Overview

- **WOW!** The task of SM
- **WHY?** The purpose of SM
- **WHAT?** The content SM
- **HOW?** The method of SM
Programming a God

SM is an attempt to estimate learning from a narrow keyhole of student's actions.

Programming human learning

Learning is the most complex cognitive process

Programming “Hello World”

Brain is the most mysterious universe

The most challenging task in CS?
Why model students?
Does one size fit all?
Student is a user

- $1+1=2$
  - Any user performing an information task is unique
  - Any user performing a difficult information task can benefit from help
  - Learning is a difficult information task

- All of us are being modelled
  - Amazon recommends us new products
  - Facebook filters out irrelevant posts
  - Google adjusts results of searches we perform
..but not just any user

• Usually is less independent, less experienced and needs more support

• Susceptible to system’s errors, can hardly recover from them on his/her own

• Margin of error in SM is rather narrow
  ◦ 7/10 search results are relevant => good personalized search
  ◦ 30% of learning tasks are irrelevant => an awful adaptive educational system
Adaptive systems

Classic loop user modeling - adaptation in adaptive systems
Intelligent vs. Adaptive

1. Intelligent but not adaptive (no user model!)
2. Adaptive but not really intelligent
3. Intelligent and adaptive
What is Being Modeled?

- Student knowledge of the subject
- Student interests
- Student goals
- Student background
- Student individual traits
- Student emotions
- Student meta-cognitive state
- Student context
How to Model User Knowledge

• Scalar model
  ◦ The user knowledge level is modeled as one value
    ◦ Example: MetaDoc, CAT

• Structural model
  ◦ What kind of knowledge?
    ◦ Declarative, procedural, episodic
  ◦ How it relates to expert knowledge?
    ◦ Overlay model -> Bug mode -> Genetic model
Overlay Model of Knowledge

• Domain model
  ◦ The whole body of domain knowledge is decomposed into set of smaller knowledge components
  ◦ A knowledge component
    ◦ Concept, topic, fact (declarative knowledge)
    ◦ Skill, rule, procedure (procedural knowledge)
    ◦ Constraint, bug, misconception (erroneous knowledge)

• User knowledge model
  ◦ Overlay of the Domain model
  ◦ Student knowledge is measured independently for each knowledge component
Vector vs. Network Domain Models

- Vector - no relationships
- Precedence (prerequisite) relationship
- is-a, part-of, analogy
  - Wescourt et al, 1977
- Genetic relationships
  - Goldstein, 1979
Vector model
Network model
Student Knowledge Estimation Approaches

• Ad Hoc (e.g., 1-100)
• Heuristic and rule-based (qualitative, e.g., weak-medium-strong)
• Simple statistical (Bush, Atkinson)
• Probabilistic and Bayesian (BN, D-S…)
• Fuzzy logic
• Neural networks
• Combined approaches and layered models
Simple overlay model
Simple overlay model
Weighted overlay model
Weighted overlay model
Sources of Input for SM

• Adaptive educational systems use exercises, questions, and other evaluation activities to model student knowledge.

• If a page is read, an example is browsed, or an exercise is solved, knowledge of (all) involved concept(s) increases.

• If an exercise is not solved, the system needs to allocate “blame” for involved concept(s).
Sources of Input for SM

- Explicit learning actions with associated learning objects

- Knowledge propagation
  - Some models rely on assumptions that knowledge components are independent
  - Others introduce prerequisite-outcome relationships (arguably, most popular)
  - Prerequisite knowledge is low => outcome knowledge is low
  - Outcome knowledge is high => prerequisite knowledge is high
Other sources of input

- Open SM
  - Ask the user

- Stereotypes-based SM
  - Seed individual SM with preset stereotypical models

- Time spent
  - Time spent reading a page – indicates that the page has been read?
  - Time spent solving an exercise – indicates that the exercise has been challenging?

- Scrolling and mouse moving behavior
  - Was the student actually engaged?

- Eye tracking
  - Visual search based learning activity
  - Detection of targeted activity in complex interfaces
  - Recognition of alarming behavior
  - Prediction of important emotions

Help with the cold-start problem (New user – empty model)
Bug models

- Each concept/skill has a set of associated bugs/misconceptions and sub-optimal skills
- There are help/hint/remediation messages for bugs
Dissecting AEH

- Components

Student model

Domain model

Content

Overlay

Indexing

goals, previous knowledge, preferences, style, context

Knowledge Elements (KE):
- topics, concepts, rules

Learning objects, Web pages, exercises, examples, etc.
Indexing
(from Content to knowledge)

- **Cardinality** (single/multiple concept indexing)
- **Expressive power** (info about the link)
Indexing
(from Content to knowledge)

• Concept based hyperspace

• Learning object indexing

• Fragment indexing
Keyword-based indexing

• Based on keywords extracted from web pages visited, bookmarked, saved or explicitly provided by the user

• Bag-of-words
  ◦ Simply a set of most popular words, can be used in different kinds of systems
  ◦ Each keyword may be also associated with a numerical weight representing its importance in the profile

• Profile vectors
  ◦ An overlay of a keyword vector used in document modeling in a specific system
  ◦ 0-1 vector
  ◦ Weighted vector

• Benefits
  ◦ Simplicity, automation

• Shortcomings:
  ◦ Words may have multiple meanings. Same idea can be expressed by different words. Because of this polysemy and synonymy, the keywords in the user profile are ambiguous, making the profile inaccurate
0-1 Keyword profiles

- Rows represent document terms
- Columns represent users

User 1 learnt a document about variables
User 2 learnt a document about loops

<table>
<thead>
<tr>
<th></th>
<th>User 1</th>
<th>User 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>loops</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>arrays</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
## Weighed Keyword Profiles

<table>
<thead>
<tr>
<th>User 1</th>
<th>User 2</th>
<th>User m</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_{11} )</td>
<td>( w_{12} )</td>
<td>( w_{13} )</td>
</tr>
<tr>
<td>( w_{21} )</td>
<td>( w_{22} )</td>
<td>( w_{23} )</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( w_{m1} )</td>
<td>( w_{m2} )</td>
<td>( w_{m3} )</td>
</tr>
</tbody>
</table>
Advanced keyword profiles

• Using lexical ontologies (WordNet) to deal with synonymy, polysemy
• Using controlled vocabularies to identify important keywords
• Using Probabilistic Topic Modeling (PSLA and LDA) to extract more representative topics

• Yet, the applicability is limited due to low accuracy and interpretability of such profiles
Simple goal model

- Learning goal as a set of topics
More complicated models

- Sequence, stack, tree
Short-term vs Long-term SM

My current cell phone company charges me $14.95 per month for service and $.13 per minute. PPS Cellular Phone Company has offered me $15.00 worth of free calls a month if I switch, but the charge is $.39 per minute.

1. How many minutes of calls can I get from PPS Cellular Phone Company for $50? What is the cost from my current company for that number of minutes?

2. How many minutes of calls can I get from my current company for fifty dollars? What is the cost from PPS Cellular Phone Company for that number of minutes?

3. What is the cost from both companies for sixty minutes of calls?

\[ \frac{14.95}{0.13} - 15 = 115.1923 \]

\[ \frac{29.95}{0.26} \]

\[ 115.1923 = t \]
Modelling Meta-cognitive State

- Modeling a learner’s understanding of their own knowledge and skills and helping them to guide their own learning

- Types of Metacognitive Skills, Knowledge, Processes:
  - Knowledge:
    - What is my current knowledge state?
    - Can I solve this exercise?
    - Where do I have a gap?
  - Regulation Skills:
    - Planning
    - Monitoring
    - Evaluating
  - Strategies:
    - Help seeking
    - Self-explanation
    - Problem solving
Help-seeking Tutor
Given: In the Hourglass shown, Segment AL intersects Segment PI at Point R.

1. If the measure of Angle ARP = 41.1 degrees, find the measure of Angle IRL.

\[
\begin{align*}
\text{m} \angle \text{ARP} &= 41.1 \\
\text{m} \angle \text{IRL} &= 128.9 \\
\end{align*}
\]

Reason: Given

2. If the measure of Angle IRL = 43 degrees, find the measure of Angle ARP.
Model of help-seeking behavior
Modelling Context
Context dimensions

- Device
  - Screen
  - Computational power
  - Connectivity

- Location and movement
  - Coordinates vs. place

- Time
  - events

- Ambient light and noise

- Weather

- New wearable devices have new sensors
  - Health, mood

Sensors:
- Accelerometer
- Gyroscope
- Magnetometer
- GPS
- Proximity sensor
- Ambient Light sensor
- Temperature sensor
- Humidity sensor
- Barometer
- Gesture sensor
- Compass
+ WWW
Example - QuizGuide

Quiz Guide

**Question 5**

```c
main ()
{
    int i = 10;
    int S = 0;
    do {
        i--;
        S += i;
    } while (3 < i);
}
```

What is the final value of `S`?

`s = _______`

**Submit**
Architecture

CUMULATE

knowledge levels

usage data

student model updates

inference

student's answers

QuizGuide

loops (while)
Quiz1
Quiz2

increment decrement
compound assignments
Quiz1

logical expressions
Quiz1

loops (do while)
Quiz1
Quiz2

conditionals (if else)

QuizPACK

Question 5

```c
main()
{
    int i = 10;
    int S = 0;
    do {
        i--;
        S += i;
    } while (3 < i);
}
```

What is the final value of S

S = ___________
Topic-based domain model
Topic

weight-1

weight-2

weight-n

Content modelling

Adaptation
Modelling formula

\[
K_i = \frac{\sum_{j=1}^{N_i} w_{ij} \frac{\sum_{k=1}^{M_j} x_{jk}}{z_{jk}}}{\sum_{j=1}^{N_i} w_{ij}}, \quad \text{where}
\]

- $K_i$ is the current level of knowledge for the $i$-th topic,
- $N_i$ is the number of quizzes participating in the $i$-th topic,
- $M_j$ is the number of questions in the $j$-th quiz,
- $w_{ij}$ is the weight (contribution) of the $j$-th quiz in the $i$-th topic,
- $x_{jk}$ is the number of correct attempts for the $k$-th question of the $j$-th quiz,
- $z_{jk}$ is the total number of attempts for the $k$-th question of the $j$-th quiz.
## Knowledge gain

<table>
<thead>
<tr>
<th>Non-adaptive semesters</th>
<th>Adaptive semesters</th>
<th>Test statistics</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge gain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M = 5.77 )</td>
<td>( M = 6.55 )</td>
<td>( U = 539 )</td>
<td></td>
</tr>
<tr>
<td>( Mdn = 6 )</td>
<td>( Mdn = 7 )</td>
<td>( p = .019 )</td>
<td></td>
</tr>
<tr>
<td>( SD = 2.06 )</td>
<td>( SD = 2.35 )</td>
<td>( r = .26 )</td>
<td></td>
</tr>
<tr>
<td><strong>Normalized knowledge gain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M = .59 )</td>
<td>( M = .68 )</td>
<td>( U = 572 )</td>
<td></td>
</tr>
<tr>
<td>( Mdn = .6 )</td>
<td>( Mdn = .7 )</td>
<td>( p = .044 )</td>
<td></td>
</tr>
<tr>
<td>( SD = .20 )</td>
<td>( SD = .23 )</td>
<td>( r = .22 )</td>
<td></td>
</tr>
</tbody>
</table>
Within system performance

<table>
<thead>
<tr>
<th>Used mostly non-adaptive</th>
<th>Used mostly adaptive</th>
<th>Test statistics</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M = .36$</td>
<td>$M = .47$</td>
<td>$t(35) = 2.31$</td>
<td></td>
</tr>
<tr>
<td>$Med = .31$</td>
<td>$Med = .51$</td>
<td>$p = .027$</td>
<td></td>
</tr>
<tr>
<td>$SD = .15$</td>
<td>$SD = .13$</td>
<td>$d = .77$</td>
<td></td>
</tr>
</tbody>
</table>
## Semester activity

<table>
<thead>
<tr>
<th></th>
<th>Non-adaptive semesters</th>
<th>Adaptive semesters</th>
<th>Test statistics</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Per semester</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity/ attempts</td>
<td>$M = 132$</td>
<td>$M = 279$</td>
<td>$U = 769$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Md = 88.0$</td>
<td>$Md = 126$</td>
<td>$p = .011$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 119$</td>
<td>$SD = 313$</td>
<td>$r = .26$</td>
<td></td>
</tr>
<tr>
<td>Coverage/quizzes</td>
<td>$M = 13.7$</td>
<td>$M = 24.5$</td>
<td>$U = 544$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Md = 14.0$</td>
<td>$Md = 23.0$</td>
<td>$p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 7.87$</td>
<td>$SD = 12.0$</td>
<td>$r = .42$</td>
<td></td>
</tr>
<tr>
<td>Coverage/ goals</td>
<td>$M = 8.77$</td>
<td>$M = 10.5$</td>
<td>$U = 874$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Md = 9.50$</td>
<td>$Md = 11.0$</td>
<td>$p = .079$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 4.65$</td>
<td>$SD = 2.60$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Sessions

<table>
<thead>
<tr>
<th></th>
<th>Non-adaptive semesters</th>
<th>Adaptive semesters</th>
<th>Test statistics</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Per session</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity/ attempts</td>
<td>$M = 12.8$</td>
<td>$M = 18.7$</td>
<td>$U = 632$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Mdn = 10.7$</td>
<td>$Mdn = 16.3$</td>
<td>$p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 10.9$</td>
<td>$SD = 11.2$</td>
<td>$r = .36$</td>
<td></td>
</tr>
<tr>
<td>Coverage/ quizzes</td>
<td>$M = 1.89$</td>
<td>$M = 3.71$</td>
<td>$U = 253$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Mdn = 1.71$</td>
<td>$Mdn = 2.79$</td>
<td>$p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = .71$</td>
<td>$SD = 2.76$</td>
<td>$r = .65$</td>
<td></td>
</tr>
<tr>
<td>Coverage/ goals</td>
<td>$M = 1.40$</td>
<td>$M = 2.08$</td>
<td>$U = 533$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Mdn = 1.32$</td>
<td>$Mdn = 1.67$</td>
<td>$p &lt; .001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = .40$</td>
<td>$SD = 1.53$</td>
<td>$r = .44$</td>
<td></td>
</tr>
<tr>
<td>Number of sessions</td>
<td>$M = 11.2$</td>
<td>$M = 18.0$</td>
<td>$U = 939$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Mdn = 8.50$</td>
<td>$Mdn = 11.0$</td>
<td>$p = .20$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$SD = 9.21$</td>
<td>$SD = 28.3$</td>
<td>$r = .13$</td>
<td></td>
</tr>
</tbody>
</table>
Learning Curve

![Graph showing the average topic-based error rate vs. learning step. The graph includes a curve labeled "Learning Curve (Topics)" and a scatter plot with markers representing the data points. The x-axis represents the learning step, and the y-axis represents the average topic-based error rate.]
ROC curve

![ROC curve graph](image-url)