CASTILLA-LA MANCHA

CO-ORGANIZED BY
THE SCHOOL OF TECHNOLOGY
OF SETÚBAL

Escola Superior Tecnologia
SETÚBAL

IN COLLABORATION WITH
AAAI

Edited by José Braz, Mario Piattini and Joaquim Filipe

ISBN: 972-98050-6-7

Depósito Legal Número: 176839/02

(Volume 1 and Volume 2)

Printed in Portugal

ICEIS Press / Escola Superior de Tecnologia de Setúbal

Campus do Instituto Politécnico de Setúbal

Rua do Vale de Chaves, Estefanilha

2914-508 Setúbal

produced using laser printers from XEROX and XETCOPY

THE DOCUMENT COMPANY

XEROX®

POSTERS

dblp.uni-trier.de
Computer Science Bibliography

Universität Trier

4. ICEIS 2002: Ciudad Real, Spain

ICEIS 2002, Proceedings of the 4st International Conference on Enterprise Information Systems, Ciudad Real, Spain, April 2-6, 2002

VOLUME I

Invited Speakers

Keynote Lectures

Tutorials

Case Studies

Databases and Information Systems Integration

Papers

- Ikram Amous, Florence Sedes: Descriptors and Meta-Documents for Mono-Media and Multimedia Documents. 11-17
- Ikram Amous, Anis Jedidi, Florence Sedes: Organising and Modelling Metadata for Media-Based Documents. 18-25
- Wilmondes Manzi de Arantes Júnior, Christine Verdier, André Flory: XML-Based Document to Query a Relational Database. 26-33
- Ladjane S. Arruda, Cláudio de Souza Baptista, Carlos A. A. Lima: MEDIweb: A Mediator-Based Environment for Data Integration on the Web. 34-41
- Thomas Birkhölzer, Jürgen Vaupel: The Role of Enterprise Architecture for Planning and Managing Future Information Systems Integration. 42-49
- Ramón Alberto Carrasco, María Amparo Vila Miranda, José Galindo: FSQL: A Flexible Query Language for Data Mining. 50-56
- Pauline Lienhua Chou, Zahir Tari: Predicate-Based Caching Scheme for Wireless Environments. 57-64
- Michael Christoffel, Bethina Schmitt, Jürgen Schneider: Semi-Automatic Wrapper Generation and Adaption: Living with Heterogeneity in a Market Environment. 65-72
- Carmen Constantinescu, Uwe Heinkel, Ralf Rantzau, Bernhard Mitschang: A System for Data Change Propagation in Heterogeneous Information Systems. 73-80
- Johann Eder, Christian Koncilia, Herbert Kogler: Temporal Data Warehousing: Business Cases and Solutions. 81-88
- José Esteves, Joan Antoni Pastor: A Framework to Analyse Most Critical Work Packages in ERP Implementation Projects. 89-98
- Erik Gyllenswärd, Mladen Kap, Rikard Land: Information Organizer: A Comprehensive View on Reuse. 99-106
- Arne Harren, Heiko Tapken: A Process Model for Enterprise-Wide Design of Data Acquisition for Data Warehousing. 107-114
- Mónica Larre, José Torres, Eduardo F. Morales, Sócrates Torres: Data Integration Using the MONIL Language. 115-120
- Wai Lup Low, Joseph Lee, Peter Teoh: DIDAFIT: Detecting Intrusions in Databases Through Fingerprinting Transactions. 121-128
- Farhi Marir, Kamel Zerzour, Karim Ouazzane: An Integrated Object Database and Description Logic System for Online Content and Event-Based Indexing and Retrieval of a Car Park Surveillance Video. 129-139
- Alberto Pan, Paula Montoto, Anastasio Molano, Manuel Álvarez, Juan Raposo, Ángel Viña: A Model for Advanced Query Capability Description in Mediator Systems. 140-147
- Valéria Magalhães Pequeno, Vânia Maria Ponte Vidal: Using Full Match Classes for Self-Maintenance of Mediated Views. 148-154
- Alexander Prosser: Proposing a Method for Planning the Materialisation of Views in a Data Warehouse. 155-162
- Claudia Raibulet, Claudio Demartini: Data Representation in Industrial Systems. 163-178
- Bruno Sadeg, Samia Saad-Bouzefrane, Laurent Amanton: D-ANTICIP: A Protocol Suitable for Distributed Real-Time Transactions. 171-178
- Daniel Sánchez, José-María Serrano, María Amparo Vila Miranda, V. Aranda, J. Calero, G. Delgado: Using Data Mining Techniques to Analyze Correspondences between Partitions. 179-186
- Rebecca Boon-Noi Tan, Guojun Lu: A Hierarchical Approach to Complex Data Cube Queries. 187-194
- Yauheni Veryha: Implementation of Fuzzy Classification Query Language in Relational Databases Using Stored Procedures. 195-202
- Zhang Xiaoou, Hung Keng Pung: An XML-Based Virtual Patient Records System for Healthcare Enterprises. 203-209

- Ale Gicqueau: Importing XML Documents to Relational Databases. 210-217
- Goce Trajcevski, Ouri Wolfson, Hu Cao, Hai Lin, Fengli Zhang, Naphtali Rishe: Managing Uncertain Trajectories of Moving Objects with Domino. 218-225

Short Papers

- Muhammad Abaidullah Anwar, S. Hameed: An Integrated Approach for Finding enRoute Best Alternate Route. 226-234
- Joseph Barjis, Samuel Chong: Data Modeling for the Purpose of Database Design Using Entity-Relationship Model and Semantic Analysis. 235-243
- Peter Bodorik, Shawn Best, Dawn N. Jutla: Toolkit for QoS Monitoring in Middleware. 244-249
- Miguel Calejo, Mário Araújo, Sónia Mota Araújo, Nuno Soares: Web Application Maker. 250-256
- Mary Garvey, Mike Jackson, Martin Roberts: Using Persistent Java to Construct a GIS. 257-262
- Oleg Yu. Gusikhin, Erica Klampfl, Giuseppe Rossi, Celestine Aguwa, Gene Coffman, Terry Marinak: Virtual Reality Web-Based Environment for Workcell Planning in an Automotive Assembly. 263-268
- Sai Peck Lee, Chin Heong Khor: Persistence Framework for Multiple Legacy Databases. 269-274
- Steve C. A. Peters: Introducing an Enterprise Resource Planning (ERP) System in a Hospital. 275-281
- Yoshiyuki Shinkawa, Masao J. Matsumoto: State-Sensitive Design of Database Transactions. 282-289
- Zahir Tari, Abdelkamel Tari, Vincent Dupin: Designing an Object and Query Cache Management System for CODAR Database Adapter. 290-295
- Jari Töyli, Matti Linna, Merja Wanne: Modeling Relational Data by the Adjacency Model. 296-301
- Frank Wang, Ruby Sharma, Na Helian, Farhi Marir, Yau Jim Yip: The Millennium Information System (MIS) for European Public Health and Environment Network (EPHEN). 302-306

Posters

- Laura M. Bartolo, Austin Melton, Monica Strah, Cathy S. Lowe, Louis Z. Feng, Christopher Woolverton: An Electronic Scientific Notebook: Metadata And Database Design for Multidisciplinary Scientific Laboratory Data. 307-310
- Matthew C. F. Lau, Rebecca Boon-Noi Tan: The Impact of Change on IT Professionals. 311-316

Artificial Intelligence and Decision Support Systems

Papers

- Pedro J. Abad, Antonio J. Suárez, Sixto Romero, Juan Antonio Ortega: Semiqualitative Reasoning for Software Development Project by Constraint Programming. 319-324
- Rafael Alcalá, Oscar Cordón, Francisco Herrera, Igor Zwir: Insurance Market Risk Modeling with Hierarchical Fuzzy Rule Based Systems. 325-333
- Li Bai, Yihui Liu: Neural Networks and Wavelets for Face Recognition. 334-340
- Larry Y. C. Cheung, Paul W. H. Chung, Ray Dawson: Supporting Engineering Design Process with

an Intelligent Compliance Agent: A Way to Ensure a Standard Complied Process. 341-349

- M. Julia Flores, José A. Gámez: Applicability of Estimation of Distribution Algorithms to the Fuzzy Rule Learning Problem: A Preliminary Study. 350-357
- Francisco Herrera, Luis Martínez: Group Decision Making Based on the Linguistic 2-Tuple Model in Heterogeneous Contexts. 358-366
- Christian Hillbrand, Dimitris Karagiannis: Using Artificial Neural Networks to Prove Hypothetic Cause-And-Effect Relations: A Metamodel-Based Approach to Support Strategic Decisions. 367-373
- Shonali Krishnaswamy, Seng Wai Loke, Arkady B. Zaslavsky: Supporting the Optimisation of Distributed Data Mining by Predicting Application Run Times. 374-381
- Raquel Flórez López: Strategic Position of Firms in Terms of Client's Needs Using Linguistic and Numerical Information through a New Model of SOFM. 382-390
- Farhi Marir, Frank Wang, Karim Ouazzane: A Case-Based Expert System for Estimating the Cost of Refurbishing Construction Buildings. 391-398
- I-Heng Meng, Wei-Pang Yang, Wen-Chih Chen, Lu-Ping Chang: Data Mining Mechanisms in Knowledge Management System. 399-404
- Md. R-Moreno, Manuel Prieto, D. Daniel Meziat Luna, J. Medina, C. Martin: Controlling and Testing a Space Instrument by an AI Planner. 405-409
- Jim R. Oliver: A Training Environment for Automated Sales Agents to Learn Negotiation Strategies. 410-417
- Abdel-Badeeh M. Salem, Taha ElAreef, Marwa F. Khater, Aboul Ella Hassanien: A Density-Based Approach for Clustering Spatial Database. 418-423
- Jose Jesus Castro-Schez, Luis Jiménez, Juan Moreno García, L. R. Luis Rodriguez Benitez: Simple Decision Support System Based on Fuzzy Repertory Table. 424-430
- Aïda Valls, Antonio Moreno, David Sánchez: A Multi-Criteria Decision Aid Agent Applied to the Selection of the Best Receiver in a Transplant. 431-438
- Alfredo Vellido: Neural Networks for B2C E-Commerce Analysis: Some Elements of Best Practice. 439-445
- Stijn Viaene, Bart Baesens, Guido Dedene, Jan Vanthienen, Dirk Van den Poel: Proof Running Two State-Of-The-Art Pattern Recognition Techniques in the Field of Direct Marketing. 446-454

Short Papers

- Lucimar F. de Carvalho, Candice Abella S. Dani, Hugo J. T. de Carvalho, Diego Dozza, Silvia M. Nassar, Fernando Mendes de Azevedo: Medical Data Base Exploration through Artificial Neural Networks. 455-459
- Carlos J. Costa, Pedro Antunes: Evaluating EMS Value - The Case of a Small Accountancy Firm. 460-466
- Monica Dascalu, Sergiu Goschin, Eduard Franti: Using Cellular Automata in Traffic Modeling. 467-472
- Sabine Delaître, Alain Giboin, Sabine Moisan: The AEX Method and Its Instrumentation. 473-481
- Alberto Díaz Esteban, Pablo Gervás, Antonio García Jiménez: Improving Access to Multilingual Enterprise Information Systems with User Modelling Context Enriched Cross-Language IR. 482-487
- Tarek F. Gharib, Mostafa G. Mostafa, Mohamed F. Tolba: High Dimensional Data Clustering Using SOFM And K-Means Algorithms. 488-493
- Melanie Gnasa, Jens Woch: Natural Language Interface to Knowledge Management Systems. 494-499
- Ana Iglesias, Paloma Martínez, Dolores Cuadra, Elena Castro, Fernando Fernández: Learning to

Teach Database Design by Trial and Error. 500-505

- S. M. Michael, Deepak Khemani: Knowledge Management in Manufacturing Technology: An A.I. Application in the Industry. 506-511
- Manuel Filipe Santos, José Neves, António Abelha, Álvaro M. Silva, Fernando Rua: Augmented Data Mining over Clinical Databases Using Learning Classifier Systems. 512-516
- Min-Jung Yoo, Jean-Pierre Müller: Using Multi-Agent System for Dynamic Job Shop Scheduling. 517-525

Posters

- Leszek Kieltyka, Robert Kuceba: The Application of Artificial Neural Networks for Heat Energy Use Prediction. 526-529
- Su Linying, Bernadette Sharp, Claude C. Chibelushi: Knowledge-Based Image Understanding: A Rule-Based Production System for X-Ray Segmentation. 530-533
- Ling Tan, David Taniar, Kate A. Smith: A Taxonomy for Inter-Model Parallelism in High Performance Data Mining. 534-539

VOLUME II

Information Systems Analysis and Specification

Papers

- Ana Almeida, Licinio Roque: Some Reflections on IS Development as Operator of Organisational Change. 543-552
- Jorge Luis Nicolas Audy: Analysis of the Relation between the Prescriptive and Descriptive Approaches of the Information System Planning. 553-561
- Maria Cecília Calani Baranauskas, Juliana P. Salles, Kecheng Liu: Analysing Communication in the Context of a Software Production Organisation. 562-569
- Sergio de Cesare, Mark Lycett, Dilip Patel: Business Modelling with UML: Distilling Directions for Future Research. 570-579
- Mohamed Dahchour, Alain Pirotte: The Semantics of Reifying n-ary Relationships as Classes. 580-586
- H. Kadri-Dahmani, A. Osmani: Updating Data in GIS: How to Maintain Database Consistency? 587-593
- Ivan Mathias Filho, Toacy Cavalcante de Oliveira, Carlos José Pereira de Lucena: A Proposal for the Incorporation of the Features Model into the UML Language. 594-601
- Alexandra Galatescu, Taisia Greceanu: Ontologies Supporting Business Process Re-Engineering. 602-609
- Félix García, Francisco Ruiz, Mario Piattini, Macario Polo: Conceptual Architecture for the Assessme NT and Improvement of Software Maintenance. 610-617
- Dániel Hoványi: Reusable Component Extraction from Intellingent Network Management Applications. 618-626
- Zaijun Hu: Semantic Augmentation through Ontology for XML Integration Server. 627-634
- Henry M. Kim, K. Donald Tham: Designing Business Processes and Communication Structures for E-Business Using Ontology-Based Enterprise Models with Mathematical Models. 635-641

- Juan de Lara, Hans Vangheluwe: Using AToM^3 as a Meta-Case Tool. 642-649
- Toacy Cavalcante de Oliveira, Ivan Mathias Filho, Carlos José Pereira de Lucena: Frameworks - A High Level Instantiation Approach. 650-657
- Vicente Pelechano, Manoli Albert, Eva Campos, Oscar Pastor: Automating the Code Generation of Role Classes in OO Conceptual Schemas. 658-666
- Geert Poels: A Functional Size Measurement Method for Event-Based Objectoriented Enterprise Models. 667-675
- Licinio Roque, Ana Almeida: The Context Engineering Approach. 676-682
- Monique Snoeck: Sequence Constraints in Business Modelling and Business Process Modelling. 683-690
- Sibylle Steinau, Oscar Díaz, Juan J. Rodríguez, Felipe Ibáñez: A Tool for Assessing the Consistency of Websites. 691-698
- Juan Trujillo, Sergio Luján-Mora, Enrique Medina: The Gold Model Case Tool: An Environment for Designing OLAP Applications. 699-707
- Mohamed Zairi, Majed Al-Mashari: An International Study of Benchmarking Spread and Maturity. 708-715
- Ilham Alloui, Sorana Cîmpan, Flávio Oquendo: Taming Process Deviations by Logic Based Monitoring. 716-724

Short Papers

- Akihiro Abe: Applying Domain Modeling and SECI Theory in Knowledge Management for Information Systems Analysis. 725-731
- Ilia Bider, Maxim Khomyakov: If You Wish to Change the World, Start with Yourself. 732-742
- Guido Dedene: On the Use of Jackson Structured Programming (JSP) for the Structured Design of XSL Transformations. 743-751
- S. Hanlon, L. Sun: A Framework for the Dynamic Alignment of Strategies. 752-757
- Cirano Iochpe, Lucinéia Heloisa Thom: Inferring Aspects of the Organizational Structure through Workflow Process Analysis. 758-763
- Huiqing Liu, Tecksin Lim: A Knowledge Object Oriented System for High Throughput Collection and Analysis of Data. 764-769
- Juan M. Luzuriaga, Rodolfo Martínez, Alejandra Cechich: Managing Enterprise Communication Networks to Improve the Requirements Elicitation Process. 770-775
- Mara Nikolaidou, Dimosthenis Anagnostopoulos: Introducing Business Process Automation in Dynamic Banking Activities. 776-781
- Almudena Sierra-Alonso, Paloma Cáceres, Esperanza Marcos, Jorge Enrique Pérez-Martínez: Incorporating Knowledge Engineering Techniques to Requirements Capture in the MIDAS Web Applications Development Process. 782-787
- Slim Turki, Michel Léonard: Hyperclasses. 788-794
- Hans Weghorn, Carolin Gaum, Daniel Wloczka: Linking Mobile Network Services to Internet Mail. 795-798

Posters

- Tauqir Amin, Hung Keng Pung: Inter-Organizational Workflow Management in Virtual Healthcare Enterprises. 799-802
- Nuno Castela, José M. Tribolet, Arminda Guerra, Eurico R. Lopes: Survey, Analysis and Validation of Information for Business Process Modeling. 803-806
- Manuel Enciso, Angel Mora: FD3: A Functional Dependencies Data Dictionary. 807-810
- Javier Garzás, Mario Piattini: Beyond Object Oriented Design Patterns. 811-814

- Luis A. Guerrero, Sergio F. Ochoa, Oriel A. Herrera, David A. Fuller: Mediated Communication in Groupware Systems. 815-818 BibTeX ML
- Harith T. Al-Jumaily, Dolores Cuadra, Paloma Martínez: An Execution Model for Preserving Cardinality Constraints in the Relational Model. 819-822 BibTeX ML
- Takaaki Kamogawa, Masao J. Matsumoto: Towards a New Business Process Architecture. 823-829 BibTeX ML
- Rafael Prikladnicki, Fernando Peres, Jorge Luis Nicolas Audy, Michael da Costa Móra, Antônio Perdigoto: Requirements Specification Model in a Software Development Process inside a Physically Distributed Environment. 830-834 BibTeX ML
- Leonilde Reis, Luis Alfredo Martins do Amaral: Integrated Planning of Information Systems and Contingency and Recovery. 835-837 BibTeX ML
- Manuela Rodríguez-Luna: Stemming Process in Spanish Words with the Successor Variety Method. Methodology and Result. 838-842 BibTeX ML
- Khodakaram Salimifard, Mike Wright: Modelling and Performance Analysis of Workflow Management Systems Using Timed Hierarchical Coloured Petri Nets. 843-846 BibTeX ML
- Andy Salter, Kecheng Liu: Using Semantic Analysis and Norm Analysis to Model Organisations. 847-850 BibTeX ML
- Ryo Sato, Kentaro Hori: Developing Quick Iterative Process Prototyping for Project Management: Linking ERP and BPE. 851-854 BibTeX ML
- Khalid Suliman Al-Tahat, Sufian Bin Idris, Tengku M. T. Sembok, Mohamed Yousof: Using Hot-Spot-Driven Approach in the Development of a Framework for Multimedia Presentation on the Web. 855-860 BibTeX ML

Internet Computing and Electronic Commerce

Papers

- Abraham Alvarez, Youssef Amghar, Richard Chbeir: Managing XML-Link Integrity for Structured Technical Documents. 863-870 BibTeX ML
- Marta Fernández de Arriba, José Antonio López Brugos: Usability and Accessibility in the Specification of Web Sites. 871-876 BibTeX ML
- Colin G. Ash, Janice M. Burn: Staged Implementation of E-Business Networks through ERP. 877-884 BibTeX ML
- Meletis A. Belsis, Anthony N. Godwin, Leonid Smalov: An Enterprise IT Security Data Model. 885-891 BibTeX ML
- Sebastián Bruque Cámara: Internet Technology as a Business Tool. 892-899 BibTeX ML
- Magdalena Payeras-Capellà, Josep Lluís Ferrer-Gomila, Llorenç Huguet i Rotger: ITHAKI: Fair N-Transferable Anonymous Cash. 900-907 BibTeX ML
- M. F. Chen, M. K. Shan: Virtual Mall of E-Commerce Web Sites. 908-914 BibTeX ML
- Michael S. H. Heng: Rethinking the Strategy of Amazon.com. 915-921 BibTeX ML
- I. Hoyle, L. Sun, Stephen J. Rees: The Design of an XML E-Business Applications Framework. 922-928 BibTeX ML
- Hamid Jahankhani, Solomon A. Alexis: E-Commerce Business Practices in the EU. 929-936 BibTeX ML
- Oliver Kelkar, Jörg Leukel, Volker Schmitz: Towards Extended Price Models in XML Standards for Electronic Product Catalogs. 937-945 BibTeX ML
- Juan de Lara, Manuel Alfonseca: Hierarchical Visualization in a Simulation-Based Educational Multimedia Web System. 946-953 BibTeX ML
- Andy Laws, M. Allen, A. Taleb-Bendiab: Normative Services for Self-Adaptive Software to Support Dependable Enterprise Information Systems. 954-961 BibTeX ML
- Kanta Matsuura, Hideki Imai: Digital Timestamps for Dispute Settlement in Electronic Commerce:

Generation, Verification, and Renewal. 962-967

- Maurizio Panti, Luca Spalazzi, Simone Tacconi, Salvatore Valenti: Automatic Verification of Security in Payment Protocols for Electronic Commerce. 968-974
- Nathalie Petit: A Method for WIS Centered on Users Goals. 975-980
- Aleksander Pivk, Matjaz Gams: A Semi-Universal E-Commerce Agent. 981-988
- Antonio Reyes-Moro, Juan A. Rodríguez-Aguilar, Maite López-Sánchez, Jesús Cerquides, D. Gutierrez-Magallanes: *Quotes*: A Negotiation Tool for Industrial E-Procurement. 989-994
- Antonio Ruiz Cortés, Rafael Corchuelo, Amador Durán: An Automated Approach to Quality-Aware Web Applications. 995-1000
- Carlos Serrão, Joaquim Marques: IM@GIX: An Electronic Commerce Digital Image Bank. 1001-1008
- Markus Sihvonen: Profile Negotiation Requirements in a Mobile Middleware Service Environment. 1009-1015
- Qiubang Li, Rajiv Khosla: Intelligent Agent-Based Framework for Mining Customer Buying Habits in E-Commerce. 1016-1022

Short Papers

- Trevor Burbridge, Jonathan Mitchener, Ben Strulo: The "Shared Data Approach" to the Integration of Dynamic Business Enterprises. 1023-1028
- Anastasia Constantinou, Vassilios Tsakalos, Philippos Koutsakas, Dimitrios Tektonidis, Adamantios Koumpis: An E-Service Infrastructure for Innovation Exploitation and Technology Transfer: The DILEMMA Project. 1029-1034
- Francisco José García Peñalvo, I. Borrego, M. J. Hernández, Ana Belén Gil, Miguel A. Laguna: An E-Commerce Model for Small and Medium Enterprises. 1035-1041
- Hamid Jahankhani, Pari Jahankhani: Distributed Online Doctor Surgery. 1042-1047
- Qusay H. Mahmoud, Luminita Vasiu: Accessing and Using Internet Services from Java-Enabled Handheld Wireless Devices. 1048-1053
- Mike Rogers, Thomas Chesney, Scott Raeburn: E-Procurement in a Rural Area. 1054-1057
- Irene Luque Ruiz, Enrique López Espinosa, Gonzalo Cerruela García, Miguel Ángel Gómez-Nieto: A System Based on Preferences for Aid to the Purchase Decision. 1058-1063
- Eun-Jung Song, Ho-Song Lee, Taeck-Geun Kwon: Design and Implementation of a Message Service Handler for EBXML. 1064-1069
- Luminita Vasiu, Alan Murphy: Design Requirements for Mobile Agent Systems. 1070-1079
- Juan Carlos Diaz y Carballo, Alan Dearle, Richard C. H. Connor: Thin Servers - An Architecture to Support Arbitrary Placement of Computation in the Internet. 1080-1085
- Kaiyin Huang, Kaidong Huang: Managing Security in Electronic Business. 1086-1091
- Claudine Toffolon, Salem Dakhli: A Global Model of Electronic Commerce. 1092-1098

Posters

- C. Chandramouli: SRIUI Model: A Design Centric Approach to User Interfaces for Shopping Carts with Emphasis on Intelligence. 1099-1102
- Paulo Novais, Luís Brito, José Neves: The Concepts of Gratitude, Delegation and Agreement in Ecenvironments. 1103-1106
- Apostolos Vontas, Philippos Koutsakas, Christina Athanasopoulou, Adamantios Koumpis, Panos Hatzaras, Yannis Manolopoulos, Michael Vassilakopoulos: Integrating Mobile Agent Infrastructures

in Operational ERP Systems. 1107-1110

- Andrew Blyth, Daniel Cunliffe, Iain Sutherland: XEON - An Architecture for an XML Enabled Firewall. 1111-1116
- Peter Burden: The Webocracy Project. 1117-1121
- Rahim Ghasemiyeh, Feng Li: Introduction to Information Technology and its Effects on Organisational Control. 1122-1125
- Mourad Ouziri, Christine Verdier: Logic and Probabilistic Based Approach for Document Data Modeling. 1126-1129

Home | Conferences | Journals | Series | FAQ — Search: Faceted | Complete | Author

LEARNING TO TEACH DATABASE DESIGN BY TRIAL AND ERROR

Ana Iglesias, Paloma Martínez, Dolores Cuadra, Elena Castro and Fernando Fernández
Computer Science Department
Universidad Carlos III de Madrid
Avda. de la Universidad, 30 28911- Leganés (Madrid) Spain
Tel: {34-91-624}9117, 9454, 9917, 9105 Fax: 34-91-6249430
Email: {aiglesia, pmf, dcuadra, ecastro, ffernand}@inf.uc3m.es

Keywords: Intelligent tutoring System, Reinforcement Learning, Pedagogical Strategies, Coaching and Tutoring.

Abstract: The definition of effective pedagogical strategies for coaching and tutoring students according to their needs in each moment is a high handicap in ITS design. In this paper we propose the use of a Reinforcement Learning's (RL's) model, that allows the system to learn how to teach to each student individually, only based on the acquired experience with other learners with similar characteristics, like a human tutor does. This technique avoids to define the teaching strategies by learning action policies that define what, when and how to teach. The model is applied to a database design ITS system, used as an example to illustrate all the concepts managed in the model.

1. INTRODUCTION

"Intelligent Tutoring Systems (ITSs) are computer-aided instructional systems with models of instructional content that specify *what* to teach, and teaching strategies that specify *how* to teach" [19]. Pedagogical strategies specify how the content is sequenced, what kind of feedback to provide, when and how to coach, explain, remediate, summarise, give an exercise, problem, question or analogy, etc.

The choice of the best strategy has been widely studied, from PMS system [16], where the student has the responsibility of his own learning, providing a set of options to choose, until SOPHIE system [4], that exposed that the experience on a given domain requires "common sense" and "casual comprehension"; this knowledge is difficult to be incorporated in pedagogical rules, which are commonly represented as a set of "if-then" production rules.

But there are many drawbacks in the definition of pedagogical rules. First, there are many pedagogical rules that the system will require to teach effectively, being too expensive to encode them, and second, it is difficult to incorporate knowledge that human tutors do not use themselves, since expert teachers cannot describe how the system should reason with such knowledge [2]. Moreover, it is difficult to determine how much strategies are necessary, the differences among them, the moment to apply them, why they fail, how to solve it, etc. A variety of methods are used to model tutoring expertise, including procedures, plans, constraints and rules [11].

To overcome the obstacles derived from the definition and election of pedagogical strategies, other artificial intelligence techniques have been incorporated, like semantic nets, neural nets, etc. The system MENO-TUTOR [12], for instance, incorporates a procedural net, where a node represents new knowledge and each link (with an associated weight) represents the strategy to follow. In other projects, Machine Learning (ML) techniques have been incorporated in the creation of student models capable of reasoning in a pedagogical way. For instance, ADVISOR system [3] uses reinforcement learning to implement a pedagogical agent that monitors the student's behaviour based on their top-level characteristics.

In this paper, we propose to eliminate the pedagogical strategy concept, at least as it has been understood in previous works, using other kind of representation of the pedagogical information. This new representation is based on a reinforcement learning model [9] that allows ITSs to adapt tutoring to student's needs, sequencing the content in an optimal way based on the student's performance, lesson objectives and the relationships between course modules, avoiding to define all static and predefined pedagogical strategies for each student. The ITS learns by *trial and error* at the same time that students learn ITS's contents. This method is endorsed by the traditional teaching, where the human tutors learn to teach students through a long process of trial and error [14].

In this work, we use a database design ITS to illustrate the working way of the reinforcement learning system. Concretely, we have used the methodology proposed by Teorey [17].

The paper is organised as follow: first, the architecture of ITS used in this work and reinforcement learning model are briefly described in sections 2 and 3 respectively. In section 4, the definition of an ITS system as a reinforcement learning problem is defined. In section 5, an example of working where a Design Database ITS described as a Reinforcement Learning problem learns to teach is shown. Finally, the main conclusions of this work are exposed.

2. ITS MODULES

A typical structure of an ITS is composed of four well differentiated modules (see Figure 1) [5]. The *student module* contains all important information about the student in the learning process: student knowledge, personal characteristics, historical behaviour, etc. The *interface module* facilitates the communication between the ITS and the student. The *domain module* contain all characteristics of the knowledge to teach. Finally, the *pedagogical module* decides *what*, *how* and *when* to teach the domain module contents, taking the better pedagogical decisions according to the user's needs.

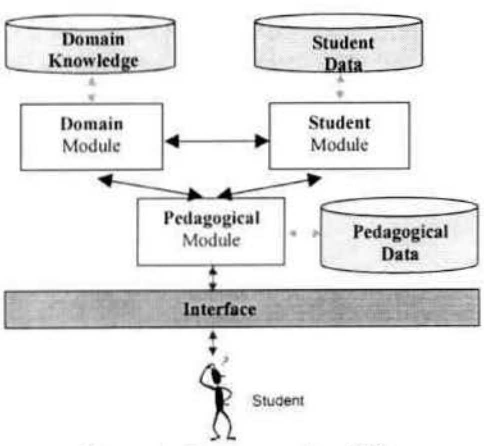

Figure 1: Structure of an ITS

3. REINFORCEMENT LEARNING

Reinforcement Learning (RL) is defined as follow [9]: "an agent is connected to its environment via perception and action, as depicted in Figure 2. On each step of interaction the agent receives as input, *i*, some indication of the current state, *s*, of the environment; the agent then chooses an action, *a*, to generate as output. The action changes the state of the environment, and the value of this state transition is communicated to the agent through a scalar *reinforcement signal*, *r*. The agent's behaviour, *B*, should choose actions that tend to increase the long-run sum of values of the reinforcement signal. It can learn to do this over time by systematic trial and error, guided by a wide variety of algorithms."

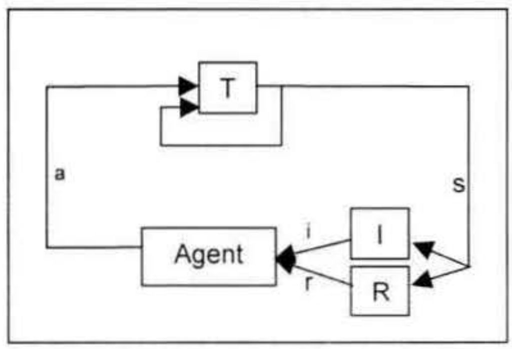

Figure 2: Reinforcement Learning Model

4. USING REINFORCEMENT LEARNING IN ITS

This paper is focused on the application of the reinforcement learning model for deciding (by learning) what, how and when to teach a concept, according to current student's needs. For the effectiveness of this proposal it is necessary, to construct a good *student model* and to classify learners according to their critical characteristics in learning the ITS's knowledge. A high variety of student models and techniques for constructing them have been studied [15], and any classification technique, like C5.0 [13], could be used to assort students. In order to bound the aims of this paper, let it be a good assortment of students.

ITS's knowledge is stored at the *domain module*. The traditional knowledge structure (topics, sub-topics, etc.) could be an advantage in the pedagogical strategy in ITS. That's why we propose to use this hierarchy to define the tutor system's objectives. Figure 3 shows a proposal of a

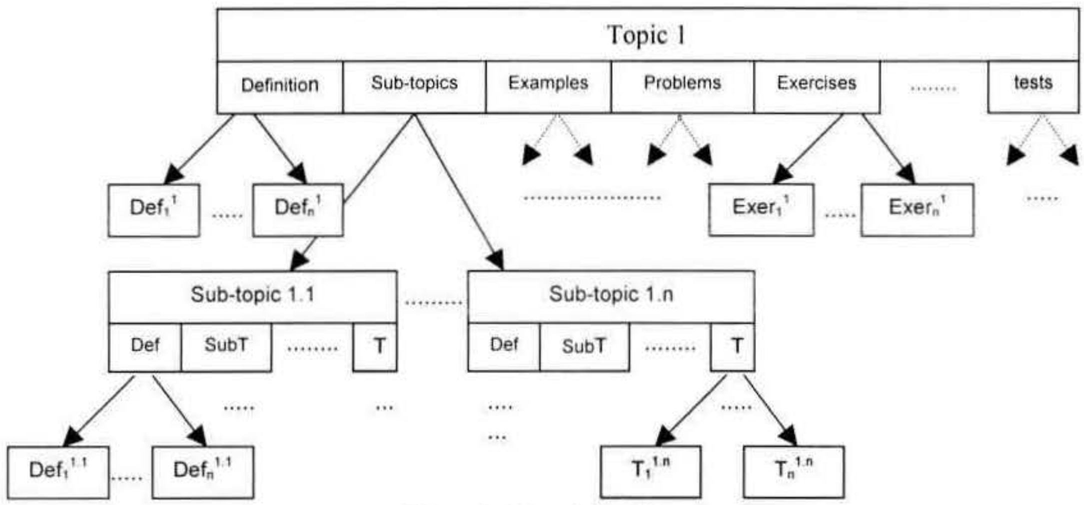

Figure 3: Knowledge Tree

hierarchical structure of knowledge, where each topic has been divided into sub-topics, and these in others sub-topics, and so on. At the same time, each node of the tree contains sets of definitions, examples, problems, exercises, etc. in several formats (image, text, video, etc.).

At the *pedagogical module,* the ITS finds the best way to teach the *knowledge items,* corresponding with the internal nodes of the tree (topics), to the current student. The definition of this problem as a reinforcement learning problem is fulfilled as follow. The agent's *state* is the current student's knowledge, represented by a vector of representative values of the student's knowledge for each topic. The ITS perceives the current student's knowledge (s_1) by evaluations (*tests*). Given one state, the system chooses an action to be executed according to the current *action policy*, *B*. The *action* corresponds with showing leaves of the knowledge tree (*definition, exercise, problem*, etc.). This action is supposed to change the current state of the student's knowledge to a new state (s_2), generating a *state transition* and a *reinforcement signal* (positive or negative) to the system. This signal is used to update the system's action policy. The system behaviour, *B*, should choose actions that tend to maximice the long-run sum of values of the reinforcement signal, choosing in this way the optimal tutoring strategy (what, when, and how; the best sequence of contents and how to teach them) to coach the current learner.

5. A DATABASE DESIGN ITS LEARNING TO TEACH

In this section, how an ITS could learn to teach using the Reinforcement Learning model is shown. Database Design (DBD) domain have been proposed as example of an ITS domain. This domain has been included too in PANDORA project[1], where we are working in this moment. The PANDORA goal is to define methods and techniques for database development implemented in a CASE tool, useful for students and practitioners, where one of the modules of this project is concerned with tutoring learning via Web.

5.1 Design Database Domain Knowledge

The knowledge model used for this system is based on the methodology proposed by Teorey [17] in which he emphasises three main phases in the database design: conceptual, logical and physical phases. Data models used are the Entity Relationship (E/R) model [6] and the Relational model [7]. The physical design is not based in any data model due

[1] This work takes part of PANDORA project (CASE Platform for Database development and learning via Internet), Spanish CICYT project (TIC99-0215).

to it depends of the DataBase Management System (DBMS).

In order to simplify the example, we focus on the "binary relationship" sub-tree, specially on the "N:M" topic (see Figure 4), supposing that "entity", "attribute", "degree" and "cardinality" topics have already been learnt by the student.

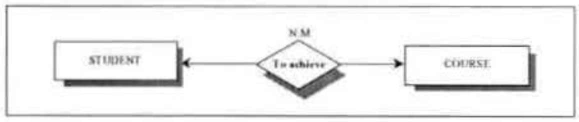

Figure 5: N:M relationship of the Definition 2

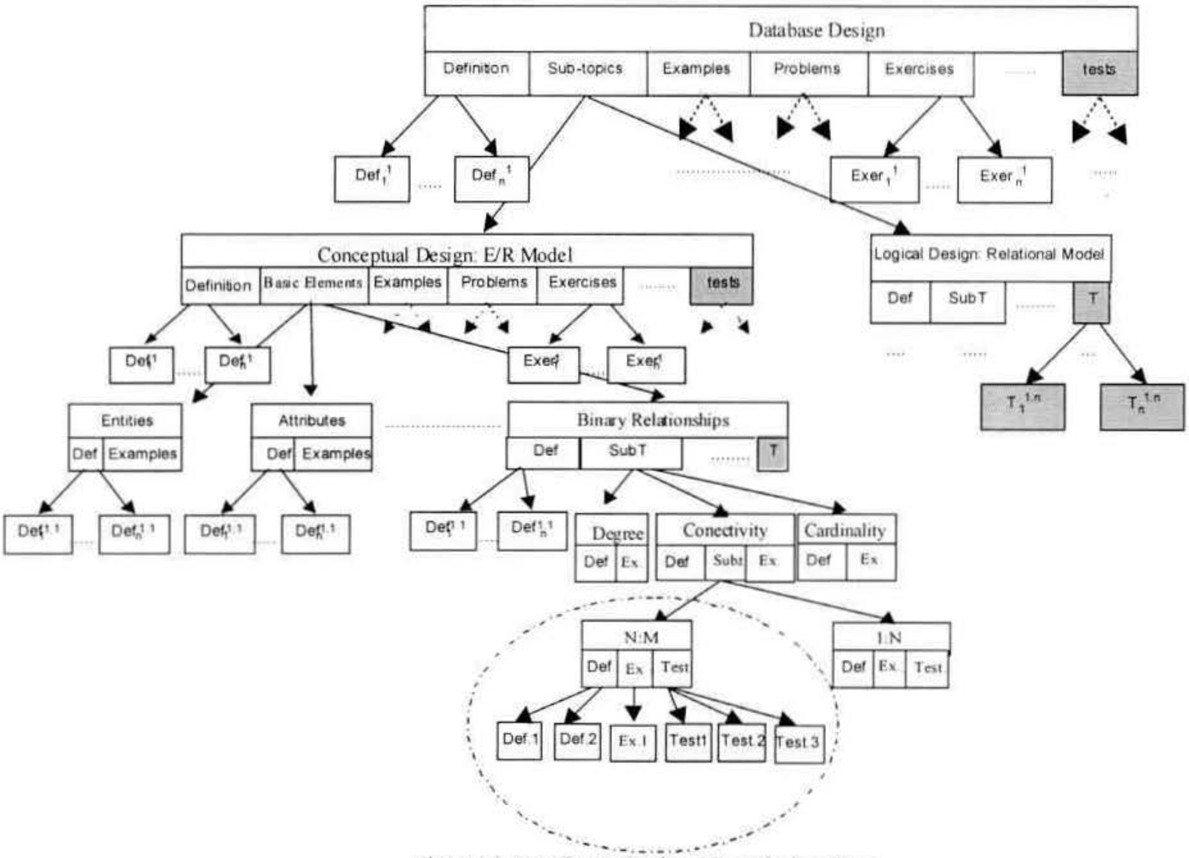

Figure 4: Database Design Knowledge Tree

The "N:M" topic contains the following definitions, examples and tests:

1) Definitions: This topic contains several definitions with different formats.
   - Definition 1 (Text format).- "The connectivity N:M of a binary relationship indicates the mapping between more than one instance of the entity types that participate in the relationship [1].
   - Definition 2 (Image and Text format).- "Let "to achieve" be a relationship between the entity types STUDENT and COURSE (Figure 5). As one student can achieve many courses and in one course many students are registered, then one instance of each entity type participates more than once in the relationship "to achieve", so the connectivity must be N:M [8]"
2) Examples:
   - Example 1 (Image and Text format): "In the domain of movies we could be interested in representing the association between actors and movies in which they have worked. One actor must act in some movies and usually in a movie more than one actor participate, so we could model a N:M connectivity as it is shown in the Figure 6."

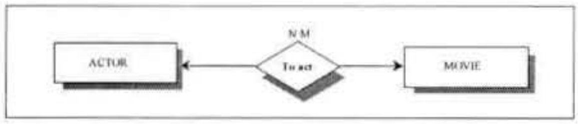

Figure 6: N:M relationship of the Example 1

3) Tests:
   - Test 1: "Read carefully the following questions and identify the N:M connectivity:
     (a) A client orders backlogs and one backlog is ordered by a client
     (b) A recipe is made with some ingredients and one ingredient may appear in various recipes
     (c) One employer may work in a department, and in a department works several employers
     (d) One book may be written by some writers and a writer may be author of various books"
   - Test 2: "If you are told that a recipe is made with some ingredients and one ingredient may appear in various recipes, which connectivity you will draw in the Figure 7?"

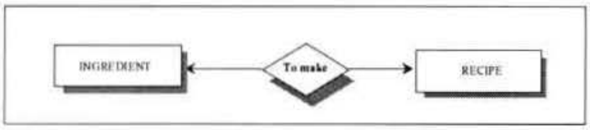

Figure 7: N:M relationship of the Test 2

   - Test 3: "Let us suppose that we want to design a database for a school. We know that one student studies at least one subject and one subject may be teached by two teachers at least. Which connectivity you will draw in the relationships of Figure 8?"

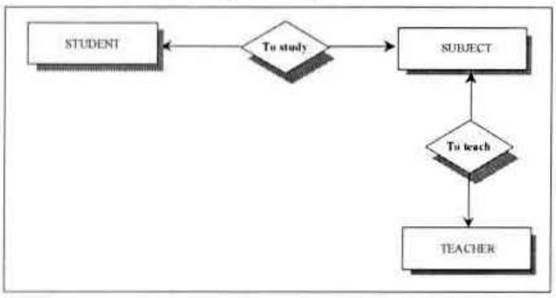

Figure 8. N:M relationships for Test 3

5.2 Reinforcement Learning Components

In this section, the components of a reinforcement learning problem in the database design intelligent tutoring system environment are defined.

1.- Set of states (S): A state is defined as the description of the student knowledge. It is represented by a vector which stores as many representative values of learner's knowledge items (internal node of the knowledge tree) as is wished the student learns. In order to simplify the example, it is supposed that these values are defined in the set {0,1}. The items of the knowledge tree are enumerated in an *pre-order* way. The zero value indicates that the student does not know the item, and the one value indicates that the item has been correctly learned. In the Figure 9, a possible state of the system is shown, based on the knowledge tree defined in Figure 4 and bounded to the "Binary Relationship" branch, where the student knows the "Degree" and "Cardinality" items, but not the others.

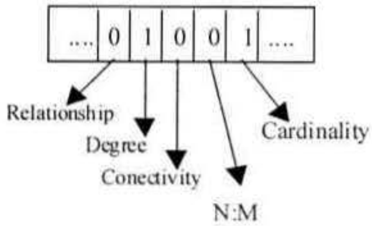

Figure 9: state vector example

2.- Set of actions (A): The actions that the tutor can execute are those that teach the ITS's subject, i.e. to show the leaves of the knowledge items defined in the knowledge tree. It is allowed to show *macro-actions*, that define a set of leaves of the tree that will be shown at the same time. It is necessary to know if a leaf of the tree has been shown to a student, because it affects on the state of the system. For instance, a student could learn a topic just after teaching him/her a leaf, but he/she could understand it thanks to the fact that previously, the system has shown him/her another topic. Thus, let us suppose a system with the macro-action defined in Figure 10. There are actions composed by only one leaf and actions composed by several leaves. Let us suppose too, that the student is in a state *s*, and the system executes the action *a1*, i.e. the system shows to the student the definition one (def1). If the student would not learnt the item, the system have to execute another action that will content the previous one. For instance, action *a4*.

a1 = to show definition 1= {def1}
a2 = to show definition 2 = {def2}
a3 = to show example1 = {ex1}
a4 = to show {def1} + {ex1} at the same time {def1, ex1}
....

Figure 10: examples of macro-actions

3.- Perception of the environment (I: S → S): This function indicates how the ITS perceives the state the student is into. To perceive the consequences of the execution of a given action, to evaluate the student's knowledge is required. The only way an ITS could perceive the knowledge state of the student is through evaluating his/her knowledge by tests (presented by shadow shapes in the Figure 4). Each knowledge item in the ITS has associated a set of tests, that will serve to evaluate how much the student knows about the item. These tests will be centred at the last action executed by the ITS, assuming that other student's capabilities evaluated before do not vary from the last evaluation. However, this could be false, varying them. This problem implies that an ITS could receive from the environment states with different levels of information, that could make the system not to differentiate some states from others, or perceiving two different states as equal (POMDPs: Partially Observable Markov Decision Problems), as it happens in teaching with human tutors. To avoid this problem, there will be necessary to accomplish continuous re-evaluations.

4.- Reinforcement (R: SxA → R): This function defines the reinforcement signals (rewards) provided by the environment. This reinforcement function supplies a maximum value upon arriving to the goals of the tutor; in conclusion, the aim of the ITS will be to maximise its long term reward. For example, a goal for the tutor would be that the user learn the "N:M" item, receiving only a positive reinforcement signal when the component of the state vector in this item is one. Only in this moment when the teaching process is considered finished. The application of the reinforcement signal is a key issue in the learning technique, because to decide when and how to apply the reward is crucial for the system learning. The ITS learning is a delayed positive reward problem, given that it is impossible to know the kindness of the actual state perceived immediately after the performance of an action (the system does not immediately know whether it has achieved its goal after executing an action, since it might be only achieved (or not) after the execution of several ones). Furthermore, a higher reinforcement signal will be applied when the student may have learnt in less time and better (positive impact on achieving the goal).

5.- Value-action function (Q: SxAxΠ → R): This function estimates the usefulness of showing leaves of the tree to a student when he is in certain knowledge state. This function provides an evaluation method for the tutor action policies. Therefore, the goal of the learning process is to find the policy such that it maximises this function, i.e., to obtain the optimal value-action function (see equation (1), where γ it is the discount parameter of future actions).

5.3 Q-Learning

The reinforcement learning algorithm proposed in this work to learn the Q function is the Q-learning algorithm [18]. This algorithm allows to learn directly from the experience, based on the *value-action* function, $Q(s,a)$, to define its action policy. One of the main characteristics of this algorithm is that it does not require that system executes optimal sequences to converge [10], being able to learn the Q function and, therefore, the optimum policy. This is very important for ITS, given that learning is based on the interaction with users, and this implies a high cost.

The Q-learning algorithm is described in Table 1. It requires a definition of the possible states, S, the actions that the agent can perform in the environment, A, and the rewards that it receives at any moment for the states it arrives to after applying each action, r. It dynamically generates a action-value table, $Q(s,a)$, that allows to follow a potentially optimal policy. The γ parameter controls the relative importance of past actions rewards with respect to new ones, and α parameter is the learning rate.

- For each pair ($s \in S$, $a \in A$), initialise the table entry Q(s,a).
  - Observe the current state, s
  - Do forever
    - Select an action, a, and execute it
    - Receive the immediate reward, r
    - Observe the new state, s'
    - Update the table entry for Q(s,a) as follows:
    - $Q(s,a) = (1-\alpha)Q(s,a) + \alpha(r + \gamma max_{a'} Q(s',a'))$
    - Set s to s'

Table1. Q-learning algorithm

In its learning, the ITS uses the state description (s) as input of the function approximator (Q), using this function to select a teaching action (a). After executing actions, the student changes its state, that has to be perceived by the ITS (using tests). Depending on the kindness of the state reached in order to attain the goals, the system will receive a positive or null reward.

To define if in ITS domain the execution of an action in a given state does or not always arrive to the same state is critical. When it does not achieve always the same state is because of noise associated to perception, and actions. This implies that the reward of executing an action in a given state, can be different depending on whether the next observed state is a goal state or not. In ITSs there is not determinism in actions, because their execution could arrive to different final states, due, for instance, to the different characteristics of learning that students have (the student can learn or not the item, with different results on the tests, etc.). Nevertheless, to simplify the example in this paper, we suppose there exist determinism, and, then, the learning rate = 1.

In the other hand, an exploitation/exploration (to use pre-acquired knowledge or to tray new alternatives respectively) strategy must be defined, because it influences the learning rate and its quality. In this case, any strategy (greedy, random, etc.) can be used, but we propose to mix them with the possibility that ITSs present in the interaction with users, proposing to the users a list of options with the following actions to be performed. In this way, at the same time, the ITS maintains the attention of the pupil, giving him/her the sensation of keeping the control of the interaction, that is an interesting characteristic of ITSs.

5.4 ITS Learning Example

An example of the ITS learning process is shown in the Figure 11, where the students states have been represented by state vectors and the states transition by arrows. The actions performed before the state transition label each arrow, and they have been marked with the initial action-value function (*Q(s,a)*) value. The *S* state is the initial state of the example and de *Goal* state is the final state. Thus, in the example, we only manage two states, *S* and *Goal*. The possible actions have been represented in Figure 10.

The *Q* values of the *Goal* state are set to zero by reinforcement learning definition and de *Q* values for *S* state are supposed to be, for instance, 0.8. So, the *Q* function values can be represented with the following table:

	A1	A2	A3	A4
S	0.8	0.8	0.8	0.8
Goal	0.0	0.0	0.0	0.0

Table 2. Example of *Q* function values

Let us suppose that the *A1* macro-action is randomly chosen and then, after its execution, the system achieve the same state, *S*, because the student is not able to learn the item (the student does not pass the related test). In this moment the *Q* function value is updated based on the following update function:

$$Q(s,a) = \gamma^{(size(a)-1)} r + \gamma^{size(a)} \max_{a'} Q(s',a') \quad (1)$$

Where *size(a)* is the number of leaves that the action *a* contains. This function is a modification of the standard Q-learning updating function, where the reinforcement can be discounted depending of the size of the macro-action. So, if the macro-action size is small, the new value will be bigger than if the macro-action size was higher. Thus, the new value of Q(S,A1) will be computed as follow:

$$Q(S,A1) = 0 + 0.9^1 * \max_{a'}\{0.8,0.8,0.8,0.8\} = 0.72 \quad (2)$$

Given that the student has not arrived to the *Goal* state, another action must be executed. Supposed that the *A4* macro-action is chosen, and this change the student state to the *Goal* state (the student pass the test). Then, the value of Q(S,A4) is computed as

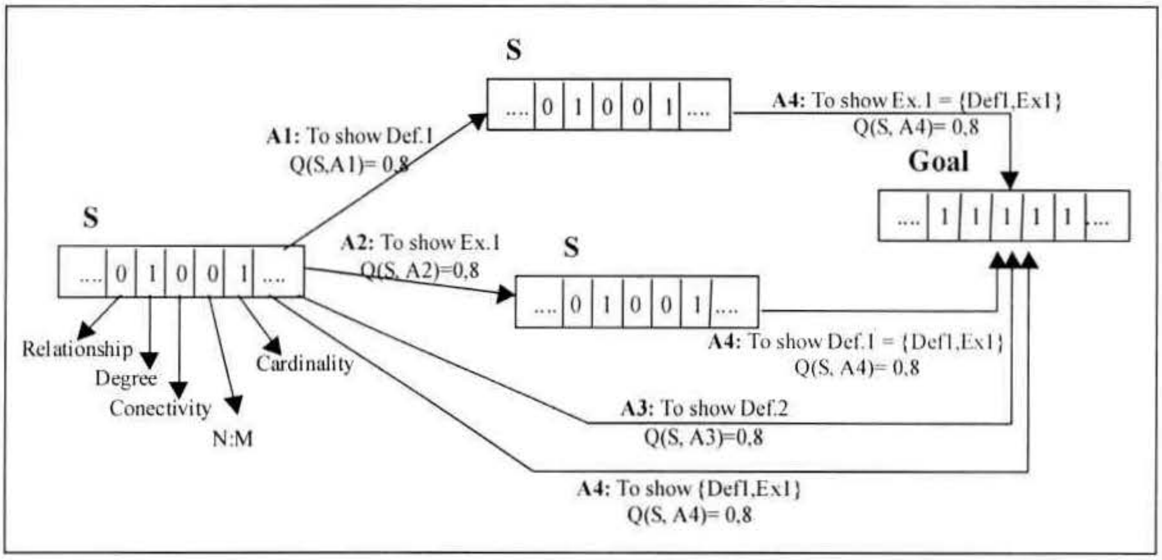

Figure 11: ITS Learning Example. Initial Situation

follow:

$$Q(S,A3) = 0.9^{1-1} * 1 + 0.9^{1} * \max_{a'}\{0,0,0,0\} = 1 \qquad (3)$$

Now let us suppose that another student with similar learning characteristics than the previous student is in the same initial situation (the S state). Suppose too that the system decides to execute the *A3* action. After executing this action, the student achieves the *Goal* state, updating the *Q(S, A3)* as follow:

We can see that even actions *A3* and *A4* achieve the *Goal* state, the value of the *Q* function for *A3* action is bigger than for *A4* action. This is due to the action *A3* only contains one leaf (a definition), and *A4* action contains two leaves (a definition and an example).

5.5 Conclusions and Further Research

In this paper we propose to use a technique that avoids the problems derived from construction of different pedagogic strategies for each student in Intelligent Tutoring Systems. This technique consists on the use of reinforcement learning in the pedagogical module of the ITS, learning to teach by trial and error and eliminating the problematic concept of pedagogical strategy.

The use of reinforcement learning techniques provides many advantages on the techniques of traditional dynamic planning. First, it avoids the cost associated to rules construction and heuristic creation for each teaching strategy adapting to the needs of each student. Furthermore, it allows the system to individually adapt to student in real time, based only on previous information of interactions with other students with similar characteristics. Thirdly, the system is able to adapt to the user not only deciding what to show, or when to show a given knowledge, but also how to do it. Finally, it is considered a general technique, being able to apply it in any ITS independently of the domain.

This paper exposes the foundation of a theoretical model, proposing a representation of the domain module in intelligent tutoring systems and instantiating it with a real example that is currently being implemented.

Currently, we are involved in the implementation work as well as in defining the evaluation with real students.

REFERENCES

[1] Batra, D. and Zanakis, H. *A conceptual database design approach based on rules and heuristics. European P.* Journal of Information Systems, Vol. 3, No. 3, pp. 228-239, 1994.

[2] Beck, J. *Learning to Teach with Reinforcement Learning Agent.* In Proceedings of the Fifteenth National Conference on Artificial Intelligence, 1998.

[3] Beck, J.; Woolf, B. and Beal, C.R. *ADVISOR: A machine learning architecture for intelligent tutor construction.* In Proceedings of the Seventeenth National Conference on Artificial Intelligence, 2000.

[4] Brown, J.S. and Burton, R.R. *An investigation of computer coach for informal learning activities.* D. Sleeman, J.S. Brown (Eds), Intelligent Tutoring Systems, Academic Press, San Diego, pp. 79, 1982.

[5] Burns, Hugh and Capps, Charles. *Foundations of Intelligent Tutoring Systems: An Introduction.* Foundations of Intelligent Tutoring Systems. Hillsdale, N.J: Lawrence Erlbaum Associates, 1-19, 1988.

[6] Chen, P. *The Entity-Relationship Model – Toward a Unified View of Data.* ACM Transactions on Database Systems, Vol. 1, N. 1, 1976.

[7] Codd E. F. *A Relational Model of Data for Large Shared Data Banks.* CACM 13 (6), June, 1970.

[8] Cuadra, D.; Castro, E. And Martínez, P. *Curso Mentor: Diseño de Bases de Datos.* Servicio de publicaciones del PNTIC (Ed). NIPO: 176-01-5 ISBN: 84-369-3473-3

[9] Kaelbling, Leslie Pack; Littman, Michael L. and Moore, Andrew W. *Reinforcement learning: A survey.* Int. J. of. Artificial Intelligence Research, 237-285, 1996.

[10] Mitchell, Tom Michael. *Machine learning* Boston: McGraw-Hill, 1997.

[11] Murray, Tom. *Authoring Intelligent Tutoring Systems: An analysis of the state of the art.* International J. of Artificial Intelligence in Education, 10, pp. 98-129, 1999.

[12] Murray, Tom. *Control of ITSs: a comparison of blackboard architectures and discourse management networks.* M.J. Farr, J. Psotka (Eds.), Intelligent Instruction by Computer: Theory and Practice, Taylor and Francis, New York, pp. 87, 1992.

[13] Quinlan, J.R. *C4.5 Programs for machine learning.* Morgan Kaufmann Publishers, San Mateo, California, 1993.

[14] Reinsmith, W.A. *Archetypal forms in teaching: A continuum.* Greenwood Press, Westprot, US, 1992.

[15] Sison, Raymund and Shimura, Masamichi. *Student Modeling and Machine Learning.* International Journal of Artificial Intelligence in Education, 9, 128-158.

[16] Sleeman, D. *A system which allows student to explore algorithms*. Proceedings of International Joint Conference on Artificial Intelligence 5, pp. 780-78, 1977.

[17] Teorey, T., Yang, D., Fry, J. *A Logical Design Methodology for Relation Databases Using the Extended Entity-Relationship Model*. Computer Surveys, Vol. 18. No. 2. 1986.

[18] Watkins, C.J.C.H. *Learning from Delayed Rewards*. PhD thesis, King's College, Cambridge, UK, 1989.

[19] Wenger, E. *Artificial Intelligence and Tutoring Systems*. Los Altos, CA: Morgan Kaufmann.