Educational Technologies
WS2006

Student Modeling

Source: Erica Melis

Approximate Plan of the Course

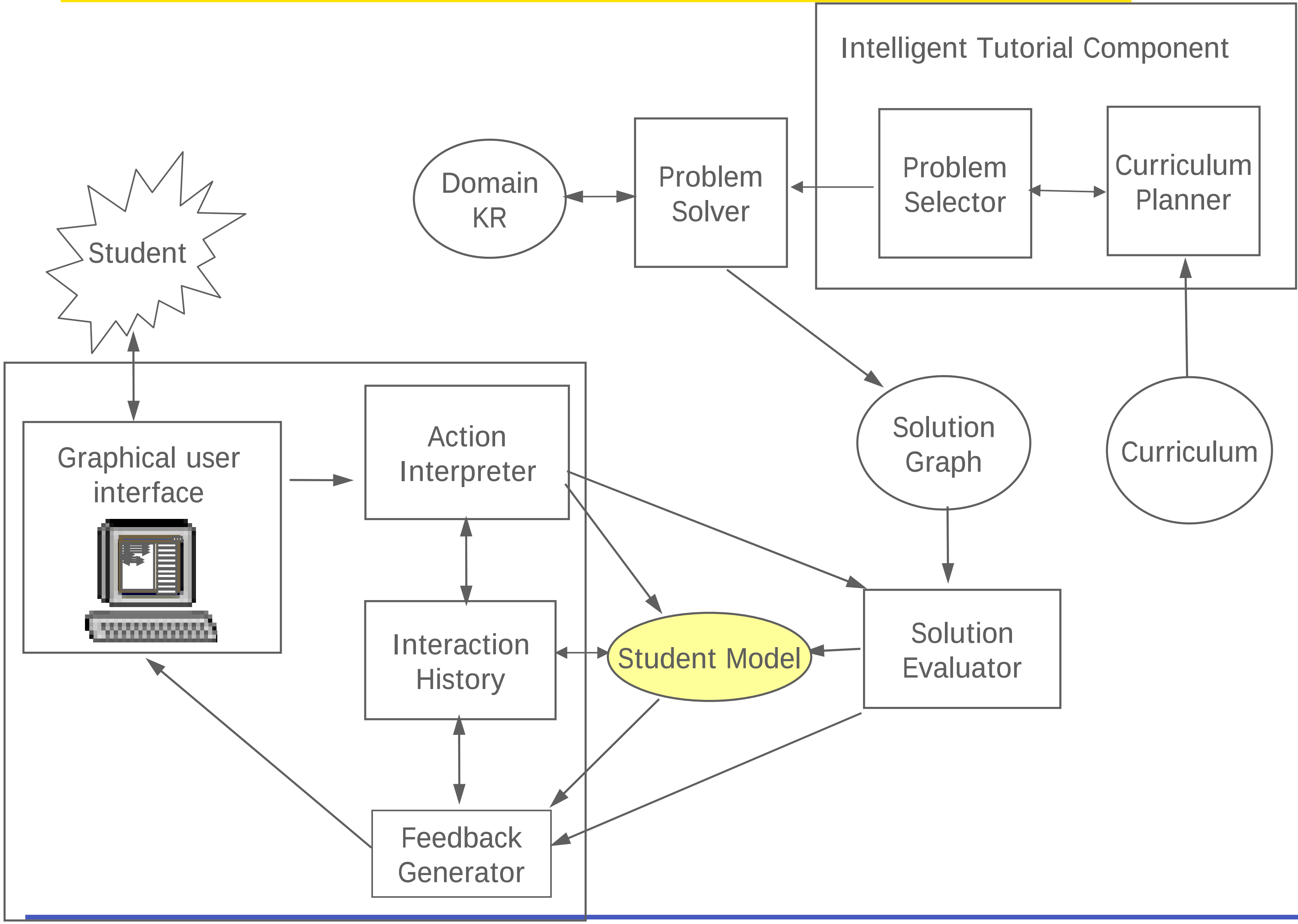
18.10. Introduction
25.10. XML- Knowledge Representation
8.11. Student Modelling
15.11. Web technologies and security
22.11. Tutorial Planning and instructional design
29.11. Media Principles
6.12. Interactive exercises
13.12. Authoring tools, CTAT
20.12. Diagnosis: model tracing and domain reasoning

10.1. Diagnosis: constraint based
17.1. Tutorial dialogues
24.1. Action analysis and Machine Learning techniques
31.1. Cognitive tools
7.2. Meta-cognitive support
14.2. student projects

Student modeling what for?

- **Domain knowledge/skill prediction** (adaptation, selection)
- **Knowledge assessment**
- **Plan recognition (suggestions)**
- **Attitude prediction and reaction**
- **Motivation monitoring**

Generic Architecture



Adaptation dimensions

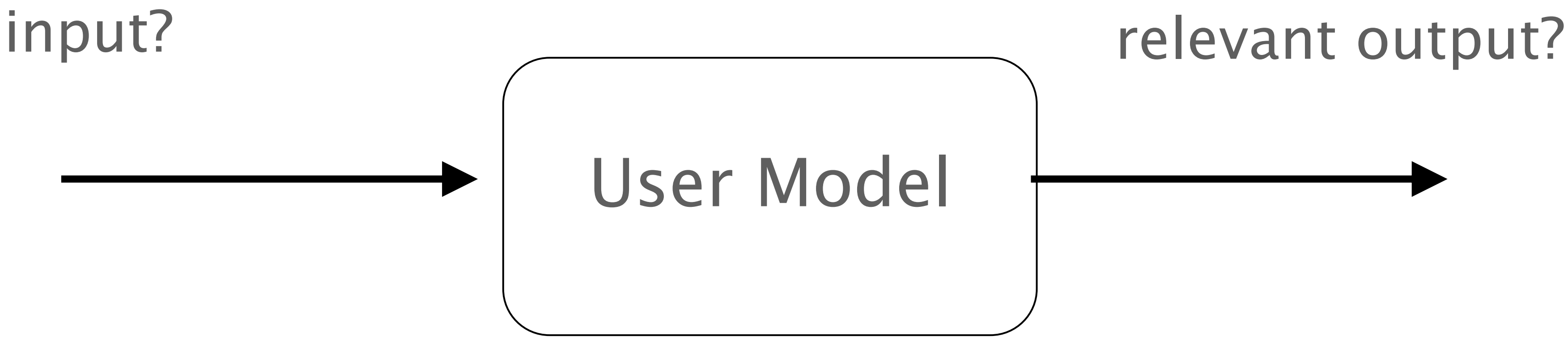
- **Technical environment**
- **Social context**
- **Personalization: difficulty…**

- **Situation**

- **Our first experience:**

  School type, field, language, goals, prerequisites, field of interest, mastery level, competencies

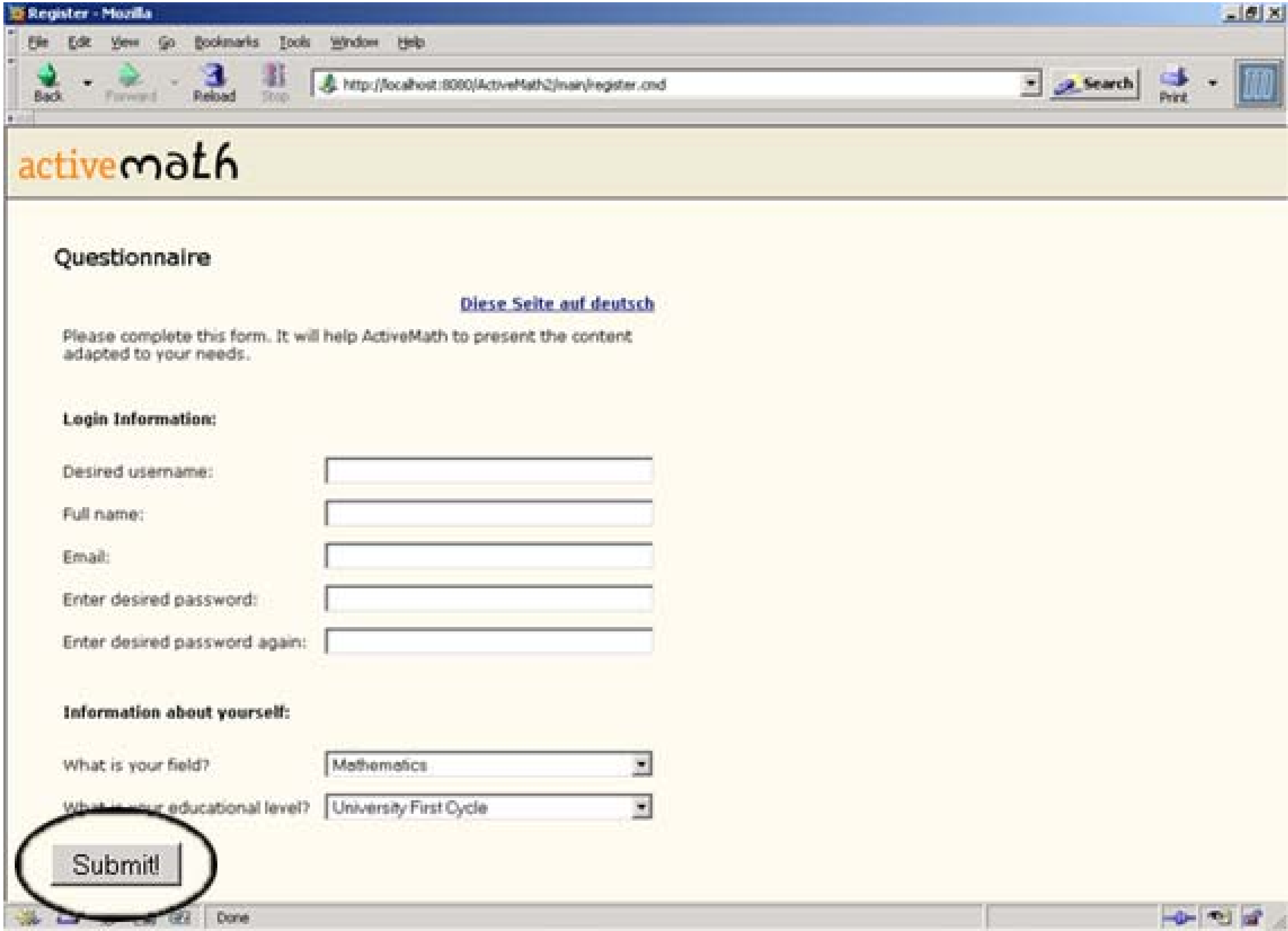
Systematic approach



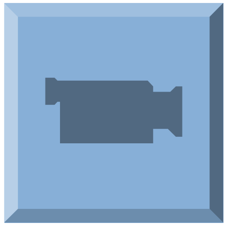
Inputs for User Model

- **Self-reports:**
  - Questionnaire
  - on-task-report ☺
  - Open Learner Model
- **teacher input**
- **behavioral data**
  - Performance
  - Number of tasks finished
  - response to test items (hesitation)
  - time-on-task
  - # help or other requests
  - low-level measures
- **context data**

Input: Online Questionnaires



My Profile – change My Data



Output: (Sub)competencies

- Solve
- Argue
- Think mathematically
  - Formulate
  - Generalize
- Model
  - Encode
  - Decode
- Represent
- Communicate
- Tools

- Knowledge
- Understanding
- Application
- Analysis
- Synthesis

Output: Individual Variables

- **Cognitive (actual)**
  - ▷ Capability, incoming knowledge
  - ▷ Cognitive style *
  - ▷ Problem solving strategies, preferences
- **Psychological (mental)**
  - ▷ Interests
  - ▷ Preferences , learning style*
  - ▷ Meta-cognitive *
- **Person-in-situation**
  - ▷ Exploratory behaviour
  - ▷ Goals
- **Affective**
  - ▷ motivation, emotion
- **Personal traits**
  - ▷ Blind, working memory capacity, attention span, reading performance

Output: Context variables

- **Social context:**
  - ▷home/classroom/museum
  - ▷Collaboative/single
  - ▷Country (geographic)
  - ▷Curriculum
  - ▷school type
  - ▷cultural (not only language)

- **Technical context** (CC/PP rdf compositeCapabilities/preferenceProfile)
  - ▷Browser (rendering)
  - ▷Hardware platform: display-size, bandwidth PC/PDA
  - ▷Software platform, Availability of tools on client, Handwriting facility

Student Modeling, techniques

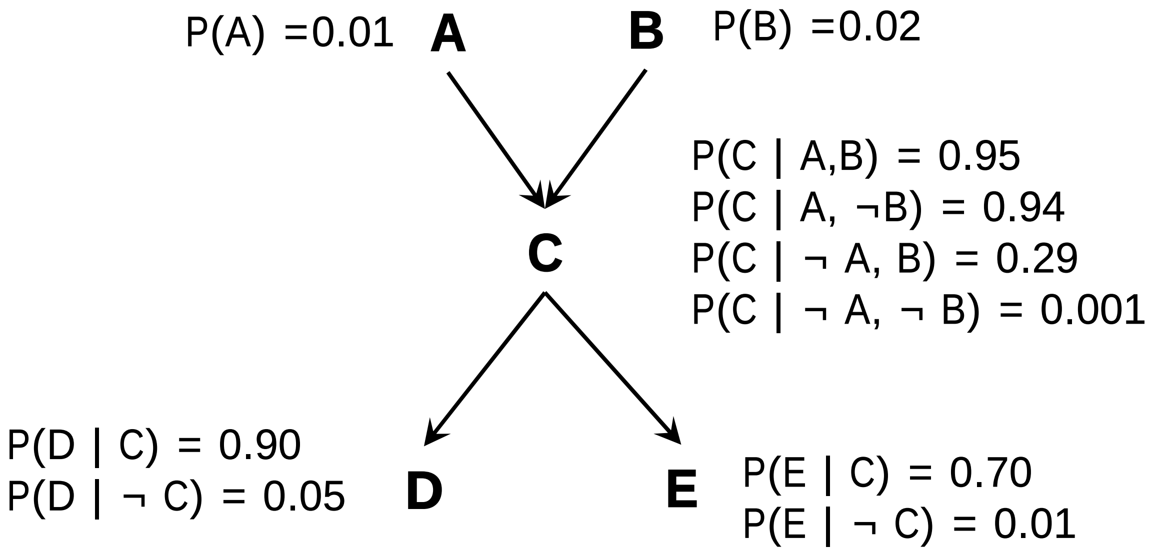
- Case-based
- Data-driven
- Baysian networks
- Bayesian updating
- Dynamic Bayesian Networks
- Stereotypes (deduction)
- Functional heuristic updating
- Social voting

knowledge tracing

model tracing

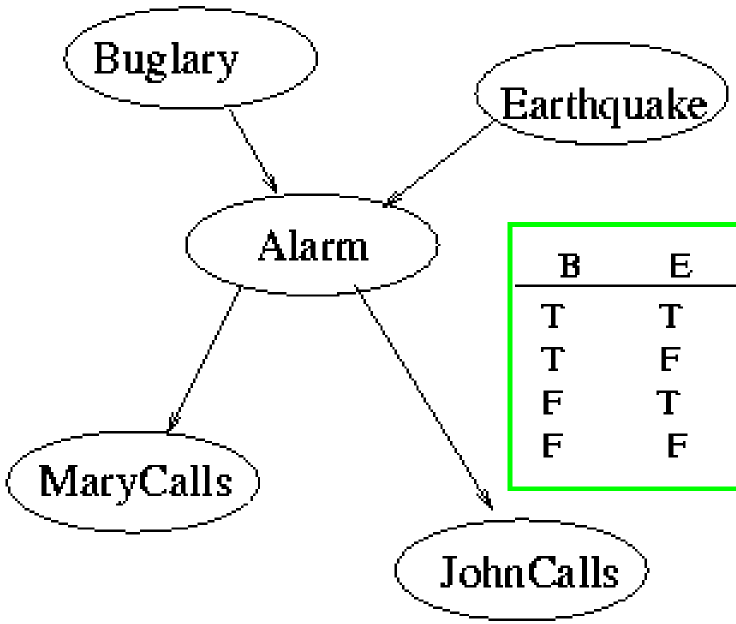
Bayesian Networks

- **Nodes: set of random variables $X_1\ X_2\ ..\ X_n$**
- **Links: probabilistic dependencies among variables**
- **Conditional probabilities: quantify the dependencies**



Bayesian Networks

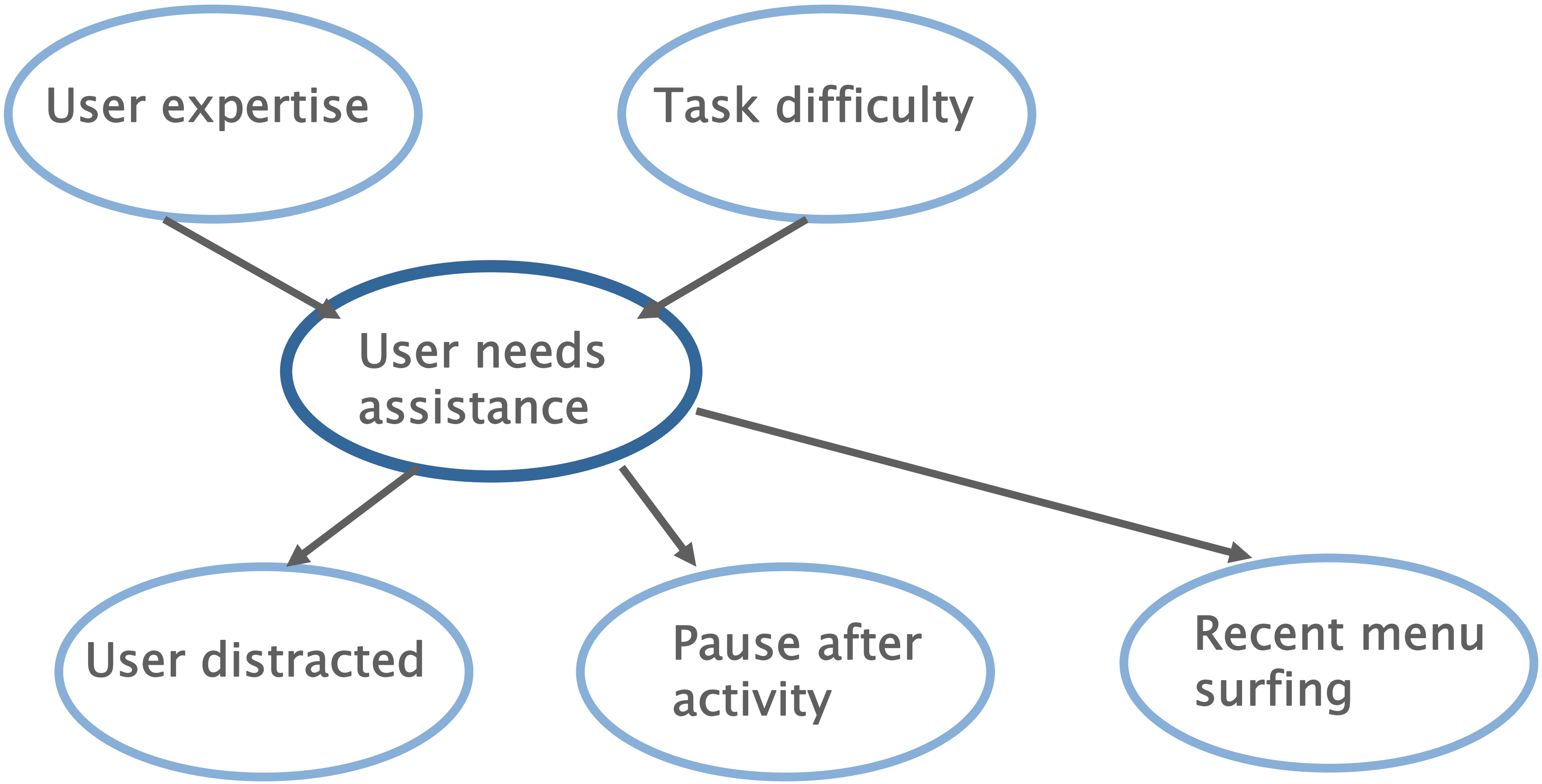
Probability distribution
events, causes, evidences
diagnostic/causal update
handle noise and uncertainty



B	E	P(A)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

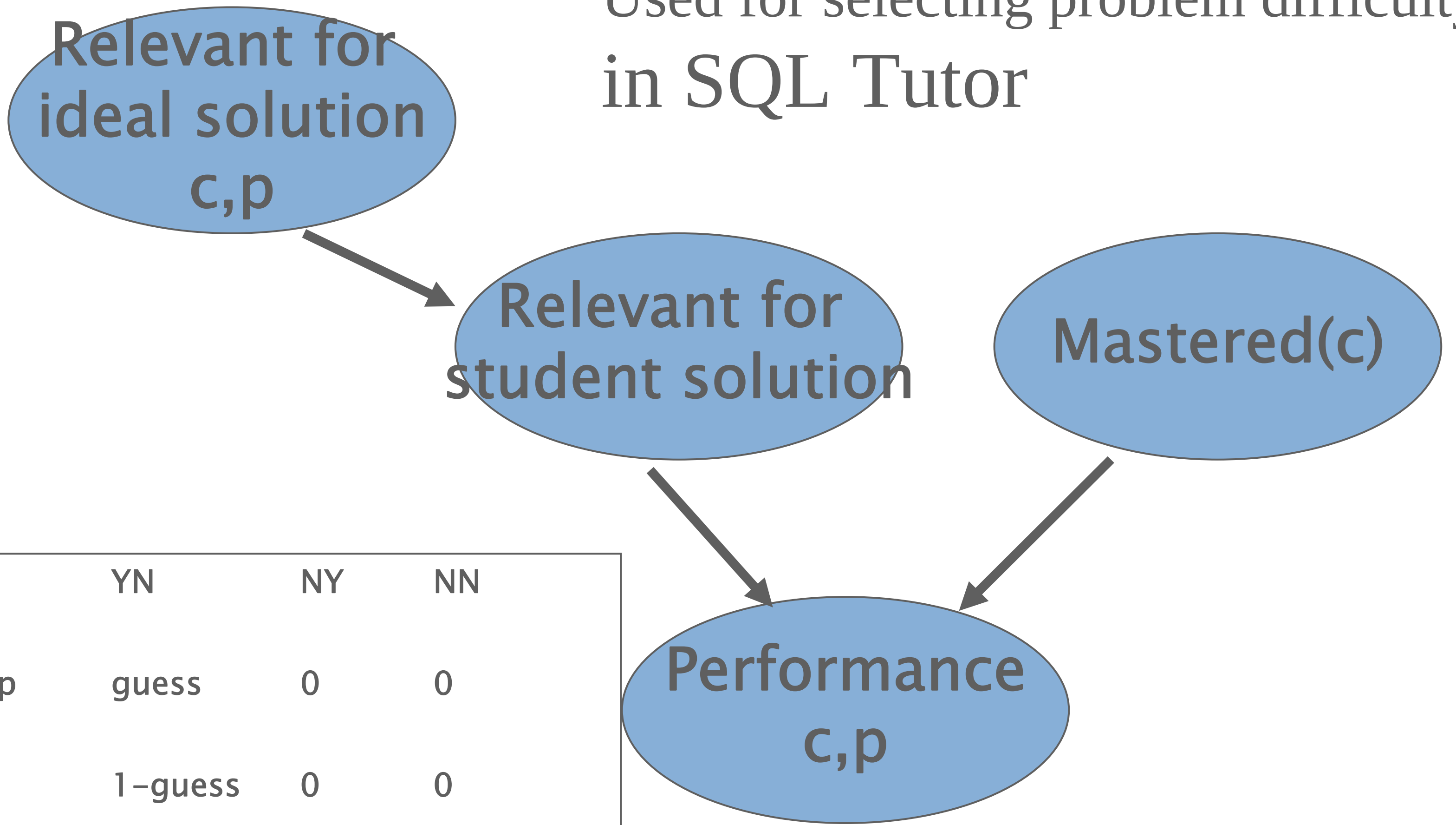
Uncertainty in cognitive modeling
Attractive alternative to modeling human expertise
Software packages (BNT, DBN)
In general NP-hard...

Bayesian Networks in **Lumiere**



BN for predicting student performance

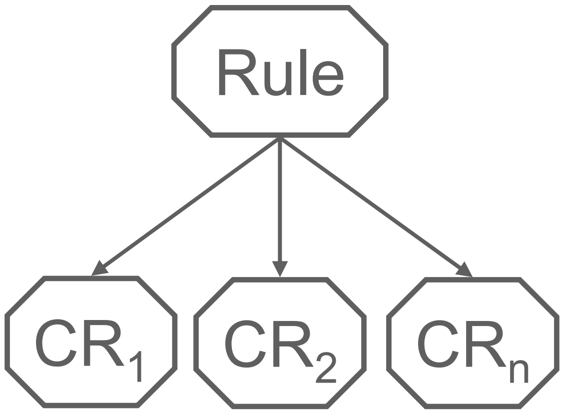
Used for selecting problem difficulty in SQL Tutor



	YY	YN	NY	NN
good	1-slip	guess	0	0
bad	slip	1-guess	0	0
not relev	0	0	1	1

Andes1 Network structure: Static part

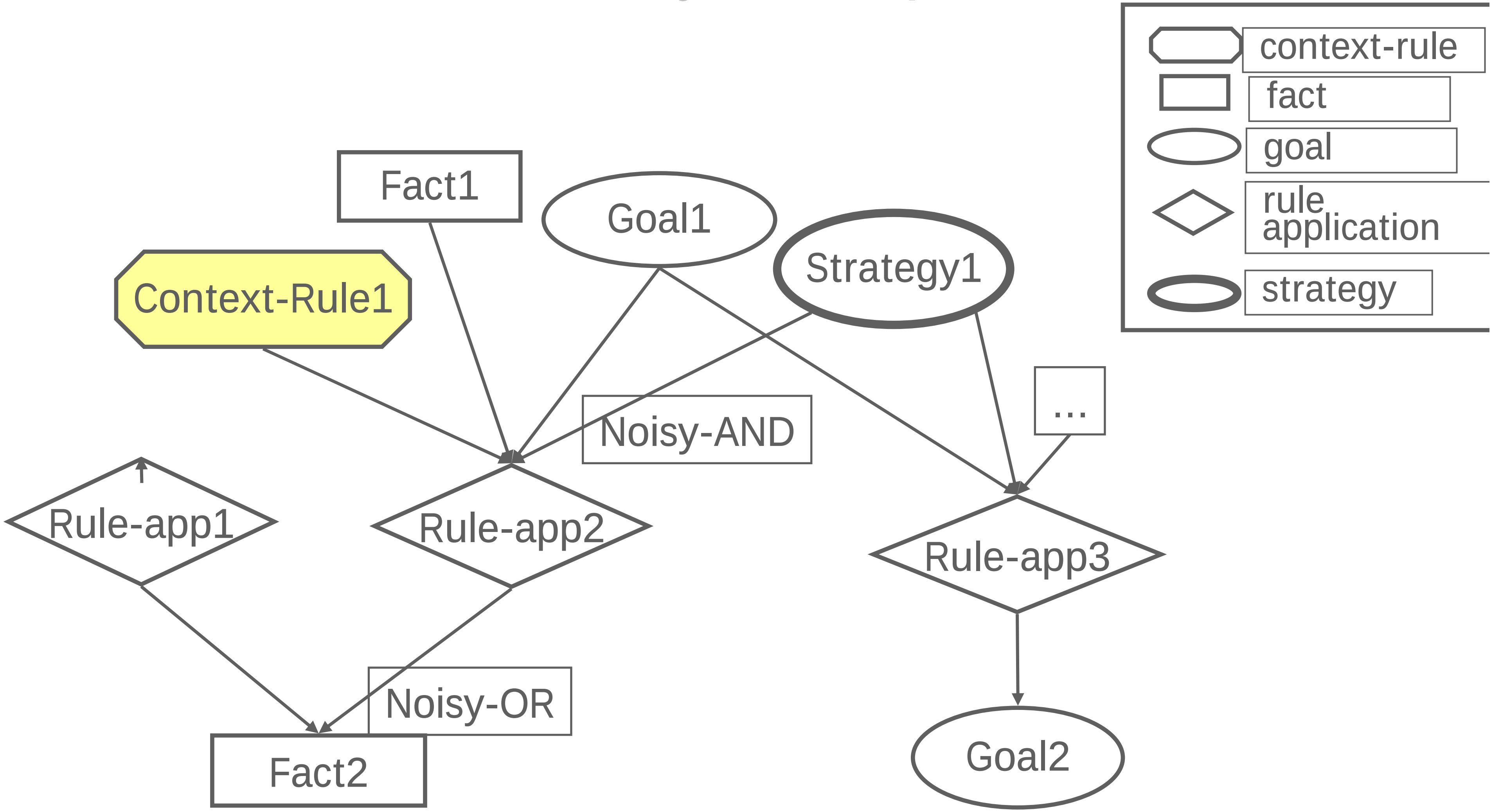
- **Rule nodes**
  - ▷P(R = T): probability that the student knows the rule
- **Context rule nodes**
  - ▷P(CR = T): probability that the student can use the rule in the corresponding context



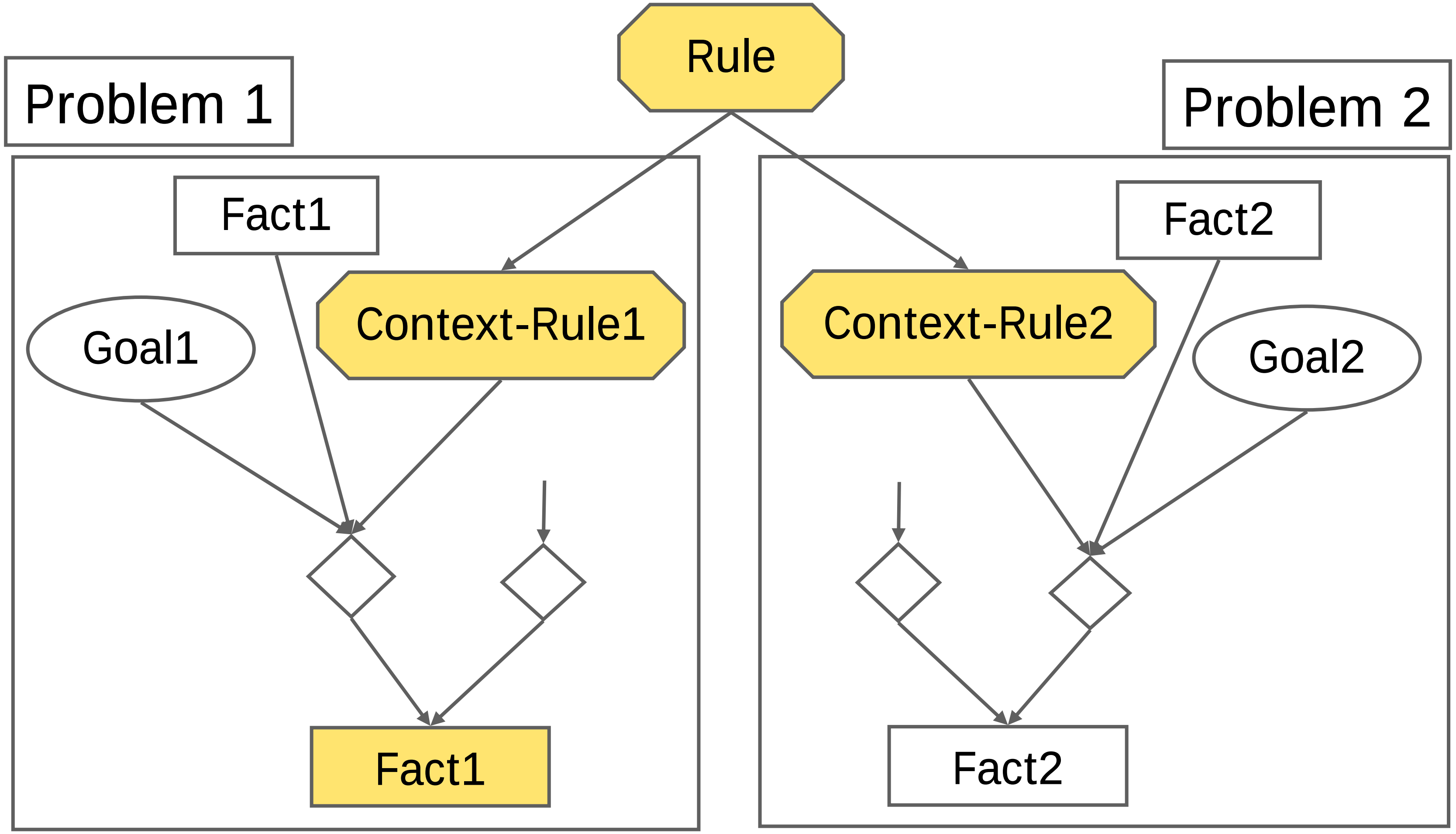
$P(CR_i=T \mid R=T) = 1$

$P(CR_i=T \mid R=F)$ estimates the level of difficulty of context i

Network structure: Dynamic part



Andes1: Propagating evidence



Andes' Model Tracing

- **Assess intention -> useful hints**
- **if error + help request -> best explanation**
- **low mastery of important rules -> mini-lesson**

Andes' Select a Hint Algo

- **Step 1: Goal inference**
  - Start with last observed student entry
  - Find closest unsatisfied goal node
  - At decision points, choose highest probability node (most likely goal)
- **Step 2: Prediction**
  - Start from goal found in step 1
  - Find path from goal to "stuck node" (p<0.8)
  - At decision points, choose lowest probability node (most help needed)

Model Tracing in **PACT/cognitive tutors**

- **EXPERT problem solving space**
  includes rules and mal-rules
- **match student‘s results with nodes in space**
- **hypothesize next possible step**
- **diagnoses errors**

Problems with Model Tracing

- **Correct domain rules provide *positive* evidence only**
- **for *negative* evidence buggy rules needed**
- **assessors most accurate if**
  - student kept along a given solution path
  - student has to provide every step

- **Bayes theoretically optimal but some parameter values hard to obtain**

Bayesian Knowledge Tracing, **PACT/cognitive tutors**

$$p(K_n) = p(K_{n-1} \mid \text{evidence}) + (1 - p(K_{n-1} \mid \text{evidence})) * p(T)$$

p(Ko) initial learning (a priori) probabiity of rule

p(T) transition prob following an opportunity of applic.

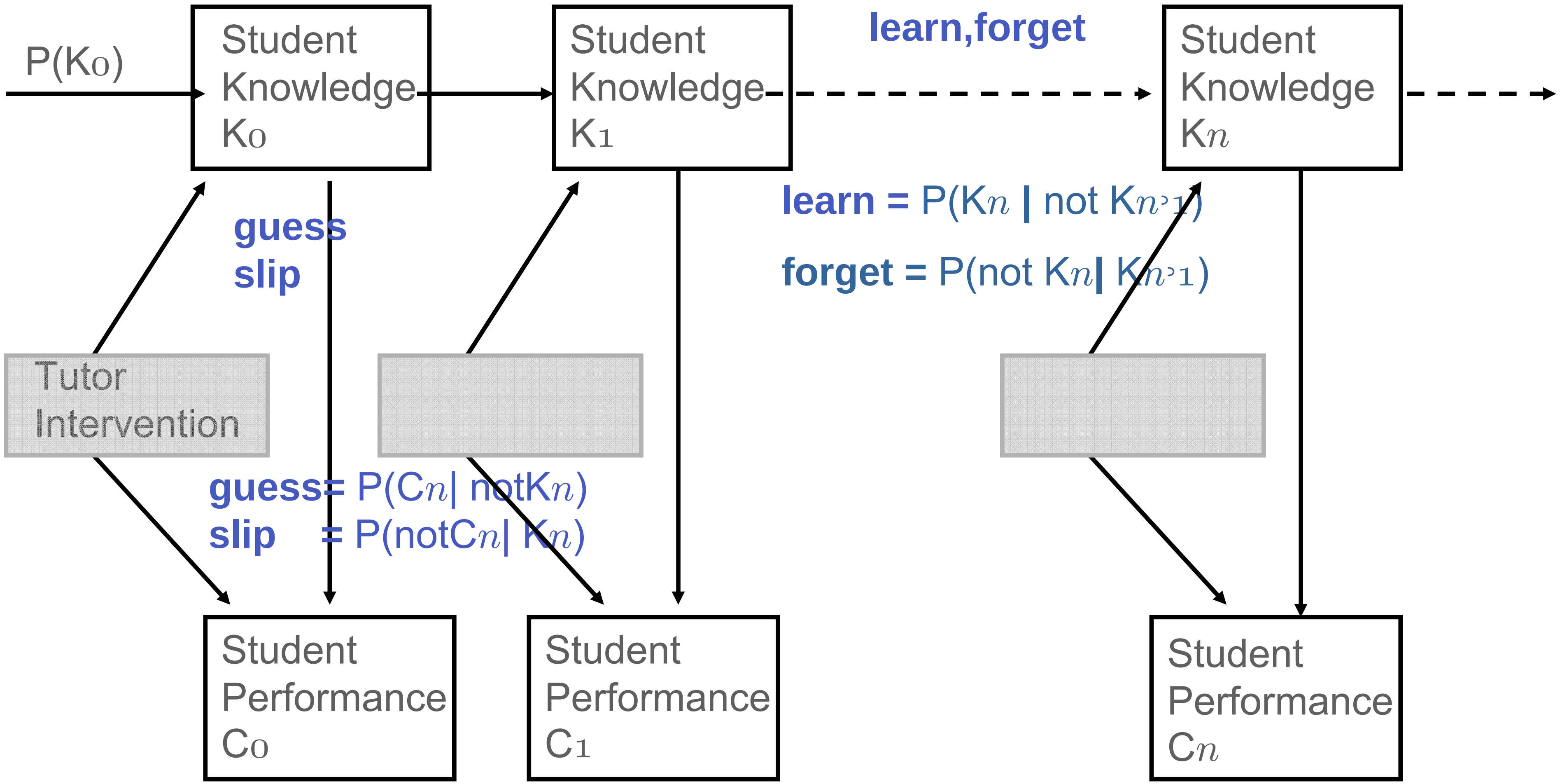
$$p(\text{correct}) = p(K) * (1-p(S)) + (1 - p(L)) * p(G)$$

p(G) guess

p(S) slip

Dynamic Decision Networks

Dynamic Bayesian Networks (BNT package)



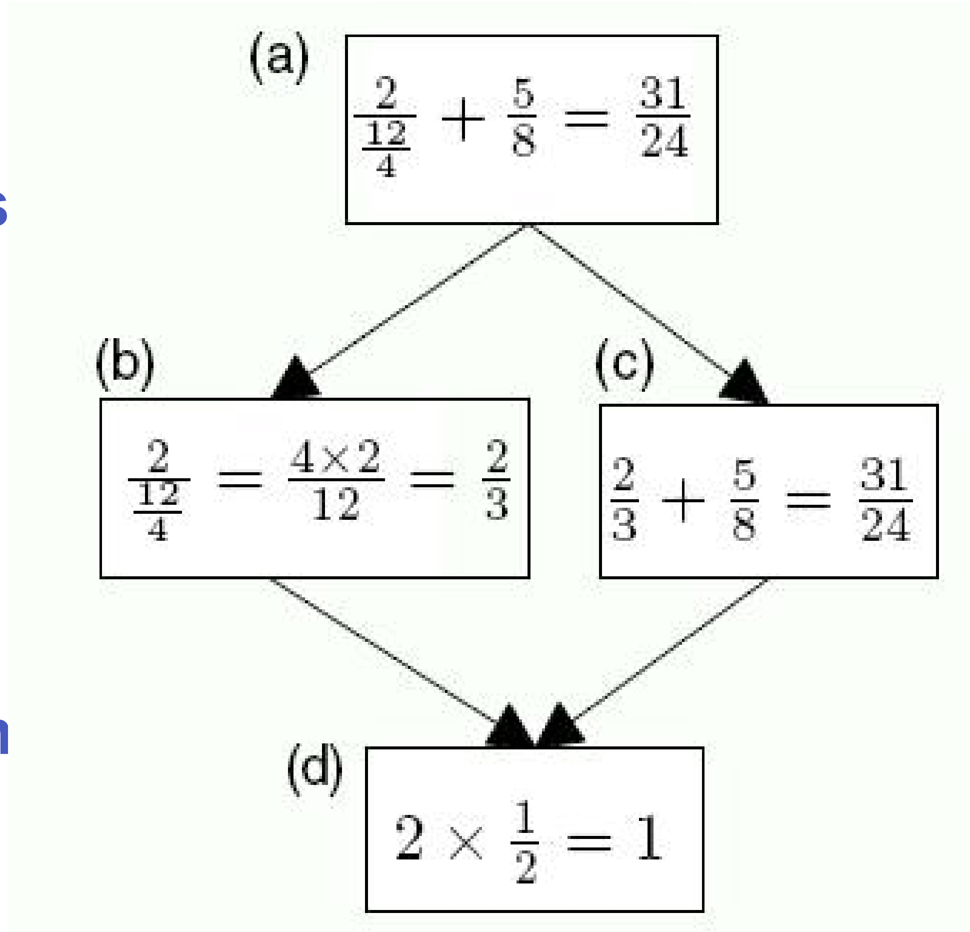
Data-driven approaches: knowledge space theory

Knowledge space theory [Falmagne1999], IRT theory, POKS Desmarais

- **Originally for adaptive testing** psychometric methods
- **Build structure from data only (observable knowledge items.** #subjects with a > c**), no modeling ~**
- **Item to item node structure**
- **Compute individual knowledge space**
- **Compute potential next KS for adaptive test**

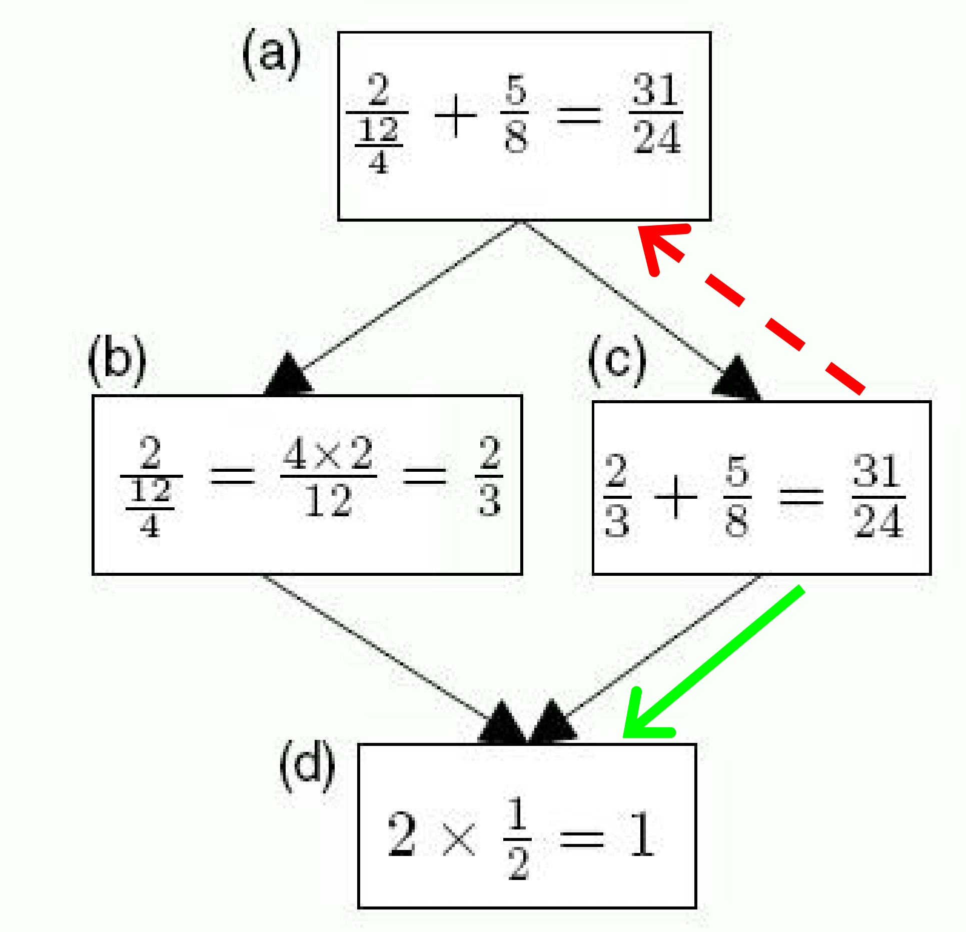
Item to item node structure

- DAG of exercises (items)
- Edges point to prerequisites
- Induce *surmise* relation
- *a* > *c*: mastery of *c* will precede mastery of *a*. if student succeeds with *a* then with *c* too
- Likely partial ordering which items learned
- DAG closed under union (and intersection) – reduces # data



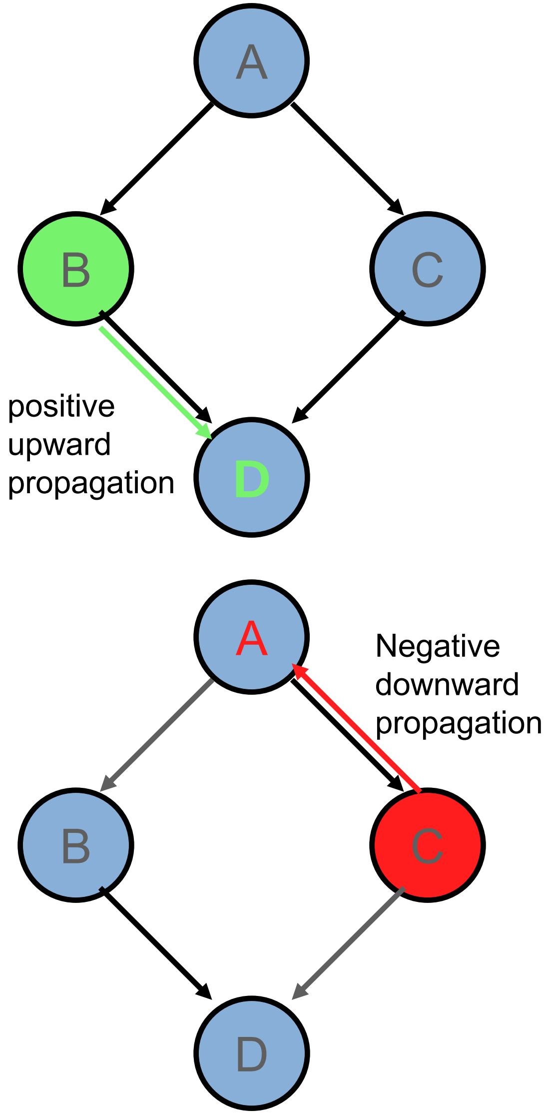
Knowledge Space, POKS structure

- Mastery of a node implies:
  Mastery of prerequisites
  (positive upward propagation)
- Failure at a node implies:
  Items that depend on this node
  are not yet mastered
  (negative downward propagatior



Knowledge Space Propagation

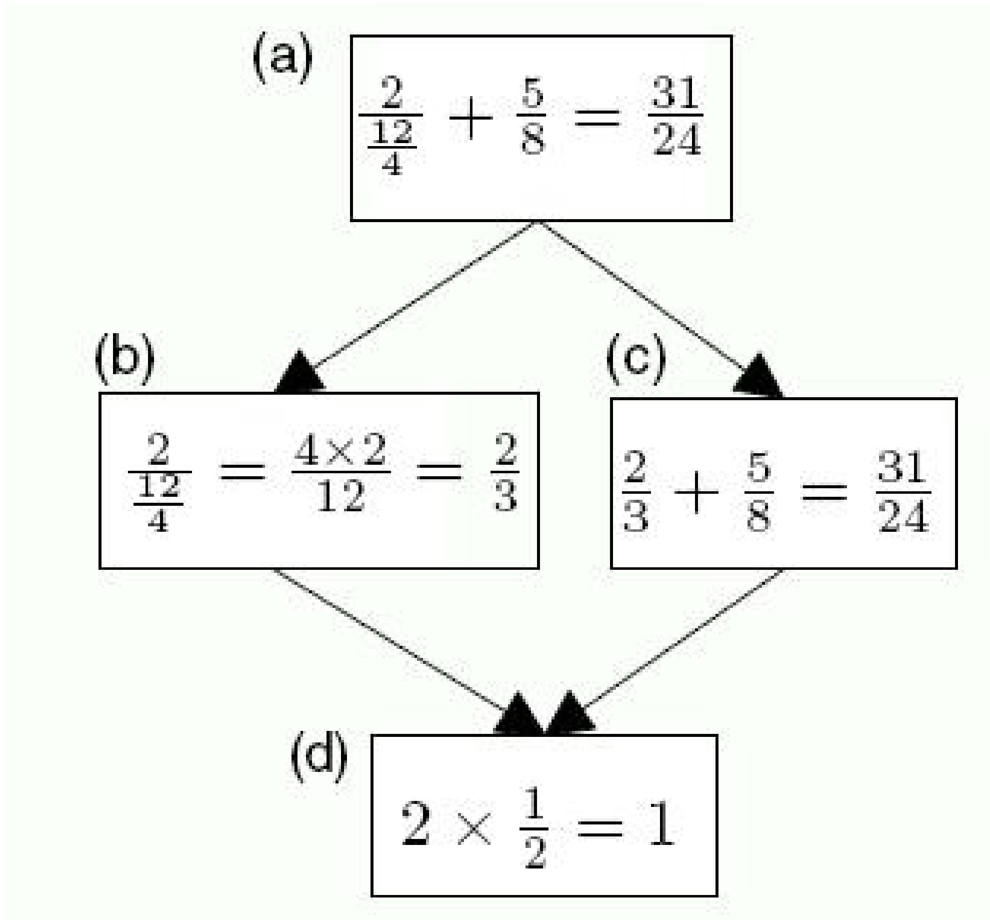
- If student solves B, prerequisite D can be assumed to be mastered
- If student fails at C, depending item A is most likely not yet mastered



Knowledge Spaces

- **Individual's knowledge state = subset of items** (mastered)
- Example knowledge states: **{Ø,{d},{c,d}{b,d},{b,c,d},{a,b,c,d}}**
- **Knowledge state determines which state to move to**

  other knowledge states unlikely
- **Determine mastery of one or few ability dimensions**



Knowledge space inference

- **Learning item-to-item structure:**
  - ▷K2: BN structural learning: greedy search over space of network topologies- given ovserved distribution:
  - ▷POKS: pairwise analysis:

    $P(B|A) \geq p_{\mathcal{C}}$(mastered)

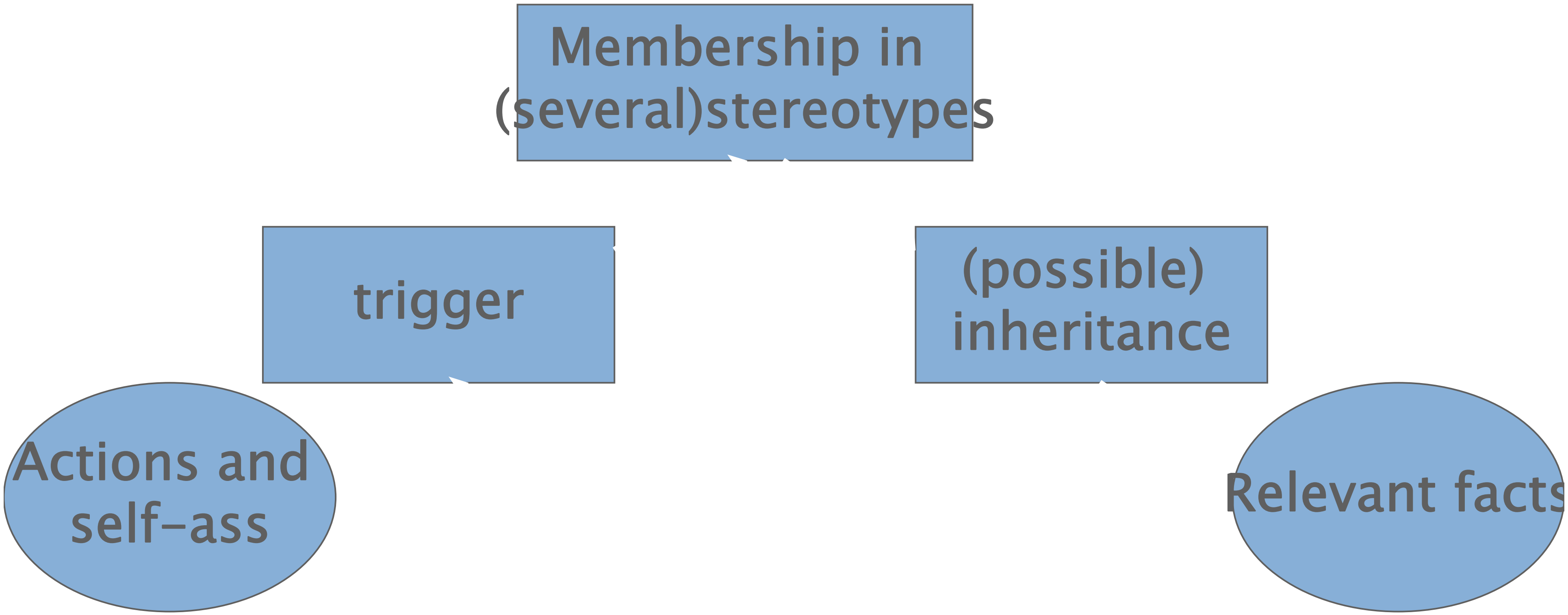
    $P(\tilde{A}|\tilde{B}) \geq p_{\mathcal{C}}$(not mastered)

    $P(B|A) \neq P(B)$

- **With network compute probability of success for items from partial evidence with Bayesian posterior** P(H | E1,E2…)
- **Compute from probability distribution: most likely next knowledge state**
  - ▷ interest is in predictive power

Stereotypes

- Stereotype: body + set of triggers
- may be arranged in hierarchies (inheritance)



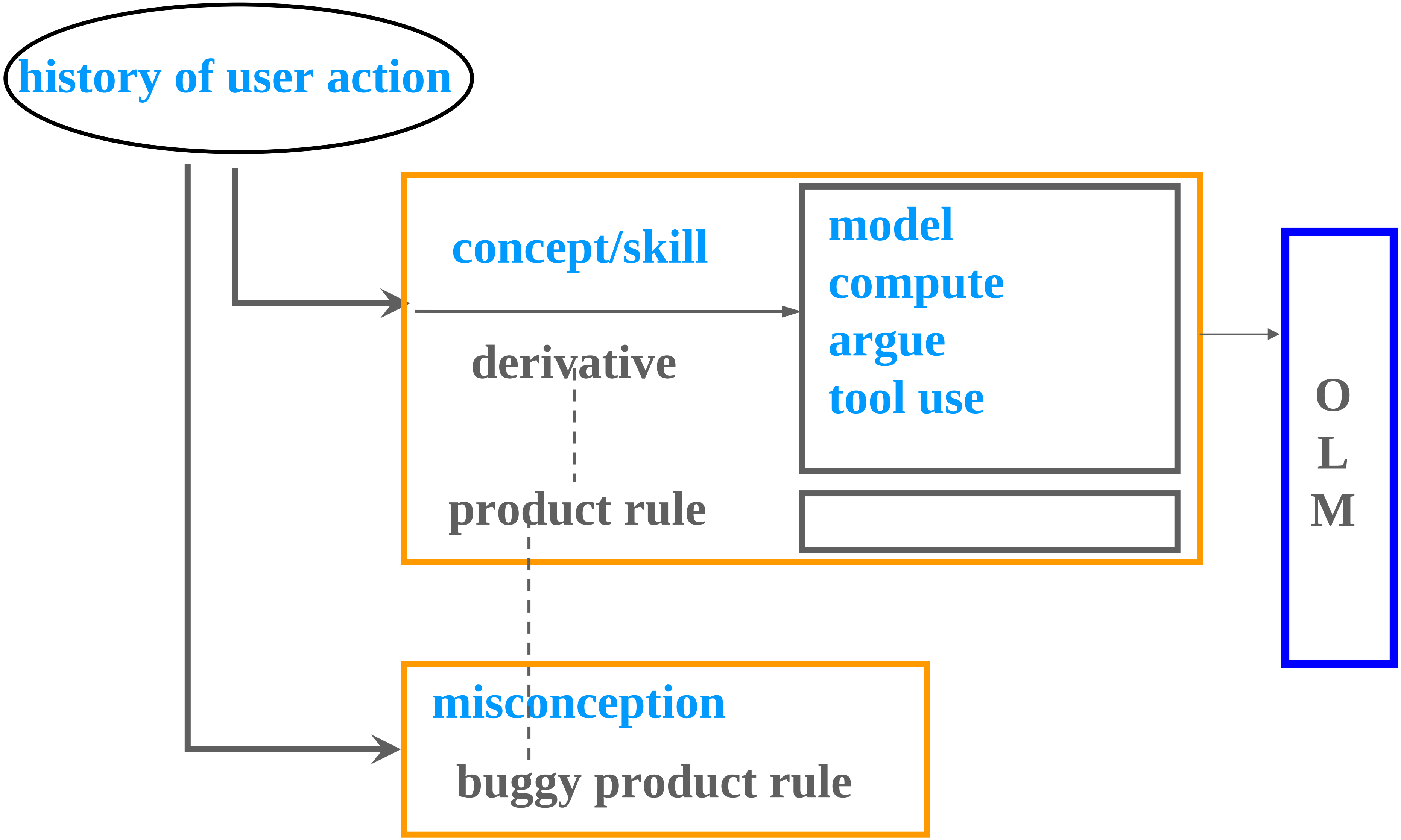
Stereotypes, rules

- **Trigger rules** (Cond Trigger)
  - ▷self-assessed ‚expert‘ THEN stereo= Expert
  - ▷at least 20 topics K=excellent THEN stereo=Expert
- **Rules for evaluation and presentation** (Cond Scoring)
  - ▷Beginner AND diff(T)=simple THEN ev(T)=10 points
  - ▷knows(T) THEN substract 20 for appropriate(T)

ActiveMath SLM – (Simple Learner Model)

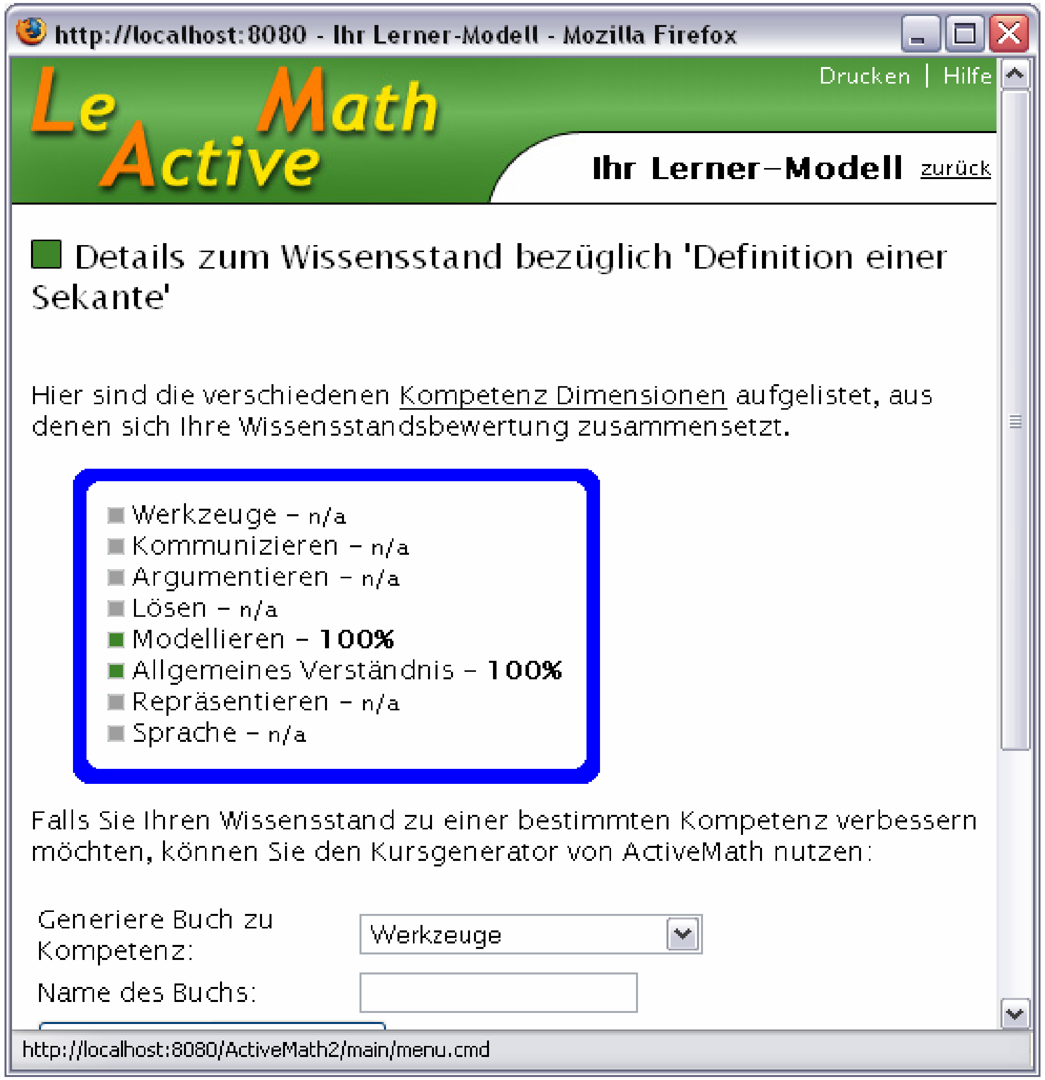
- **Goals**
  - Grasp learners‘ knowledge
  - Regard hard evidence only
  - Keep it simple
  - Consider learners’ improvements

ActiveMath Learner Model for Competencies



Individual Competencies in ActiveMath

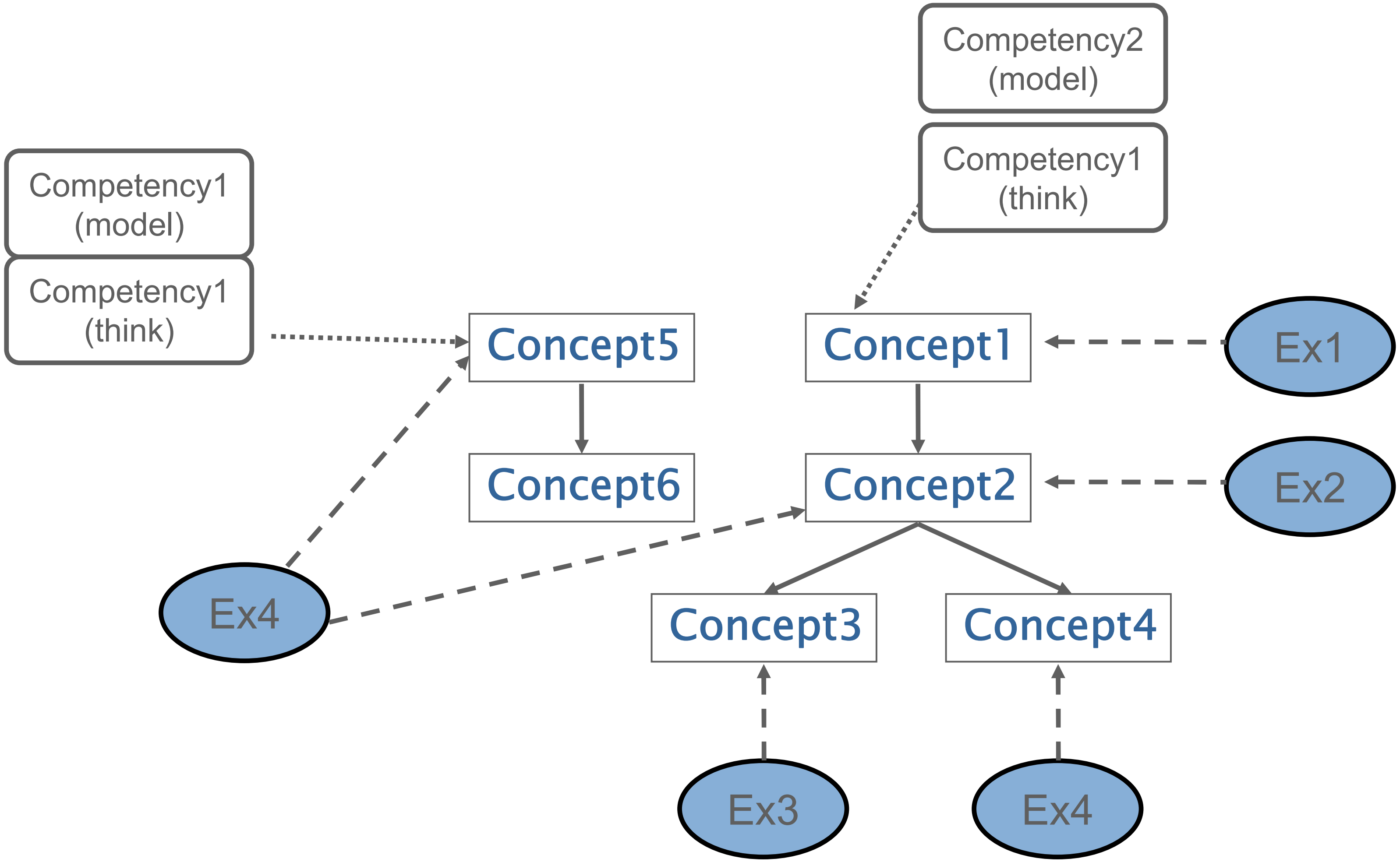
- **Multiple dimensions for Learner Model**
  - Competencies (think, model…)
  - Concepts (sum-rule, power-rule…)
- **Exercises**
  - For specific concepts
  - Steps need competencies



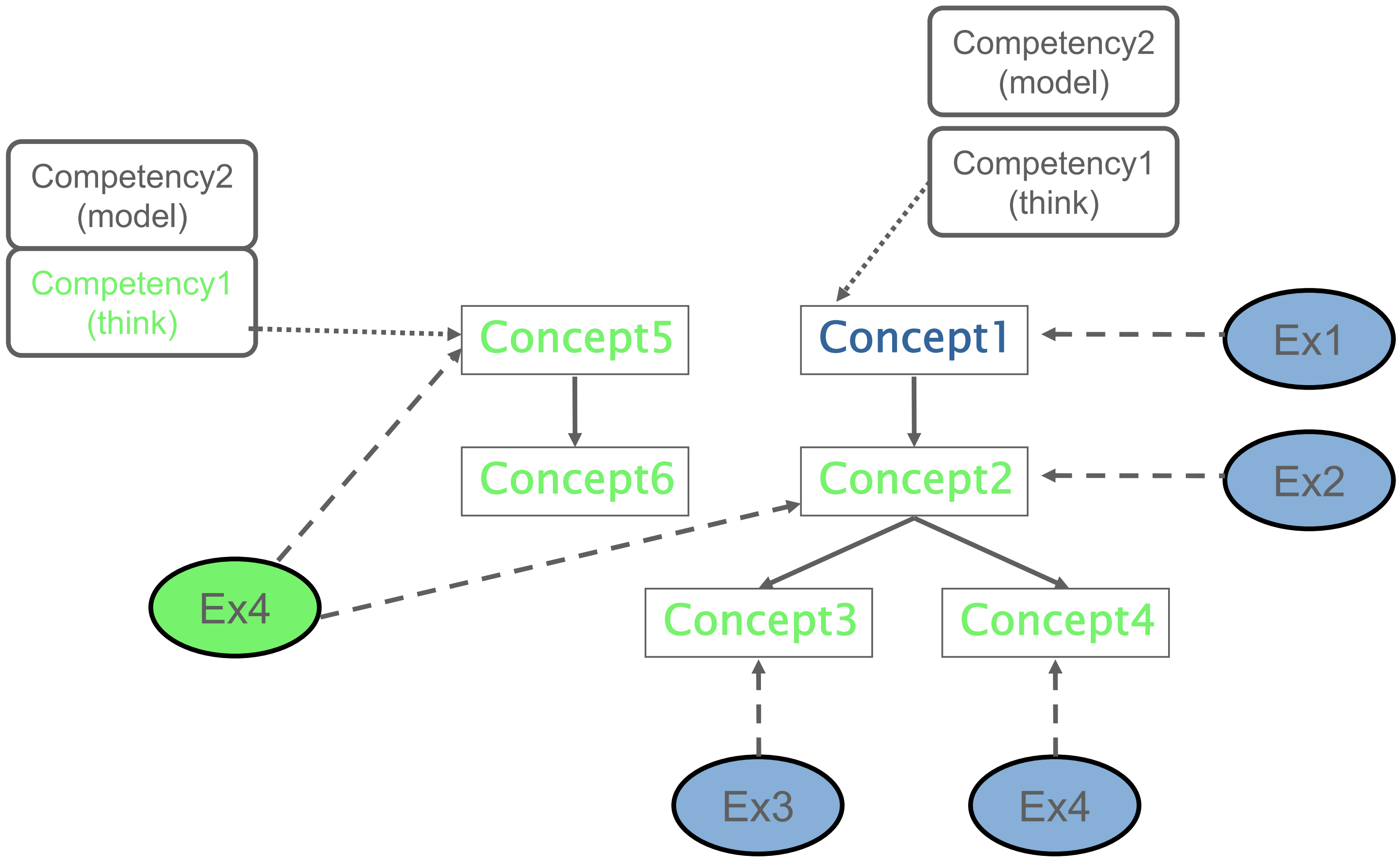
ActiveMath SLM – (Simple Learner Model) cont.

- **Assumptions**
  - Concepts are related, often hierarchical dependencies
    - E.g. percentage calculation depends on fractions
      $\Rightarrow$ to master a concept the prerequisites have to be mastered
    - Direct vs. indirect evidence vs. self-assessment (decreasing trust)
      SLM tries to assess using strongest evidence:
      $\Rightarrow$ 1st direct, 2nd indirect evidence, 3rd self-assessment
  - Exercises are hard evidence
    - Modulo guessing (not considered in SLM)
  - Simplicity: Average over exercise results
  - Learners' improvements
    - Consider only last n exercises (currently n = 4)

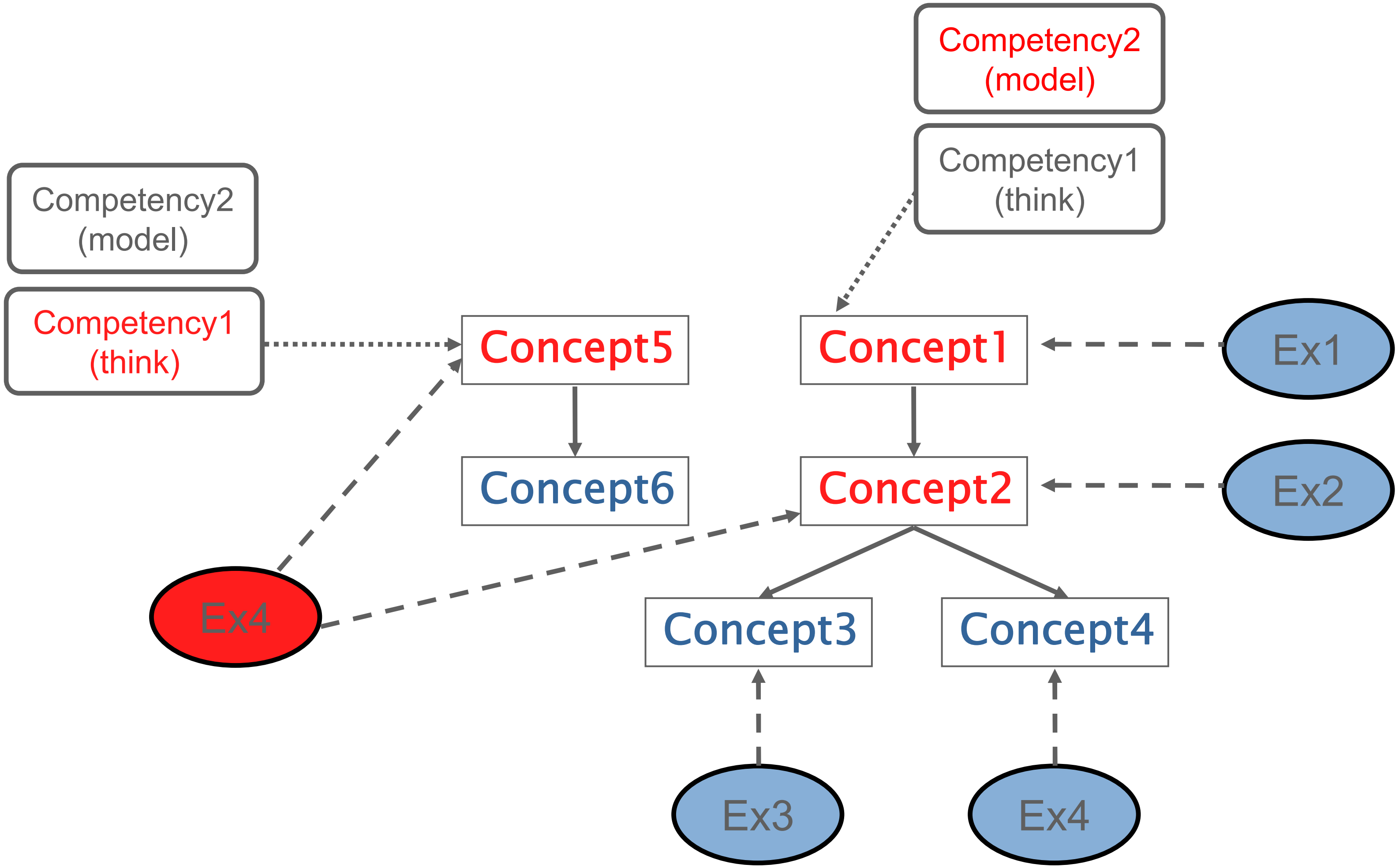
Concept dependencies



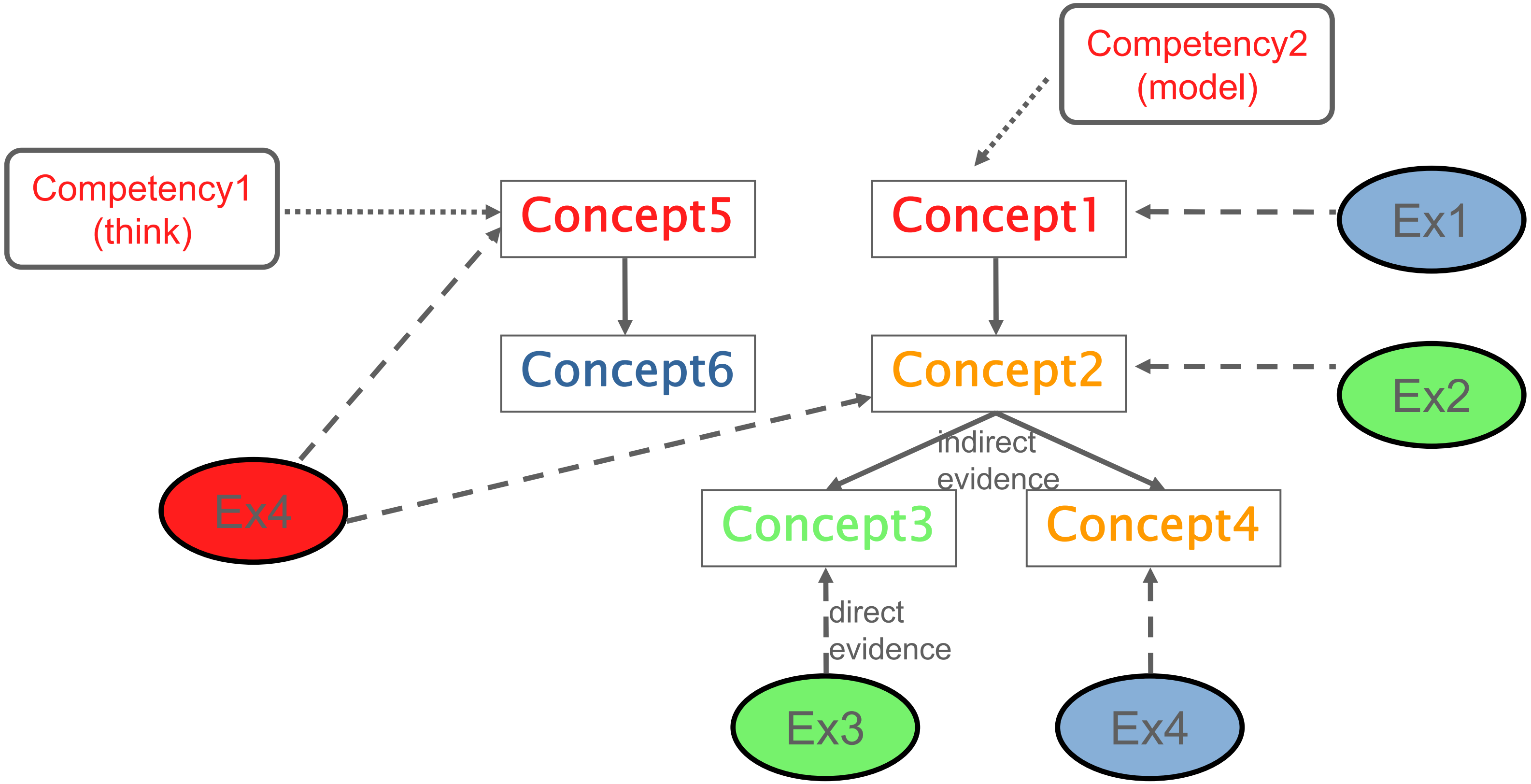
Concept-competency updating (positive)



Concept-competency updating (negative)



Concept-competency updating (combined)



- Direct evidence overwrites indirect evidence

Creating BN: observable – hidden variables

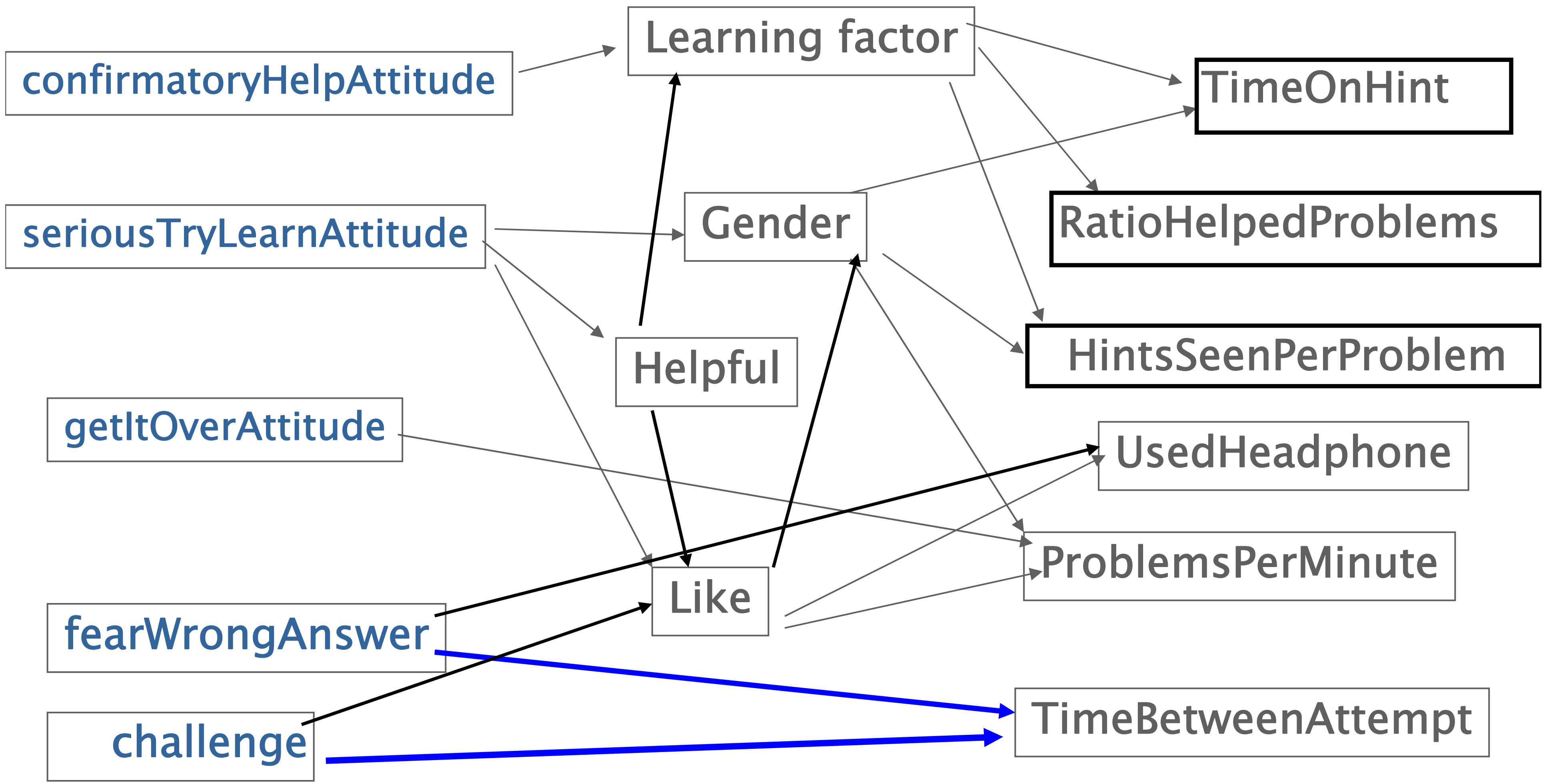
- **log data**
  - av.Time, av.mistakes,#problems
  - Help request, av.hints, time-on-help
  - Help-timing (before,after,correct)
- **Pretest-posttest result**

- **questionnaire data**
  - Gender
  - Motivational
  - Affective
  - Past experience
  - Perception of system
  - Attitude towards help
  - Learning attitude
  - Like challenge

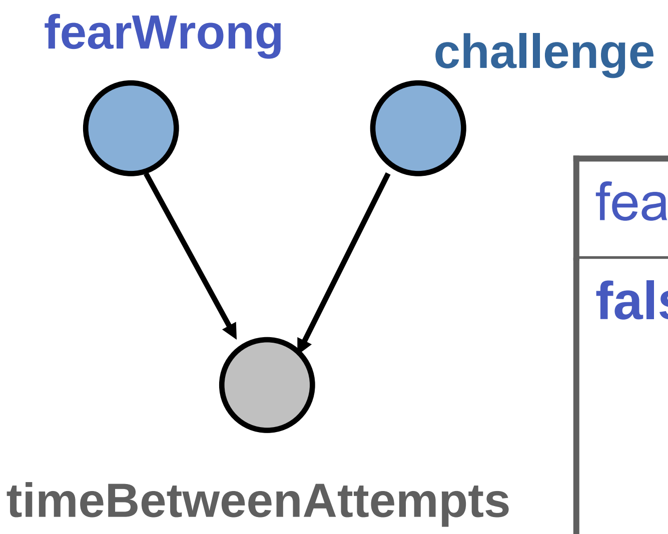
Methodology

- From bi-variate Pearson‘s correlations
- Eliminate links between observables
- Give single direction to links
- For links between non-observables create unidirectional links
- Eliminate circles
- (grouping of factors to reduce dimensions)
- Build Bayes net:
- discretize variables
- simplify (chi-square, if child removed and accuracy not affected)
- CPT for those vars

Correlations: observable – non-observables



Computing CPTs: maximum likelihood method



fearWrong	challenge	timeB	**cases**	**P**
false	**false**	**Low**	**43**	**0.64**
		High	**24**	**0.36**
	true	**Low**	**35**	**0.42**
		High	**48**	**0.58**
true	**false**	**Lpw**	**8**	**0.50**
		High	**8**	**0.50**
	true	**Low**	**7**	**0.32**
		High	**15**	**0.68**

Situational Variables

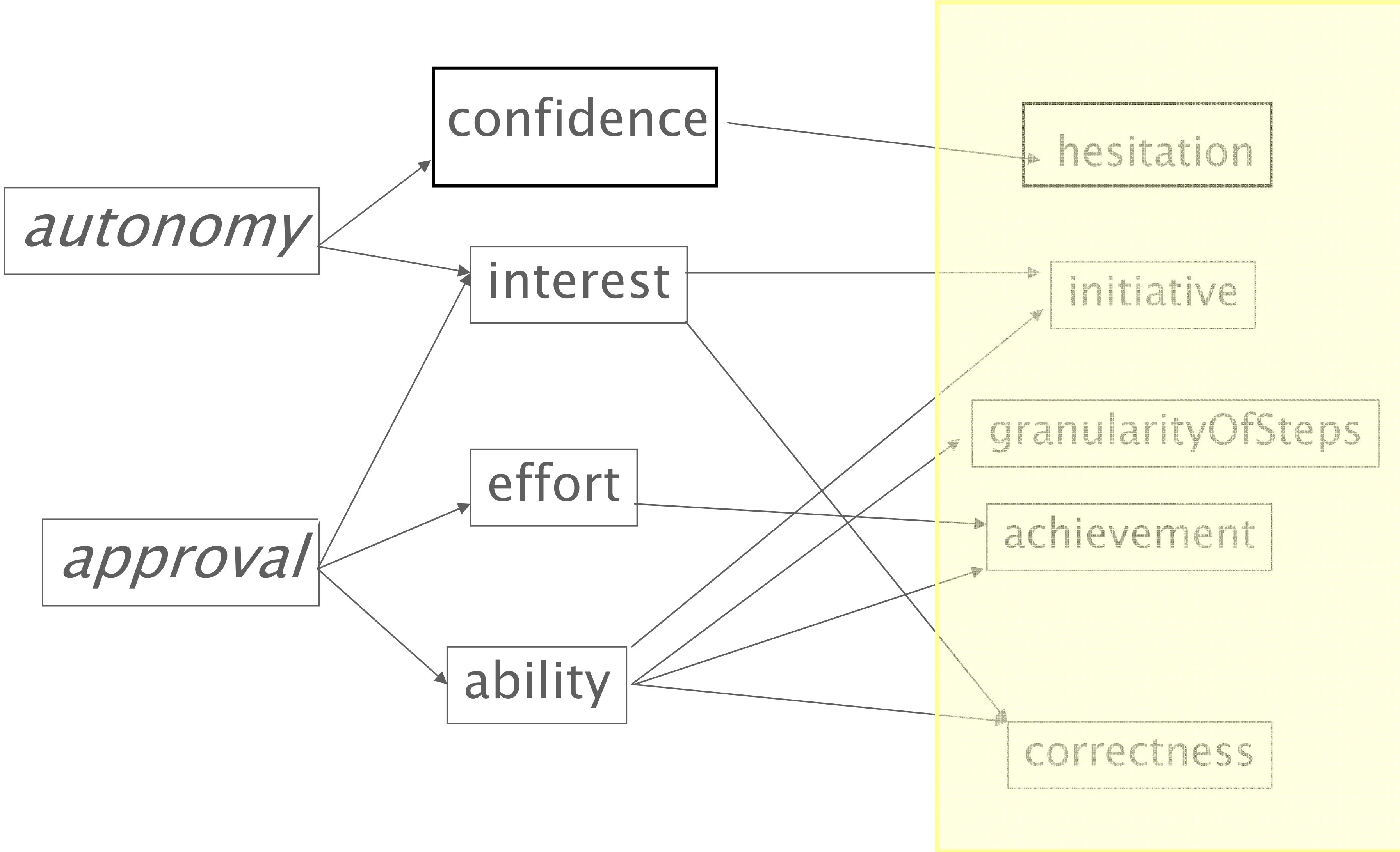
- **Situational variables based on information from LeAM**
  - student interest; student confidence; student effort; student aptitude; level of student hesitation; level of student achievement
- **Numerical Autonomy and Approval values based on the values of the situational variables (SM-ler):**
  - level of recommended guidance to be given to the student (Autonomy)
  - level of explicit approval recommended for a student (Approval)

Implications of the Data Analysis

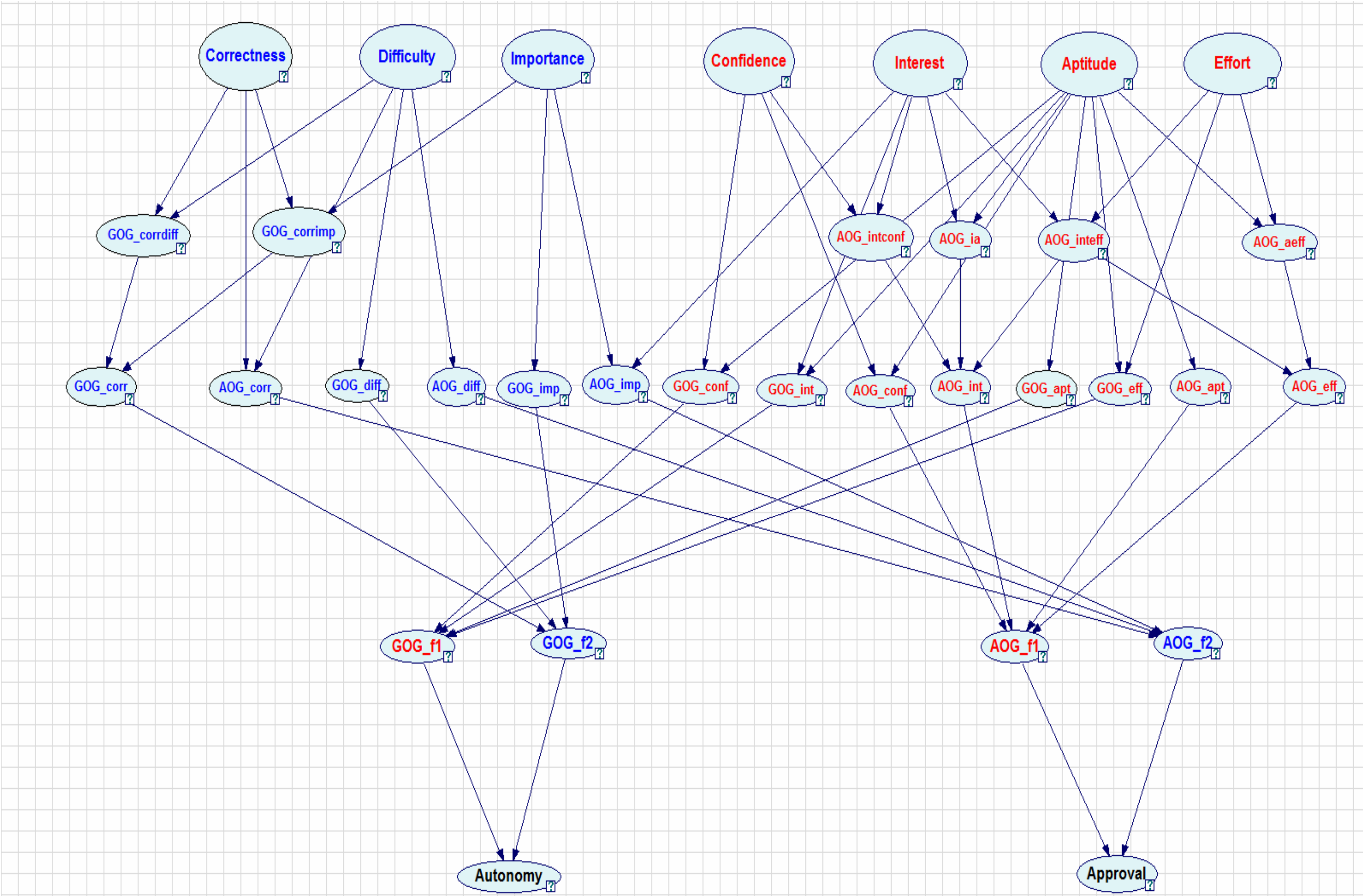
- **Reduced number of factors - the final set being:**
  - Difficulty of material
  - Importance of material
  - Student Confidence
  - Student Interest
  - Student Aptitude
  - Correctness of student's answer
  - Student Effort

- **Data-driven design and implementation of the Bayesian Network (SM-ler) for LeAM**

Bayes Net for Motivational Variables, simpl.



SM: the Bayesian Model



Situational Model: the BN explained

- 7 top nodes represent the set of situational factors modelled
- 2 lowest level nodes represent Autonomy and Approval
- GOG and AOG = Guidance oriented goals & Approval oriented goals resp.
- In some cases it is straight forward to establish what the goals are. In some cases it is not so easy - for example, highly confident students are perceived by our tutors as over confident, but such perception is only the result of other contextual information – in this case student aptitude.
- We rely on analysis of annotated verbal protocols
- Annotations: GOG and AOG

Social voting techniques: I-Help

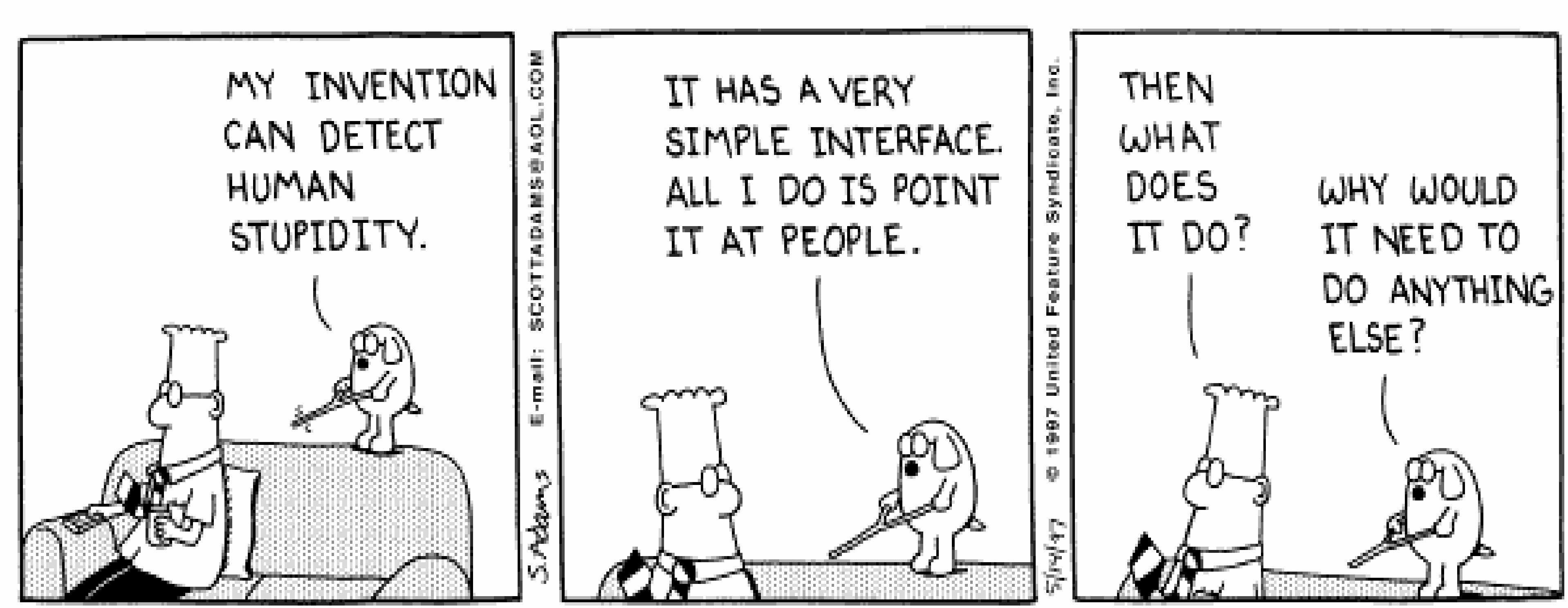
- knowledge
- interests
- eagerness
- helpfulness
- cognitive style
- preferences in helper
- preferred people
- banned people
- banned topics
- help-load

Info from:

user
peers
I-Help private
I-Help Public

For finding and ranking candidate helpers
for protecting users

Putting the model to use



Test of User Models

- **Never fully reliable...**
- **Empirical test with users**
  - protocols..teacher assessors
- **introspectable and modifiable user models**
  - student/teacher revises
  - dialogue about student's beliefs
- **test with artificial users**

Other probabilistic approaches : Belief descriptors