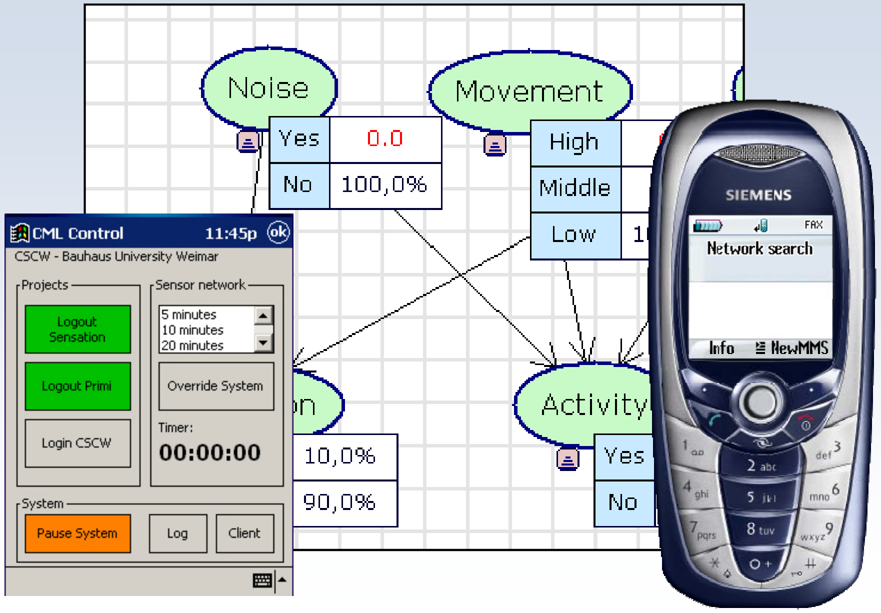
Bauhaus University Weimar



Research, Technology:
Bayesian Networks, Mobile Interface

Project Sensation

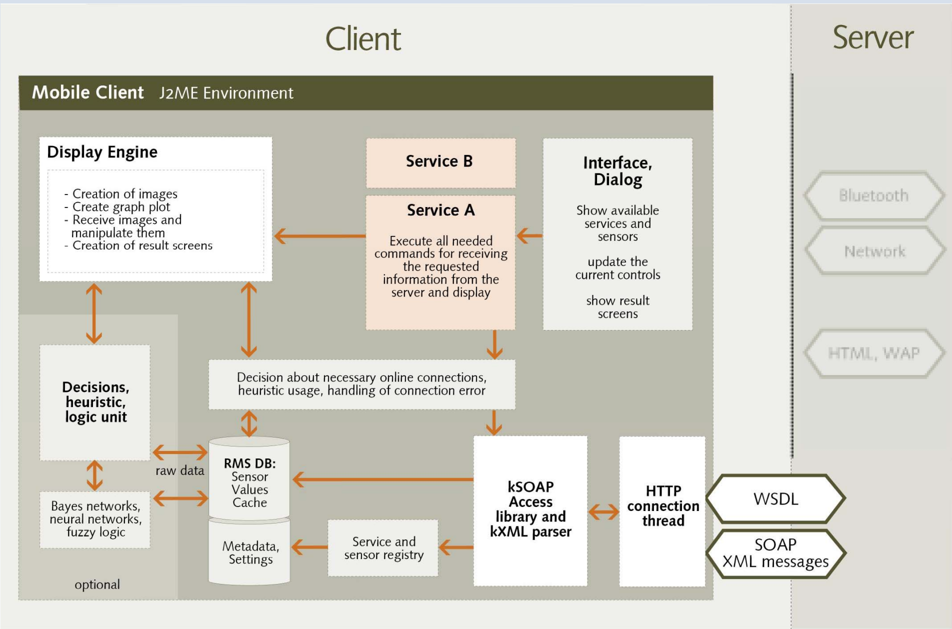
November 2004
Nicolai Marquardt
CML Cooperative Media Lab
CSCW, Prof. Tom Gross, Tareg Egla
Bauhaus University Weimar

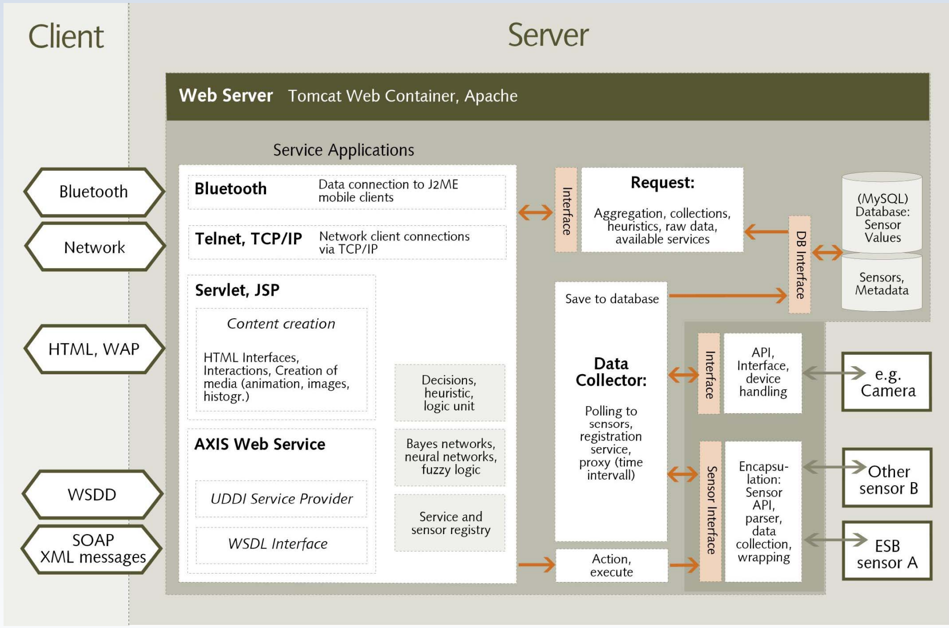
Outline

1. Infrastructure Overview
2. Introduction: Bayesian Networks
3. Using Bayesian Networks
4. Mobile Interface
5. IR Control, Interaction

1. Infrastructure Overview

1. Infrastructure Overview



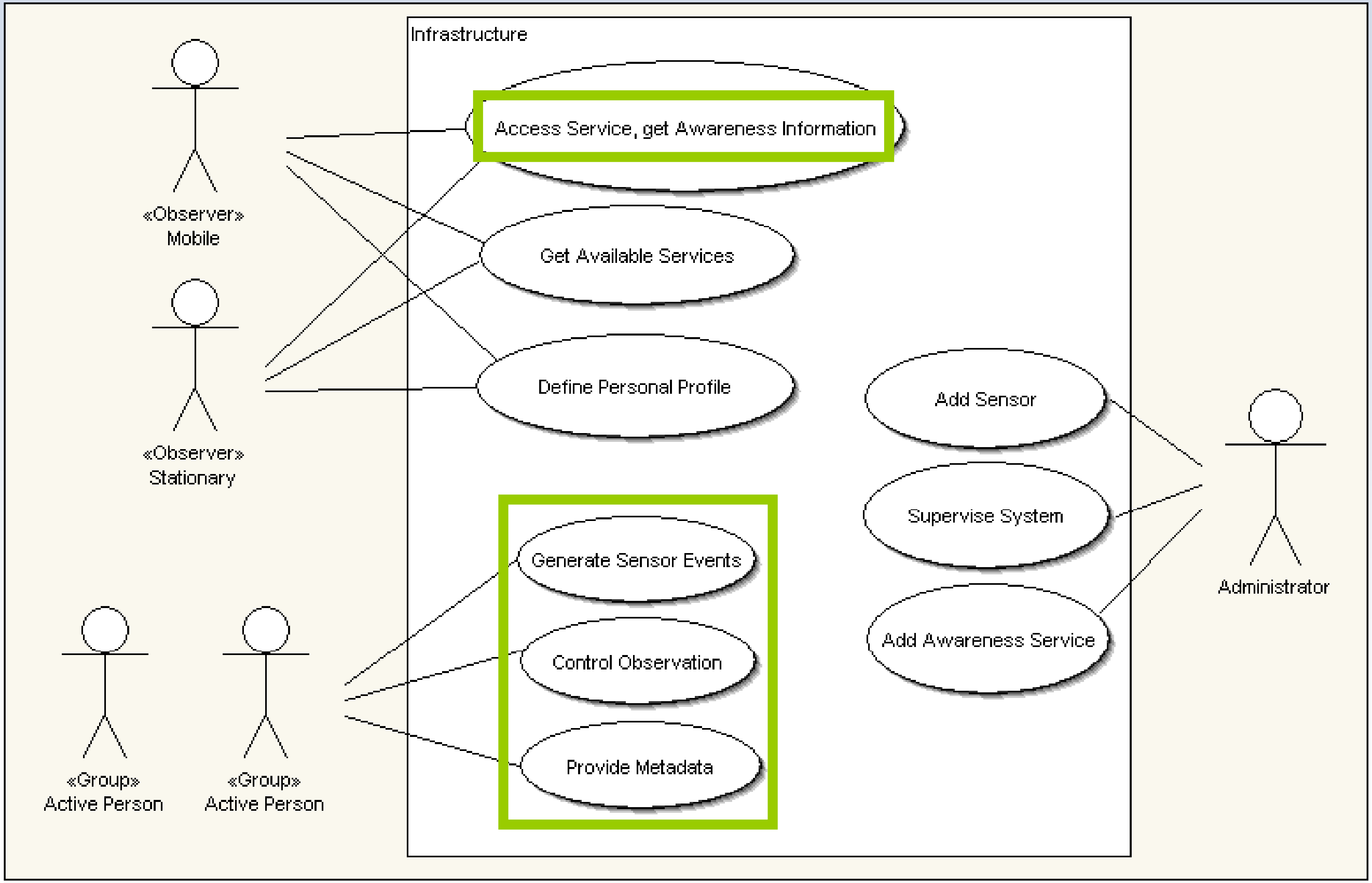


1. Infrastructure Overview

The objectives for the sensor infrastructure:

1. Easy integration of different sensors (and implementation hiding)
2. Different interfaces and their evaluation
3. Using web services, WSDL
4. **Implementation of clients, especially for mobile devices**
5. Intelligent connection decisions: save costs and time
6. **Predict values, use heuristics, "sensor paths", Bayesian Networks**
7. Proxy for web service access via mobile devices
8. Independence from Hardware, OS and Programming Language

1. Infrastructure Overview



1. Infrastructure Overview

The main issues and solutions:

Demand, Problem	Solution
Data Collector ⇔ Sensor Board (ESB), reduce the COM data transmissions	a) Sensor access timestamp, proxy b) Only the collector has access to sensor board data
Server ServiceAccess ⇔ DB, reduce the read and write operations to the database	Internal static member variables for caching, class can control the update of internal data
Client (mobile) ⇔ Web Service, reduce the requests/transmissions	a) Transmission of data collections (Integer array) b) Prospective calculations (estimate)
Client (mobile): save online time and costs	Intelligent decision logic for the mobile client: can decide if it is necessary to update the data cache
Different clients, different platforms	Various connection possibilities: Web Service, Bluetooth, TCP/IP, etc. → but the same methods (interface)
CPU time extensive XML parsing, especially with mobile clients	Concept: proxy service provides translation and parsing service → separate protocol

2. Introduction: Bayesian Networks

2. Introduction: Bayesian Networks

Probability theory (statistics):

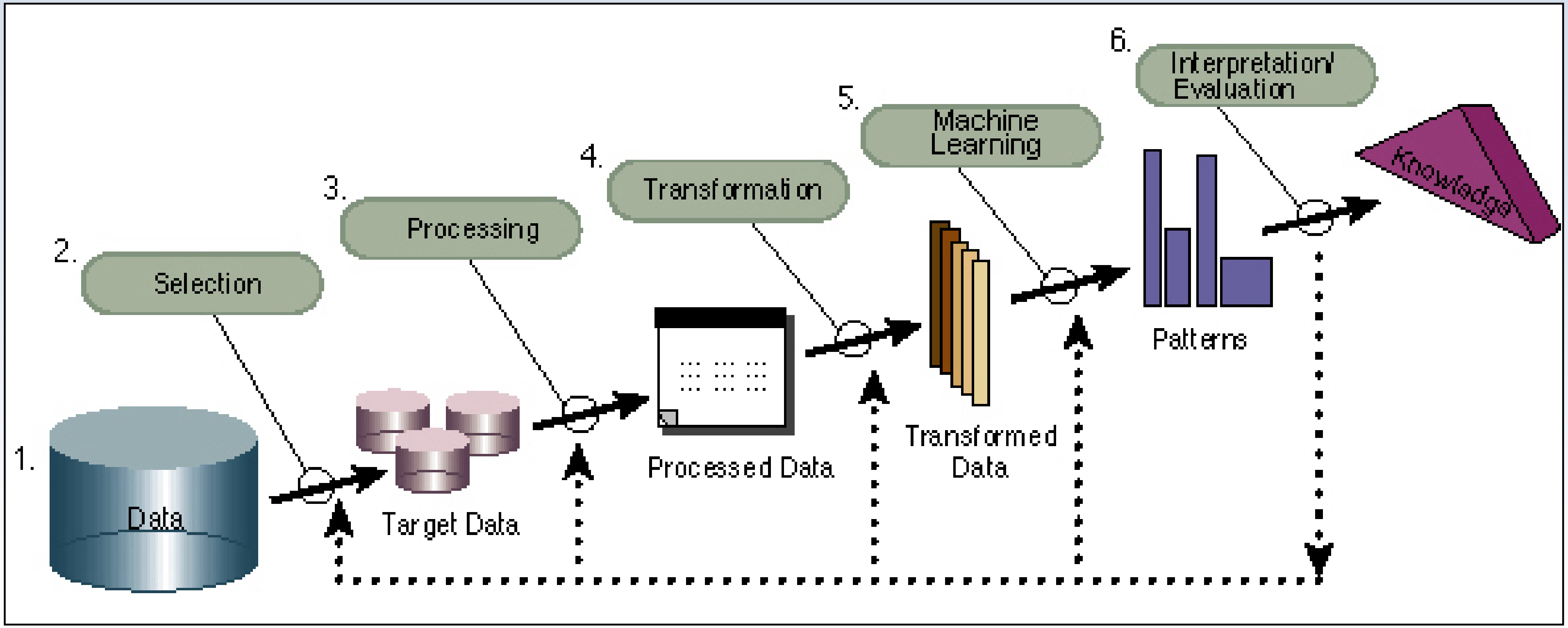
- Our objectives in general:
  calculations and predictions with sensor data

- Data pre-calculations:
  - Classify and group the sensor values (time slots)
  - Build average values, use arithmetic means, probability distribution
  - handle divergent values (detection difficult)
  - Interpolation of values

Estimate prospective values

2. Introduction: Bayesian Networks

Overview of the Steps That Compose the Knowledge Discovery Process:



Reference: Kensas State University, adapted from Fayyad, Piatetsky-Shapiro, and Smyth (1996)

2. Introduction: Bayesian Networks

Methods for machine learning:

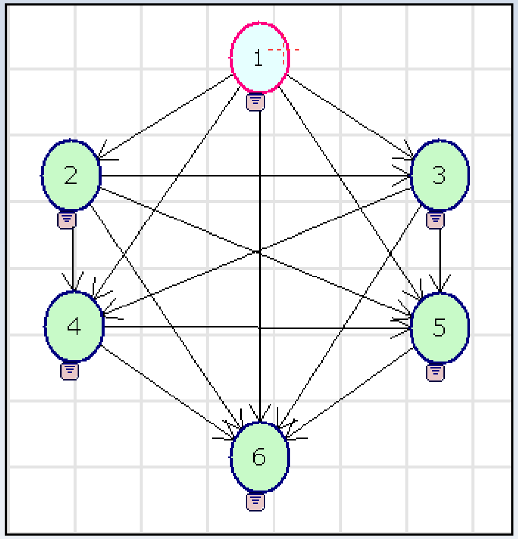
- Neuronal Networks, Fuzzy Logic, Hidden Markov Models (HMM), building Simulations
- Recommended: Bayesian Networks
- Process in general:
  1. Building network
  2. Training/learning phase (known values, examples)
  3. Inference, requests

- Applicable to: computer vision, speech recognition, determination of errors, predicting values and other fields of Data Mining

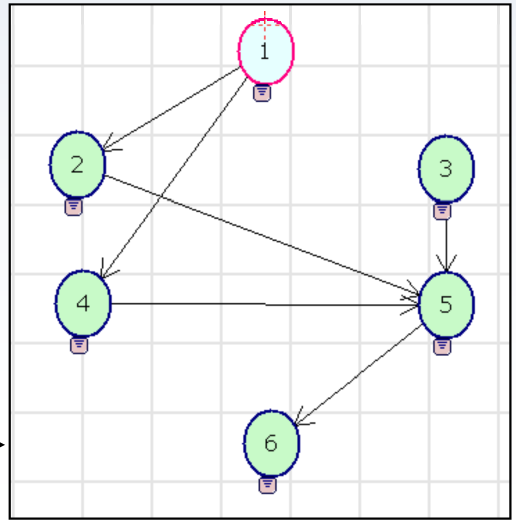
2. Introduction: Bayesian Networks

Components of Bayesian Networks [Murphy]:

- Can build abstract representations for complex circumstances in the real world (with probabilities)
- Directed, acyclic graph (DAG)



- Nodes:
  - Discrete variables or true/false
  - Variables of interest: feature of an object, state, capacity, occurrence of an event
- Edges:
  - Informational or causal dependencies
  - Dependencies are quantified by conditional probabilities for each parent node
- Build a network of conditional probabilities
- Conditional independence



2. Introduction: Bayesian Networks

Advantages of Bayesian Networks:

- Usable for complex simulations of the real world
- Networks can hold knowledge in a form of collections of probabilities, and so they can simulate human intuition and conclusions
- Easy to apply for model and simulation tests
- Readable by humans (in contrast to neuronal networks)
- Preserve knowledge of experts

Disadvantages of Bayesian Networks:

- Less exactly than neuronal networks, but more efficient
- It is difficult to examine the solutions of the network
- It is also difficult to get the probability knowledge

2. Introduction: Bayesian Networks

Deduction (top-down reasoning):

- Conclude from the given event to all conditional events in the graph structure
- Use conditional probability:

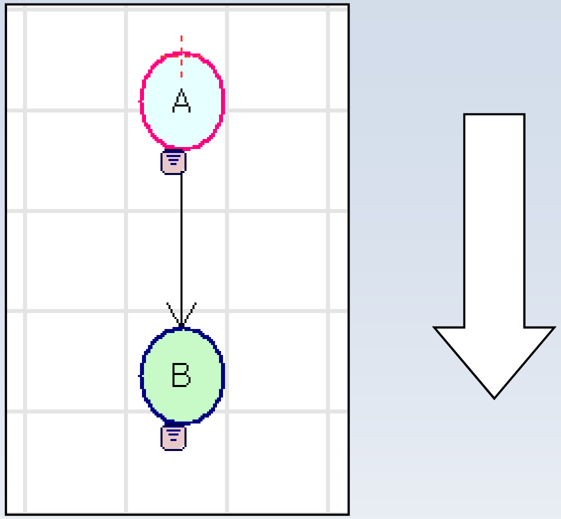
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A,B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

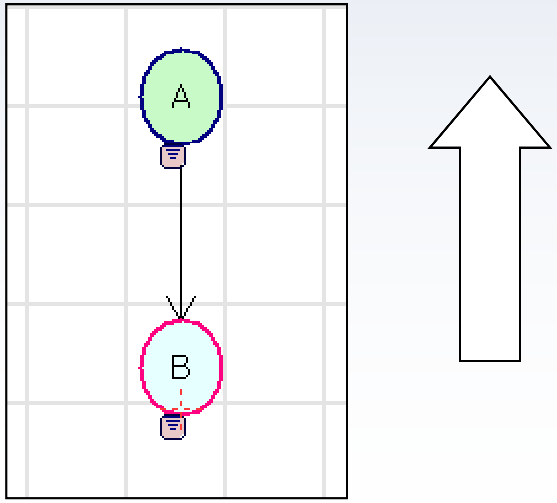
$$P(X|Y,E) = \frac{P(Y|X,E)P(X|E)}{P(Y|E)}$$

Induction (bottom-up reasoning):

- Conclude from an observed event to all the possibly generating events in the structure (uncertain assumption)
- Bayes Theorem:

$$P(B|A) = \frac{P(A|B)P(B)}{\sum_{b} P(B=b)P(A|B=b)}$$



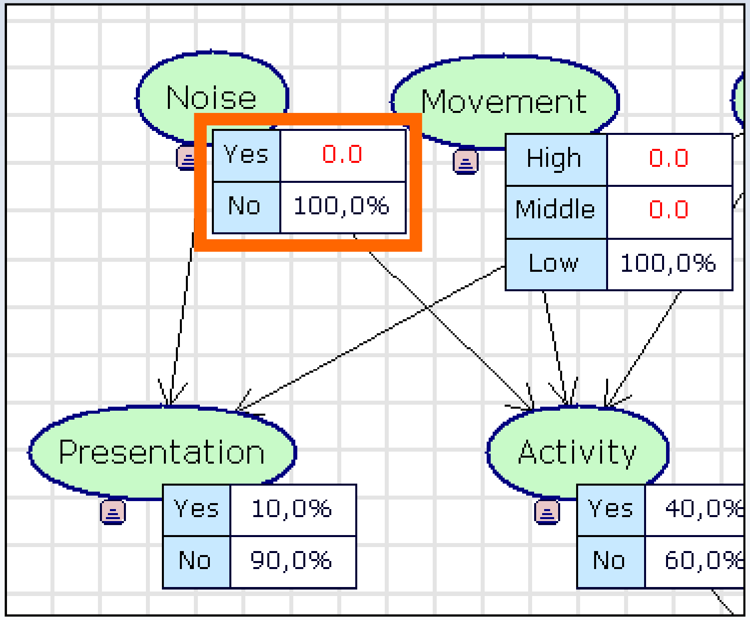


2. Introduction: Bayesian Networks

Querying the Bayesian Network; Inference [Koch] [Hackerman] [Murphy]:

- Definition: input variables, discrete values, mapping of continuous values
- Probabilistic inference
- Input: set node property values to true or false (or discrete value)
- Observe the changing effects in the network

- Methods:
  - Exact methods: use graph structure
  - Statistical methods: Gibbs Sampling, Markov Chains Monte Carlo, other Monte Carlo calculations



2. Introduction: Bayesian Networks

Compactness of the network/Inference [Koch] [Hackerman] [Murphy]: :

- In each probabilistic network we can calculate the resulting full joint distribution of a graph with complete edge occupation

$$P(X_1, X_2, \ldots, X_m) = \prod_{i=1}^{n} P(X_i \mid X_{i+1}, \ldots, X_m)$$

$$P(X_i \mid X_{i+1}, \ldots, X_m) = P(X_i \mid Parents(X_i))$$

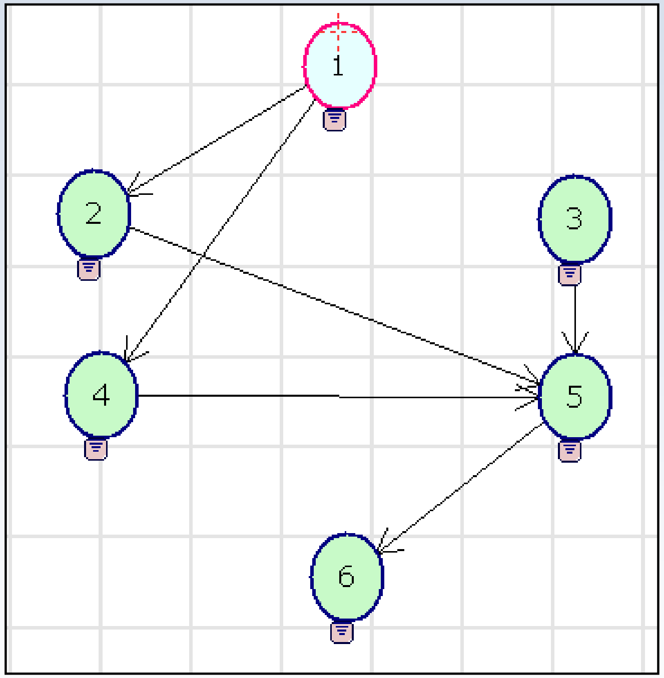
$$\Rightarrow P(X_1, X_2, \ldots, X_m) = \prod_{i=1}^{n} P(X_i \mid Parents(X_i))$$

In Bayesian Networks

- This is called the compactness in Bayesian Networks; Markov Condition
- We also need less values in the probability tables

2. Introduction: Bayesian Networks

Example:

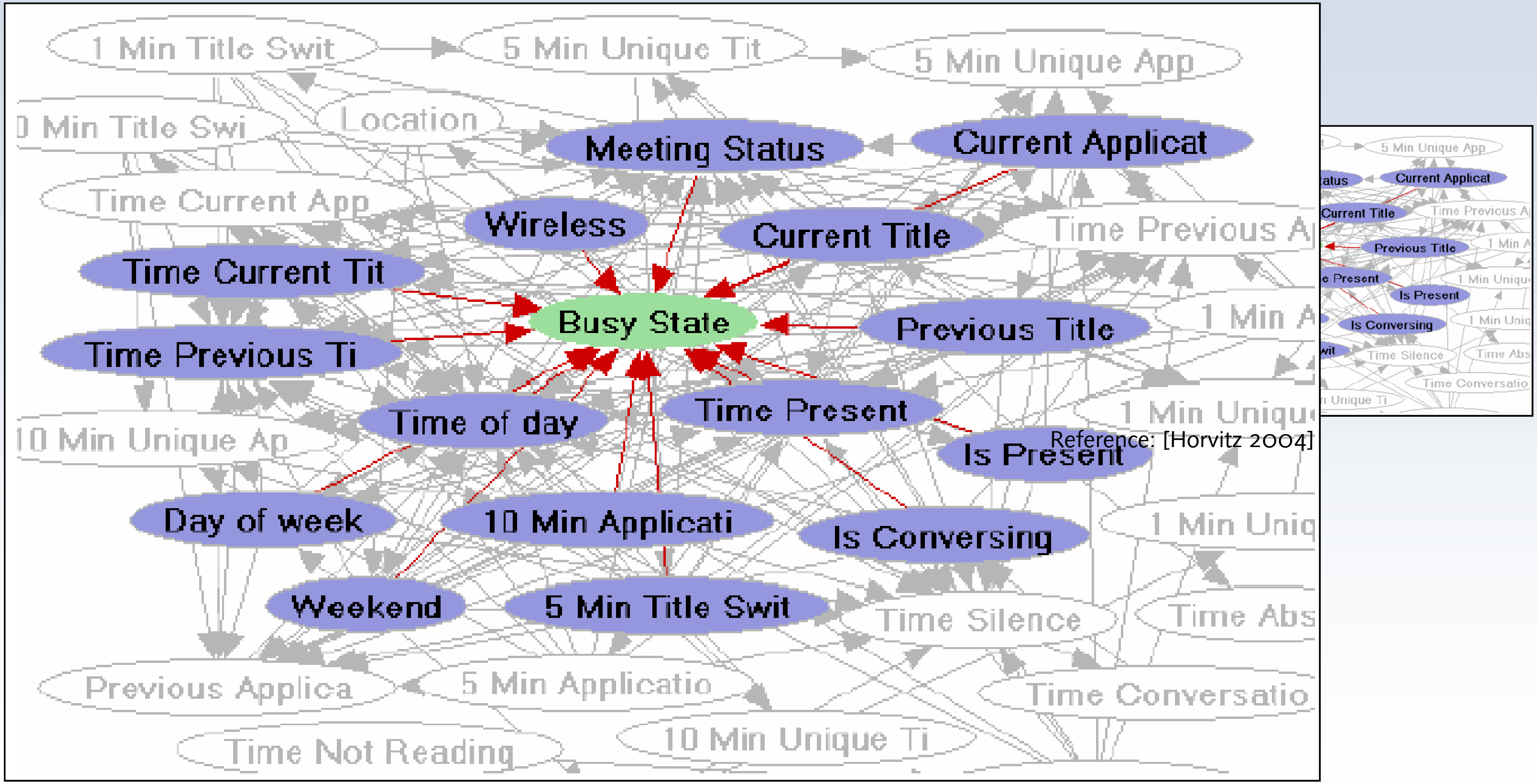


$$P(X_1, X_2, \ldots, X_m) = \prod_{i=1}^{n} P(X_i \mid Parents(X_i))$$

$$P(N_1, N_2, N_3, N_4, N_5, N_6) =$$
$$P(N_1) \cdot P(N_2 \mid N_1) \cdot P(N_4 \mid N_1) \cdot$$
$$P(N_5 \mid N_2, N_3, N_4) \cdot P(N_6 \mid N_5)$$

⇕

2. Introduction: Bayesian Networks



Reference: [Horvitz 2004]

2. Introduction: Bayesian Networks

Learning in Bayesian Networks [Pearl, Russell]:

- Local score metrics: Optimization problem where a quality measure of the network structure needs to be maximized
- Update conditional probabilities using gradient-based EM methods (Lauritzen, Binder)
- Works similar to the adjustment of weighting factors in a neuronal networks
- Learning structure of the network: trade off network complexity to fit the data (Friedman)

2. Introduction: Bayesian Networks

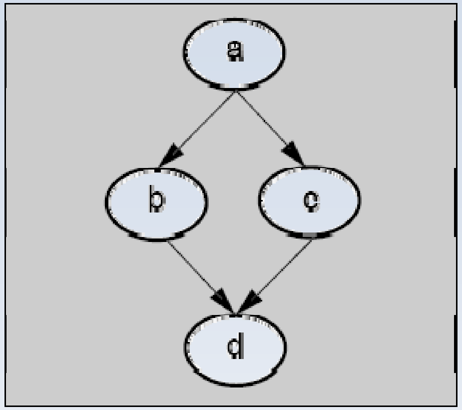


Figure a), Reference: [Charniak]

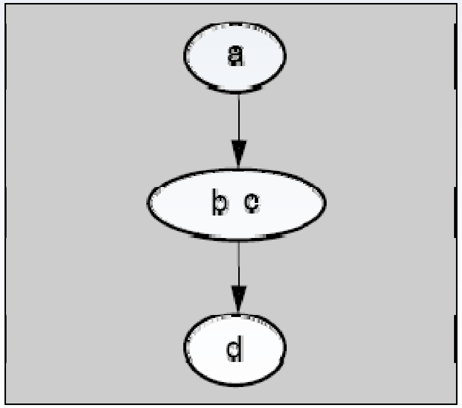


Figure b), Reference: [Charniak]

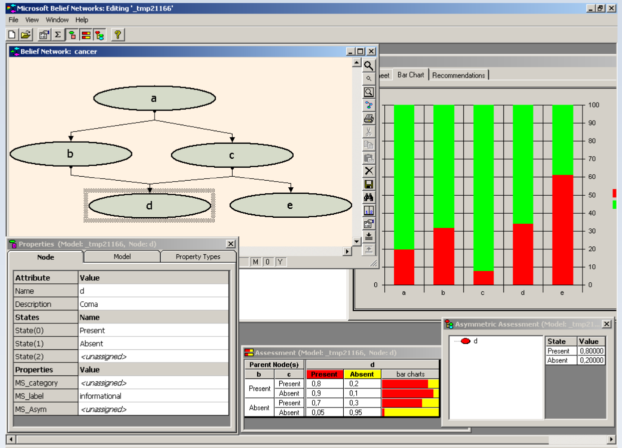
Optimization example –
Multiply connected networks:

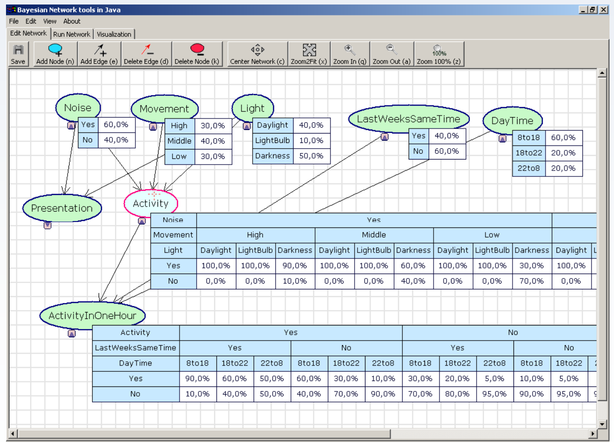
- Several ways to perform this task
- Clustering technique:
  - Combination of nodes until the resulting network graph is singly connected (Fig. b)
  - Values of node [b c] are the cross product of the b and c values
  - There are also mathematical calculations to determine the values of b and c again
  - Lauritzen, Spiegelhalter ('88), Jensen ('89)
- Various other algorithms try to simplify the graph structure to apply the common Bayesian Network solution finding algorithms

3. Using Bayesian Networks

Toolkits

- Microsoft MSBNx [MSBNx b]
  - XML, COM API, ActiveX Controls, GUI Editor
- JavaBayes [JavaBayes]
- jBNC [jBNC]
- Bayesian Networks in Java [BNJava], WEKA
  - Kansas State University
  - 8 network formats: XML-BIF (Bayesian Network Interchange Format), XBN, Ergo ENT, etc.
  - Data formats: Excel, WEKA, etc.
  - Exact and approximate inference algorithms



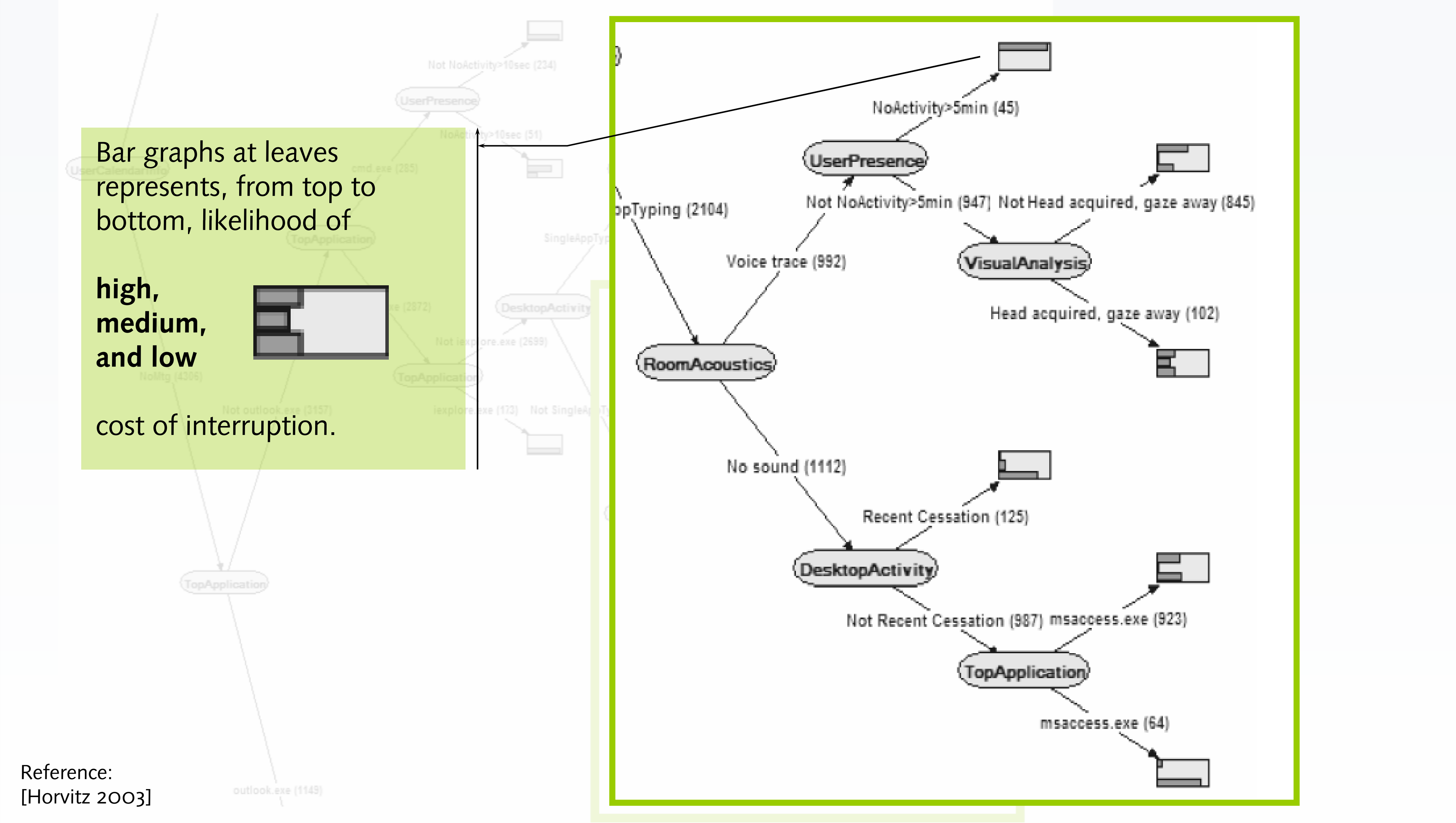


3. Using Bayesian Networks

Bar graphs at leaves represents, from top to bottom, likelihood of

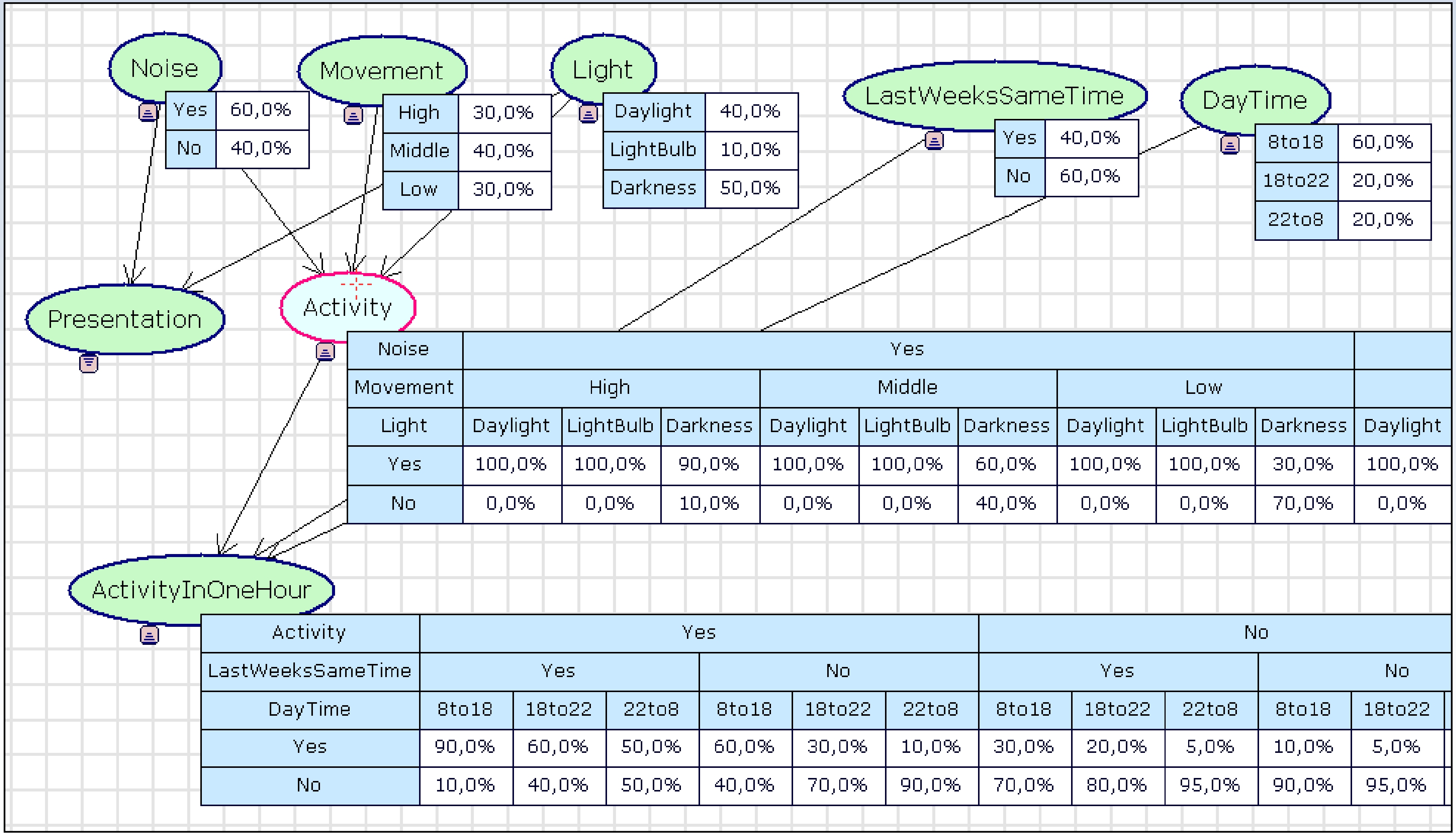
high, medium, and low

cost of interruption.



Reference:
[Horvitz 2003]

3. Using Bayesian Networks



Noise	
Yes	60,0%
No	40,0%

Movement	
High	30,0%
Middle	40,0%
Low	30,0%

Light	
Daylight	40,0%
LightBulb	10,0%
Darkness	50,0%

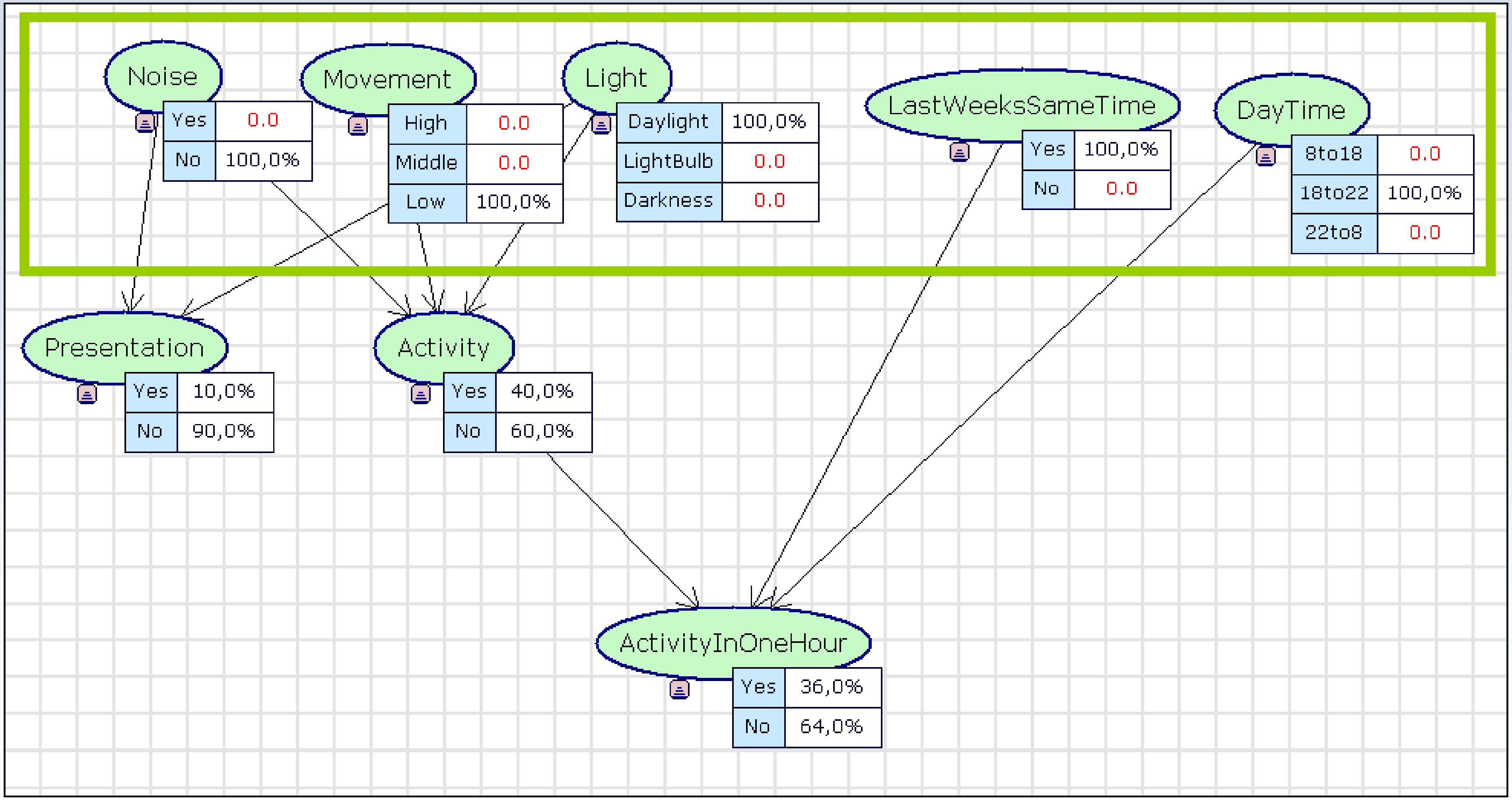
LastWeeksSameTime	
Yes	40,0%
No	60,0%

DayTime	
8to18	60,0%
18to22	20,0%
22to8	20,0%

Noise	Yes									
Movement	High			Middle			Low			
Light	Daylight	LightBulb	Darkness	Daylight	LightBulb	Darkness	Daylight	LightBulb	Darkness	Daylight
Yes	100,0%	100,0%	90,0%	100,0%	100,0%	60,0%	100,0%	100,0%	30,0%	100,0%
No	0,0%	0,0%	10,0%	0,0%	0,0%	40,0%	0,0%	0,0%	70,0%	0,0%

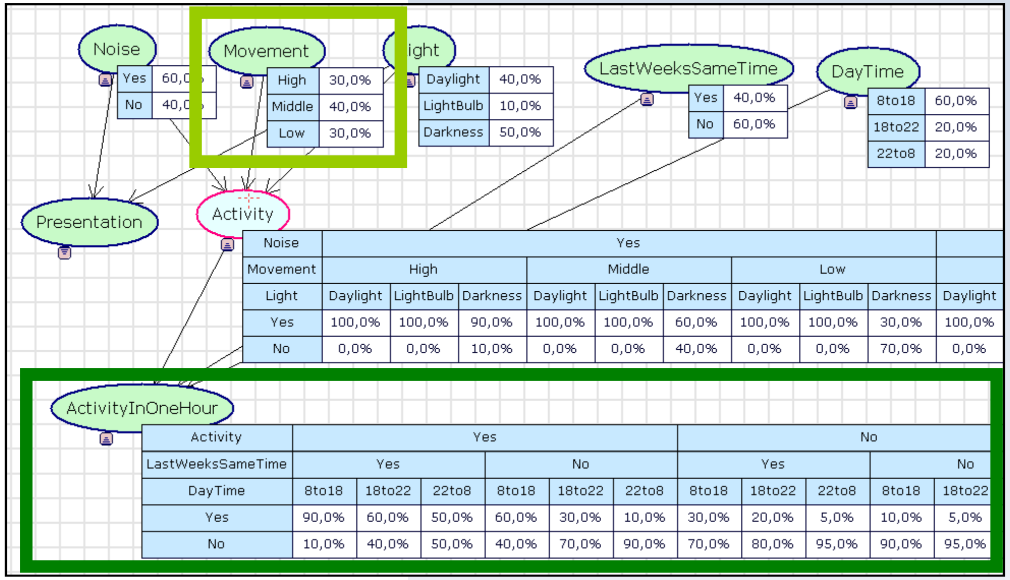
Activity	Yes						No				
LastWeeksSameTime	Yes			No			Yes			No	
DayTime	8to18	18to22	22to8	8to18	18to22	22to8	8to18	18to22	22to8	8to18	18to22
Yes	90,0%	60,0%	50,0%	60,0%	30,0%	10,0%	30,0%	20,0%	5,0%	10,0%	5,0%
No	10,0%	40,0%	50,0%	40,0%	70,0%	90,0%	70,0%	80,0%	95,0%	90,0%	95,0%

3. Using Bayesian Networks



3. Using Bayesian Networks

```
<?xml version="1.0" encoding="US-ASCII" ?>
- <!--
   Bayesian network in XMLBIF v0.3 (BayesNet Interchange Format)
   Produced by BNJ 3.0 (http://bndev.sourceforge.net/
  -->
  <!-- DTD for the XMLBIF 0.3 format  -->
  <!DOCTYPE BIF (View Source for full doctype...)>
- <BIF VERSION="0.3">
  - <NETWORK>
      <NAME>CarStarts</NAME>
    - <VARIABLE TYPE="nature">
        <NAME>Noise</NAME>
        <OUTCOME>Yes</OUTCOME>
        <OUTCOME>No</OUTCOME>
        <PROPERTY>position = (-438,-272)</PROPERTY>
      </VARIABLE>
    - <VARIABLE TYPE="nature">
        <NAME>Movement</NAME>
        <OUTCOME>High</OUTCOME>
        <OUTCOME>Middle</OUTCOME>
        <OUTCOME>Low</OUTCOME>
        <PROPERTY>position = (-137,-268)</PROPERTY>
      </VARIABLE>
    - <DEFINITION>
        <FOR>ActivityInOneHour</FOR>
        <GIVEN>Activity</GIVEN>
        <GIVEN>LastWeeksSameTime</GIVEN>
        <GIVEN>DayTime</GIVEN>
        <TABLE>0.9 0.1 0.6 0.4 0.5 0.5 0.6 0.4 0.3 0.7 0.1 0.9 0.3 0.7 0.2 0.8 0.05 0.95 0.1 0.9 0.05 0.95 0.05 0.95</TABLE>
      </DEFINITION>
    - <DEFINITION>
        <FOR>Presentation</FOR>
        <GIVEN>Noise</GIVEN>
        <GIVEN>Light</GIVEN>
        <TABLE>0.1 0.9 0.2 0.8 0.8 0.2 0.1 0.9 0.1 0.9 0.05 0.95</TABLE>
      </DEFINITION>
    - <DEFINITION>
        <FOR>DayTime</FOR>
        <TABLE>0.6 0.2 0.2</TABLE>
      </DEFINITION>
    </NETWORK>
  </BIF>
```



Software Demonstration I: Bayesian Network

JavaBayes: Network Example

Start tool:

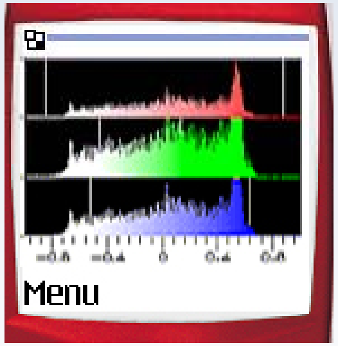
4. Mobile Interface, Visualization

4. Mobile Interface

The application scenarios (last week):

1. Awareness information
2. Connect context information with the availability service of mobile phones
3. Room profile
4. Notification service, threshold

5. Availability/usage patterns
6. Workgroup scheduling assistance (because of working room awareness)



4. Mobile Interface, Scenario

Requests of (mobile) clients:

(this list will be extended; it is an first attempt for the Web Service methods)

1. Current activity and prospective activity in the lab for the next three hours (transmission: 30 minute slots, 1+6 values, probabilities, int 00 - 99)
2. Current probabilities for some kind of events:
   working, telephone, face-to-face, meeting, presentation, discussion
   (transmission: 6 values with probabilities, or 1 value with an index of the most probably activity at the moment)
3. Room profile of the day/week (half-hour slots, 48 values)
4. Activity in a public room at the moment and the next hour (5-minute slots, 12 values, probabilities 00-99)

4. Mobile Interface

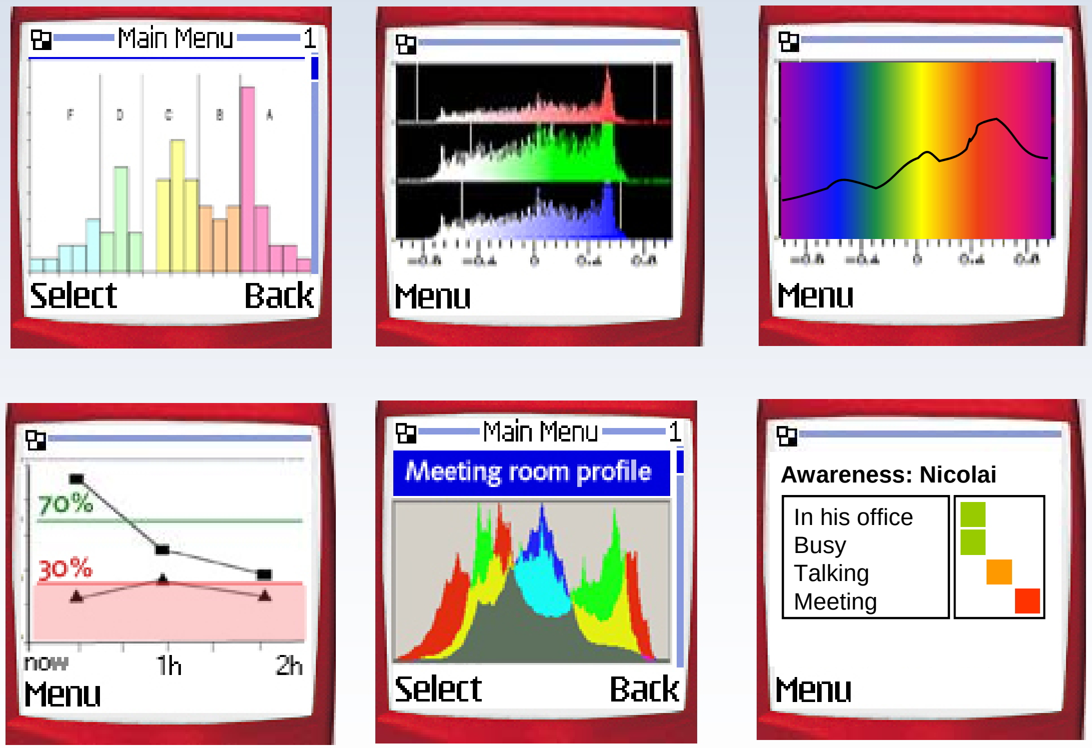
Display awareness information on mobile devices (J2ME Platform):

- Text or graphic usage depends on information volume
- Compact information on one screen
- Detail on demand; from general to specific
- Overview and context
- Histogram-like graphs (Begole, SUN Research)

→ Graph plot engine for J2ME, Library

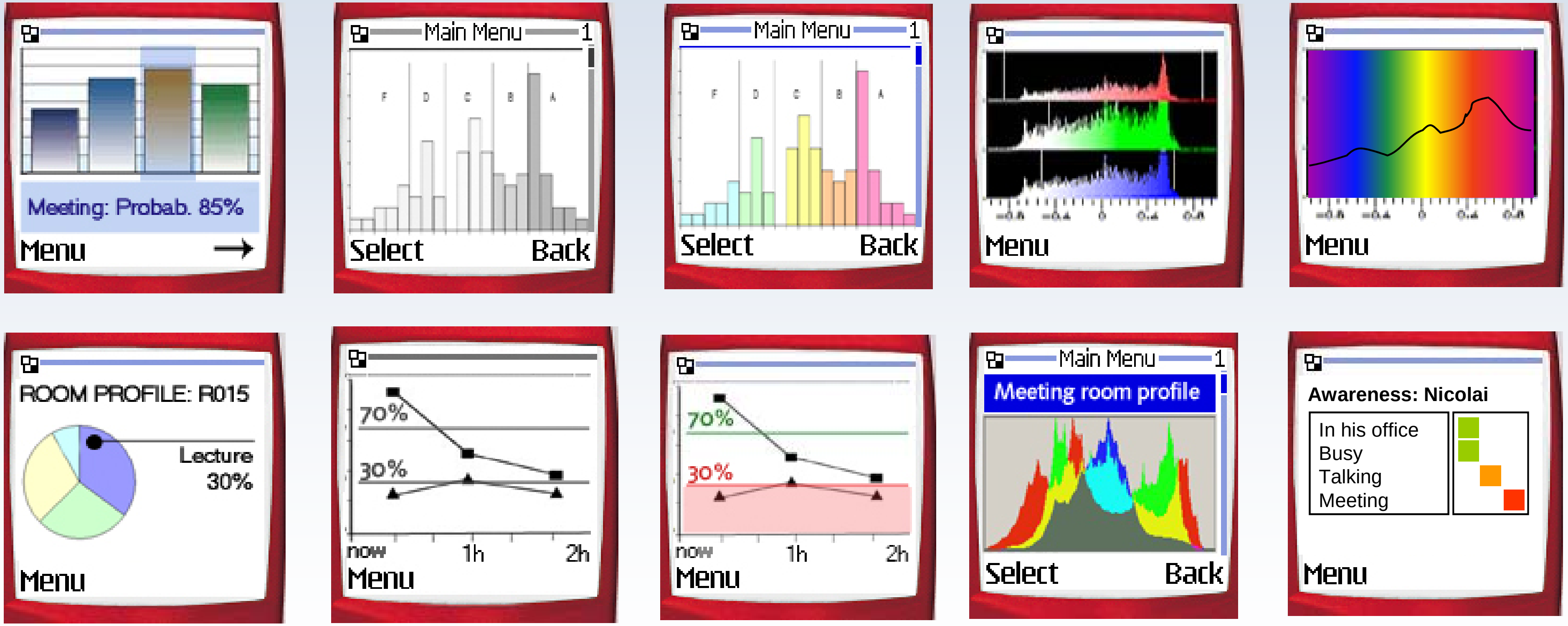
→ Display Manager

Design concepts for awareness visualization (not yet implemented):



4. Mobile Interface

Design concepts for awareness visualization (not yet implemented):



5. IR Control, Interface

5. IR Control, Interface

The acquired data:

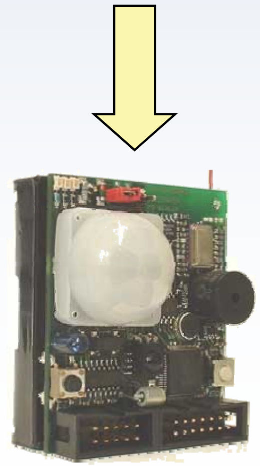
Sensor data: temperature, movement, vibration, light intensity	⇨	Anonymous information: location, rooms, area, office
Interaction with devices, computer, software, IM, etc.	⇨	Personalized: software usage patterns, working time with software
RFIDs	⇨	Personalized: exact location, availability time at locations, rooms
Special interaction device: Hardware controls, PDA Interface	⇨	Meta information: current project, group members, override heuristics, etc.

5. IR Control, Interface

– **First approach: Use an handheld device (Palm) to interact with the ESB IR receiver/transmitter.**

– Infrared Communication: ESB and protocol
  – Uses RC5 remote protocol
  – IR Codes: RC5 (Phillips), Sony Pulse-Width, Sony 12, Sony 15, Daewoo, Panasonic, Onkyo, JVC
  – Instruction set: 2048 commands, divided into 32 addresses with each 64 instructions
  – RC5 Technical specifications: [De Vleeschauwer] and the ESB websites [ESB IR] [ESB Term]
  – Link collection for IR: [Engdahl]

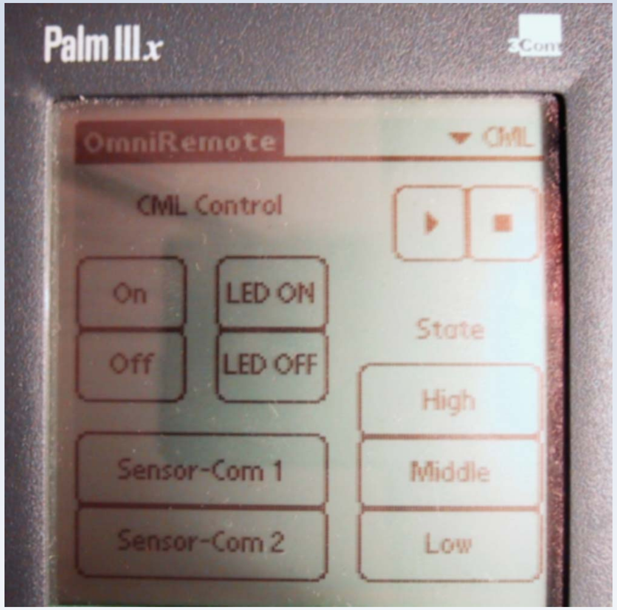
→ IR development for palm size devices: PPC [DevNet] and Palm [PalmSource], devloper kits (C++, eVC, eVB)



5. IR Control, Interface

- **Prototype using existing Palm remote software, training mode for application to learn RC5 commands**

- Communication with the ESB
  - ESB commands: `sir xxxx`, `rir`
  - Address 00 + Command: 0001 – 000F
  - Combination of two bytes (hex coded)
  - These commands are used for Palm <> PC communication



  - Received IR bytes available via `rir, rsr` and the polling mode `saf 32` → Parser

  - [IR: C(3) A(0)] →
    `\\[IR:\\sC\\([0-9]*\\)\\sA\\([0-9]*\\)\\]`

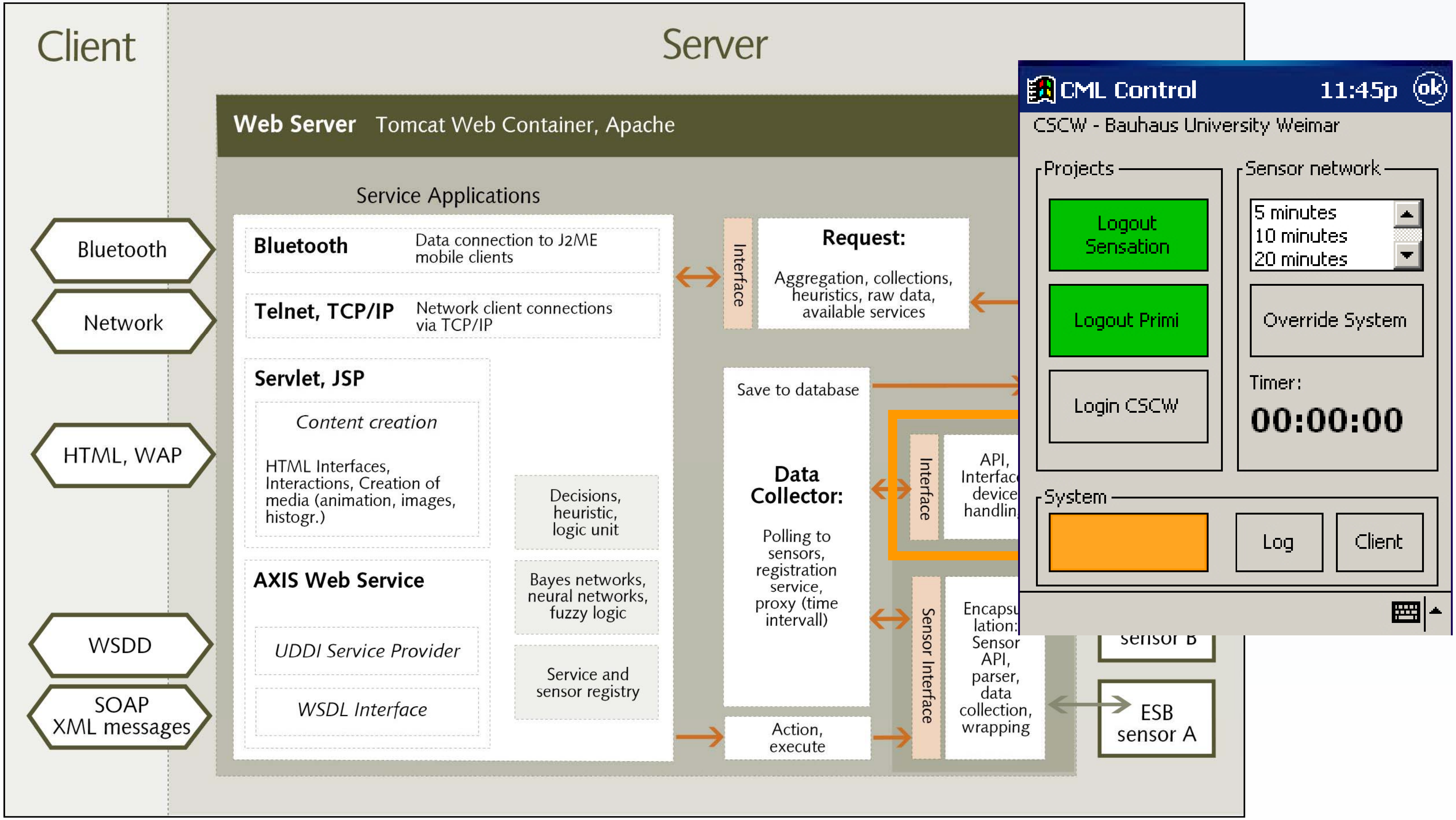
5. IR Control, Interface

Software Demonstration II:
IR Control, Parser

Ecplise: Sensor Project, IR Control

Start tool:

5. IR Control, Interface



ToDo

Assignments for the next week:

1. Working with various Bayesian Networks, discover the different toolkits and API (group?)
2. Extend the console application for ESB
3. Bluetooth tests (Nokia ⇔ Apple)
4. Tests for the J2ME mobile graph toolkit, especially to display awareness information in CSCW and UbiComp context

Furthermore:

– Prototype of interaction interface (PocketPC); if we need such an application
– “Sensor” Plug-in and (W)LAN communication

Literature, References

Bayesian Networks Literature:

[Murphy] **Kevin Murphy: A Brief Introduction to Graphical Models and Bayesian Networks, MIT, 1998**
http://www.ai.mit.edu/~murphyk/Bayes/bayes.html (last visited: 15.11.2004)

[Pearl, Russell] **Judea Pearl and Stuart Russell: Bayesian Networks**
Technical Report, November 2000, to appear in M.Arbib (Ed.), Handbook of
Brain Theory and Neural Networks, MIT Press, 2001

[Charniak] Eugene Charniak: Bayesian Network without Tears, The American Association for Artificial Intelligence, http://www.aaai.org (last visited: 15.11.2004)

[Congdon] Peter Congdon: Bayesian Statistical Modelling,
John Wiley & Sons, Ltd., Chichester, West Sussex, 2001

[Koch] Karl-Rudolf Koch: Einführung in die Bayes Statistik, Springer Verlag, Berlin, Heidelberg, 2000, Kapitel: 5.4. Bayes-Netze, S. 157ff, Kapitel: 3.1. Bayes-Strategie, S. 63ff

[Heinsohn] Jochen Heinsohn und Rolf Socher-Ambrosius: Wissensverarbeitung – Eine Einführung
Spektrum Akademischer Verlag, Heidelberg, Berlin, 1999

[Heckerman] David Hackerman: A Tutorial on Learning With Bayesian Networks
Microsoft Research, MSR-TR-95-06, March 1995
ftp://ftp.research.microsoft.com/pub/tr/tr-95-06.pdf (last visited: 14.11.2004)

[XMLBIF] The Interchange Format for Bayesian Networks (last visited: 14.11.2004)
http://www-2.cs.cmu.edu/afs/cs/user/fgcozman/www/Research/InterchangeFormat/

Literature, References

Toolkits for Bayesian Networks (API, Documentation):

[BNJava]	Bayesian Networks in Java http://bnj.sourceforge.net/, Kansas State University, Laboratory for Knowledge Discovery in Databases (KDD) (last visited: 15.11.2004)
[JavaBayes]	Java Bayes, Version 0.346 http://www-2.cs.cmu.edu/~javabayes/Home/index.html (last visited: 16.11.2004) University of Sao Paulo
[jBNC]	Bayesian Network Classifier Toolbox http://jbnc.sourceforge.net/ (last visited: 15.11.2004)
[Kadie et al.]	C.M. Kadie, D. Hovel, and E. Horvitz. MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks. Microsoft Research Technical Report MSR-TR-2001-67, July 2001.
[MSBNx a]	Microsoft Bayesian Network Toolkit, MSBNx Documentation http://research.microsoft.com/adapt/MSBNx/MSBNxManual.asp (last visited: 15.11.2004)
[MSBNx b]	Microsoft Bayesian Network, Download, Documentation, Information http://research.microsoft.com/adapt/MSBNx/ (last visited: 15.11.2004)

Literature, References

References: CSCW and Ubiquitous Computing:

[Horvitz 2004] Eric Horvitz et al.: BusyBody: Creating and Fielding Personalized Models of the Cost of Interruption, CSCW '04, November 6-10, 2004, Chicago, ACM

[Horvitz 2003] Eric Horvitz and Johnson Apacible: Learning and Reasoning about Interruption, Microsoft Research, ICMI'03, November 5–7, 2003, Vancouver, British Columbia, Canada

[Begole 2004] James Begole et al.: Lilsys: Sensing Unavailability, Sun Microsystems Research, CSCW '04, November 6-10, 2004, Chicago, ACM

[Chen, Kotz 02a] Guanling Chen, David Kotz: Context Aggregation and Dissemination in Ubiquitous Computing Systems, Dartmouth Computer Science Technical Report, Dartmouth College, 2002, In Proceedings of the Fourth 1EEE Workshop on Mobile Computing Systems and Applications. IEEE Computer Society Press, June 2002

[Beigl et al.] Michael Beigl, et al.: Typical Sensors needed in Ubiquitous and Pervasive Computing Telecooperation Office, University of Karlsruhe

[Schmidt et al.] Albrecht Schmidt, et al.: There is More to Context than Location: Environment Sensing Technologies for Adaptive Mobile User Interfaces", Proceedings of Workshop on Interactive Applications of Mobile Computing (IMC'98), Rostock, 1998

[Jonsson, Mattsson] M. Jonsson and J. Mattsson. *Building extendable room based sensor clusters for ubiquitous computing environments*. Technical Report 2002-019, DSV, Stockholm University, 2002

[Hong, Landay] Jason I. Hong and James A. Landay: An Infrastructure Approach to Context-Aware Computing, University of California at Berkeley

Literature, References

References: IR communication, ESB, PDA programming and regular expressions:

[ESB IR]	ESB: recir.c description http://www.inf.fu-berlin.de/inst/ag-tech/scatterweb_net/ESB/sensorboards/doc/html/recir_8c.html (last visited: 8.11.2004)
[ESB Term]	C++ Terminal Documentation terminal.c, terminal.h http://www.inf.fu-berlin.de/inst/ag-tech/scatterweb_net/ESB/sensorboards/doc/html/terminal_8c.html (last visited: 8.11.2004)
[De Vleeschauwer]	David De Vleeschauwer: Information about Infrared codes, I2C bus, pic16F84 kits, (last visited: 8.11.2004) http://users.pandora.be/davshomepage/
[Engdahl]	Tomi Engdahl: Infrared remote control technology http://www.epanorama.net/links/irremote.html (last visited: 10.11.2004)
[DevNet]	PocketPC Developer Network, Spb Software, http://www.pocketpcdn.com/sections/irdaled.html, and: http://www.pocketpcdn.com/sections/irdaled.html (last visited: 9.11.2004)
[PalmSource]	Palm OS Programmer's Companion: Beaming via IR http://www.palmos.com/dev/support/docs/palmos/Beaming.html
[Jakarta RegExp]	Apache Project: Jakarta Regular Expression Version 1.3 http://jakarta.apache.org/regexp (last visited: 10.11.2004)
[JakartaRE Syntax]	Apache Project: Class RE, (last visited: 8.11.2004) http://jakarta.apache.org/regexp/apidocs/org/apache/regexp/RE.html
[Darwin]	Ian Darwin: Java Cookbook, First Edition, June 2001, Chapters 4 – Pattern Matching and Regular Expressions, Chapter 11 – Programming Serial and Parallel Ports

Thank You

For Your Attention!