Research and Education by using IBM Software Products

dr. András Pataricza Associate professor, Dept. Measurement and Information Systems, Budapest University of Technology and Economics Director, Research and Technology Center, University of Veszprém Former Deputy Director for Education, Information Technology Center, Budapest University of Technology and Economics













Typical academic approach

 Model based thinking search for general solutions use mathematics, as thinking aid remain platform/technology independent attack core problems How to introduce a new technology into E&R? St. Petersburg András Pataricza 7

MM MW W . /M > II



Example: data mining

 Data: byproduct of (e-)Business processes Berkeley: 1 ExaByte/year new data More data in the forthcoming 3 years, as in the entire history of the mankind Knowledge Discovery: huge number of records in Date Bases extraction of the hidden knowledge from the operational log files general conclusions / models St. Petersburg András Pataricza 9

Concepts of Data mining

Knowledge: new, non-trivial relevant easy to understand Part of the business intelligence solution package of IBM DB2, Intelligent Miner, OLAP, text, scoring...

St. Petersburg



Principle of clustering



Data mining process











Typical questions

What percentage of the lectures did you attend?	How easy can you make notes at the lectures?
 More than 80 % 60-80% 40-60% 20-40% Less than 20% Your grade point average in the previous semester (5.0= excellent)? Above 4,0 Between 3,5 and 3,99 Between 3,0 and 3,49 Between 2,5 and 2,99 Below 2,5 	 It is extremely easy It is easy to make notes. We can make notes partially. It is difficult to make notes. Impossible to make notes. I did not try to make notes. I did not try to make notes. Does the lecture teach to solve problems? It completely does. Mostly it does. It does more or less. It does not. It is not necessary for this subject.
St. Petersburg	András Pataricza 18





Typical learning process Small-scale application Working principle/ mathematical background Getting started: technology basics Application development System engineering Research insufficiencies from your scientific/application host environment design methodologies András Pataricza 20 St. Petersburg





Elementary neuron

$$y = f\left(\sum_{i} w_{i} x_{i}\right) = \operatorname{sgn}\left(\sum_{i} w_{i} x_{i}\right) = \operatorname{sgn}(s)$$



Remark:

- IBM performs extensive basic research
 - Visualization: high-dimensional data in large, complex data sets
 - Graph manipulation
 - Optimization Solution Library: LP, QP, Network, Stochastic Programming, MIP
 - Data mining in life sciences
- www.alphaworks.ibm.com

Use in E & R

- IBM SW can be used to illustrate the principles
 - Marketing focuses on the typical applications
 - By understanding the operational principles:
 - unorthodox fields of applications



Computer science

Natural sciences (mathematics)

The education pyramide

St. Petersburg



 Storing data in DB2 to support the use of the Intelligent Miner tool



Conclusions

- Based on structure of web pages and the document structure of Lotus
 - Downloads of multiple pages because of framesets
 - After downloading a view document will be opened
 - Every student in our branch has now a similar lab practice
 - understanding of data mining basics
 - understanding of web servers

St. Petersburg

Typical learning process Small-scale application Working principle/ mathematical background Getting started: technology basics Application development System engineering Research insufficiencies from your scientific/application host environment design methodologies András Pataricza 27 St. Petersburg



Technologies

All the technologies are
 standards based
 principles covered by usual courses

but

St. Petersburg

- additional, value added services
- domain specific extensions
- tuned for efficiency, performance



IBM approach from a university perspective

IBM software is highly standards compliant

- Education:
 - platform independence
 - good support by books
 - universal skills
- Open software
 - Support of LINUX
 - Java, XML technologies
 - Open projects, like Eclipse
 - AVAILABLE TO ALL STUDENTS

St. Petersburg

Typical learning process

- Small-scale application
- Working principle/ mathematical background
- Getting started: technology basics
- Application development
- System engineering
- Research
 - insufficiencies from your scientific/application
 - host environment
 - design methodologies

St. Petersburg



Skills required

- Efficiency needs a deep training
 - Irregular courses: separate unit
 - "e-Business Academy"
 - large set of course materials
- Self-learning
 - project lab
 - diploma theses
 - IBM documentation is sufficient!

St. Petersburg





Typical learning process

- Small-scale application
- Working principle/ mathematical background
- Getting started: technology basics
- Application development
- System engineering
- Research
 - insufficiencies from your scientific/application
 - host environment
 - design methodologies

St. Petersburg

System design



 Topmost level of abstraction
 Model-based thinking and interaction
 Knowledge and skill fusion

Interdisciplinary

St. Petersburg







 IBM participates in the standardization
 Limited support in IBM tools
 Interfaces to external tools: XML, JAVA

St. Petersburg

OMG and IBM (incomplete list)



Typical learning process

- Small-scale application
- Working principle/ mathematical background
- Getting started: technology basics
- Application development
- System engineering
- Research
 - insufficiencies from your scientific/application
 - host environment
 - design methodologies

St. Petersburg





- Valuable results in the *academic world*
- *Industrial utilization Only* at specialized development sites
- Vital necessity:
 - *Complexity* of IT products
 - Increasing requirements for
 - dependability
 - QoS

St. Petersburg

Novel approach



Example: BPM based robust E-business application development



Example: reservation of lodging

Functionalities: Registration Search Reservation Advance payment Modification (service, period) Cancellation Process management András Pataricza 46 St. Petersburg





Challenges



St. Petersburg



Simple workflow





Error propagation analysis



Algorithm Based Fault Tolerance



Optimization



- Avarage workload
 - 18 use cases
 - e.g.: use case #3: search - reservation - advance payment (21/day)
 - processes to be handed
 - 2894 / day

St. Petersburg

Performance estimation









Research infrastructure



Budapest Model

- Separate, dept ranked unit
 - Own staff
- Supervisory council
- Loose relations to faculties



Veszprém Model

- Initially: Center of Competence within a department
- Next step: R&T Center
 - Rank of a dept.
 - No regular education duties
 - Staff: innovative (post)graduates
 - Temporal assignment of university staff
 - Partial organizational independence



St. Petersburg

Designated R&T

Faculty of Infomatics

University of Veszprém



