Educational Standard Assignment:
Some Findings Working with CAT & SAT
NSDL 2010 Annual Meeting

René Reitsma¹, Anne Diekema²
Byron Marshall¹, Trevor Chart¹

¹Oregon State University
²Utah State University
Educational Standard Assignment: Some Findings Working with CAT & SAT… Overview

• Need for automated educational standard assignment in TeachEngineering.org.

• Part 1: Comparative analysis of standard assignment by CAT and human catalogers (René & Anne).

Automated Standard Alignment in TeachEngineering

- www.teachengineering.org:
  - 578 hands-on science and math K-12 activities.
  - 339 lessons
  - 54 multi-lesson, curricular units
- *Explicit alignments*: by author, supervised by collection catalogers:
  - cover only one state → mean 4.5 stds./document.
- Similar coverage across all states: $917 \times 4.5 \times 50 = 200,000+$ assignments.
  - 200+ per document
  - $917 \times 4.5 \times 10 = 40,000+$ annual updates
BIG!! thank you to CNLP and friends for CAT.

FYI, ‘new’ CAT (August 2010) is really fast 😊 and includes ITEEA* & Common Core Math.
Part 1: Content Assignment Tool (CAT) & Explicit Standard Assignment in TeachEngineering

- 4,165 *explicit* alignments in TE
- 400,000+ (*unsupervised*) CAT assignments (science, math, ITEEA, common core math).
- Q-1: How are CAT assignments different from human (*explicit*) assignments?
- Q-2: Do the differences tell us something about how humans assign these standards in the first place?
- Q-3: Do the differences inform CAT and/or human improvements?
- BTW: What do we really mean when we say that a standard and a curricular item ‘align?’ (Reitsma, Marshall, Zarske (IPM – 2010))
(Inductive) Method & Data

• Approach: build networks of standards; layout the networks, interpret their spatial arrangements:
  – Networks are based on how standards have been assigned to curriculum.
  – Any two jointly assigned standards are considered ‘linked.’

• Compare and contrast the networks for clues.

• Data:
CO 2007 Science Standard Assignments

<table>
<thead>
<tr>
<th></th>
<th>Human Catalogers (CO Curriculum)</th>
<th>CAT (CO curriculum cataloged by humans)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curricular items assigned</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>Assignments</td>
<td>324</td>
<td>139</td>
</tr>
<tr>
<td>Mean number of assignments per curricular item</td>
<td>3.78</td>
<td>1.61</td>
</tr>
<tr>
<td>Standards covered</td>
<td>63</td>
<td>47</td>
</tr>
<tr>
<td>Standard reuse rate</td>
<td>5.14</td>
<td>2.96</td>
</tr>
</tbody>
</table>
CO 2007 Science Standard Assignments... Cont.’d

<table>
<thead>
<tr>
<th>Human cataloger assignments</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25</td>
<td>299</td>
<td>324</td>
</tr>
<tr>
<td>No</td>
<td>114</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>139</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- CAT recall = 25 / 324 = .077*
- CAT precision = 25 / 139 = .18*

*if the humans did it right (?)
‘Curricular units’

– Human network is denser and more clustered.
– Human clusters are curricular units
– Human clusters link through common standards.
– CAT: open structure; less clustering. Has no knowledge of curricular units.
Weighted or unweighted?

- FR diagrams consider the network unweighted; \textit{i.e.}, all links have equal value/weight.
- Two weights:
  - TF/IDF-like: weigh a standard link inversely proportional to the size of its company.
  - ‘Fidelity:’ weigh a link between standards proportional to their mutual fidelity across the collection.
- Compute the KK network layouts
• Resulting KK diagrams showed essentially the same properties as the FR diagrams (hierarchical cluster analysis of two-dimensional positions)
CO Standards: ‘Method’ vs. ‘World’

- **World** standards (W): express facts and principles about the empirical world.
  - *E.g.*, S103EC87: *Light and sound waves have distinct properties: frequency, wavelength and amplitude.*

- **Method** standards (M!): express ways and means of conducting science.
  - *E.g.*, S103ECE9: *A controlled experiment must have comparable results when repeated.*

- Some method standards ‘contaminated’ with world terms and/or examples (M):
  - *E.g.*, S103ECD4: *Technology is needed to explore space (for example: telescopes, spectroscopes, spacecraft, life support systems).*

- Question: How do CAT and human catalogers compare on World vs. Method?
## Standards: ‘Method’ vs. ‘World…’ Cont.’d

<table>
<thead>
<tr>
<th></th>
<th>Standards</th>
<th>M &amp; M! standards</th>
<th>M &amp; M! standards %</th>
<th>M! standards</th>
<th>M! standards %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humans</td>
<td>63</td>
<td>21</td>
<td>33.33</td>
<td>16</td>
<td>25.40</td>
</tr>
<tr>
<td>CAT</td>
<td>47</td>
<td>9</td>
<td>19.15</td>
<td>3</td>
<td>6.38</td>
</tr>
</tbody>
</table>
---

CAT under-assigned *method*.

Humans: *method* standards as curricular hubs

CAT central method hub: S103EC77:

"physical properties of solids, liquids, gases and the plasma state and their changes can be explained using the particulate nature of matter model“
Part 1: TeachEngineering & CAT Conclusions

• Once again, thanks for CAT! TeachEngineering needs it.

• Tools such as CAT can benefit from contextual knowledge; e.g., that certain lessons are part of a larger set of lessons or a curricular unit.

• TeachEngineering curriculum is organized around both world and method standards. Hence, it would be nice if tools such as CAT become better at recognizing method standards.

• Contrast in standard re-use rate sends a signal to human catalogers not to be ‘complacent.’