A Knowledge Metric
with Applications to Learning Assessment

Rafik Braham

Prince Research Group - University of Sousse
ISITCom - Hammam Sousse, Tunisia

KEOD 2010 – Oct. 26, 2010
Motivation

- Two questions in student Learning Assessment:
  1. How do we know they know?
  2. How much do they know?

- a Knowledge Metric (for Assessment)
- a Knowledge Reference (e.g. textbook, course content)

technical rather than epistemological questions.
Outline

- Introduction - Learning
- Knowledge levels
- Knowledge entropy
- Knowledge similarity
- Knowlets in practice
- Case study
- Related work
- Conclusion - Future directions
Learning

Study results: five categories of learning:

- A quantitative **increase in knowledge**: learning is acquiring information or ‘knowing a lot’
- **Memorizing**: learning is storing information that can be reproduced
- **Acquiring facts, skills, and methods** that can be retained and used as needed
- **Making sense or abstracting meaning**: learning involves relating parts of a subject to each other and to the real world
- **Re-interpreting and understanding reality**: learning involves comprehending the world by reinterpreting knowledge.
Learning and Bloom’s taxonomy

- Learning types: adult/child (Piaget)
- Learning domains:
  - \Levels_of_learning_Bloom.mht

Strong relationship between learning and knowledge
Knowledge levels: knowlets

- **Knowledge of Level 1**: this is basic Knowledge. It describes concepts, items or objects, for example animal, tree, person ...

- **Knowledge of Level 2**: Here we have properties and relations defined on concepts. Elements of Knowledge at this level require two K-elements of Level 1. Examples: a parrot is-a bird; Coca-Cola is-a soft-drink; Mozzarella cheese is-made-in Italy, lions are-faster-than humans. It includes simple relations of the type 5=2+3 and 5>4 as well.
Knowledge levels: knowlets (cont’d)

- **Knowledge of Level 3**: this level incorporates three cases:
  - Rules and inferences, for example:
    
    \[
    \text{hasUncle} \leftarrow \text{hasParent} \land \text{hasBrother}
    \]
  - Logical structures of the type IF-THEN
  - Equations

- **Knowledge of Level 4**: this is the highest level. It includes logical structures of the form IF-THEN-ELSE such as those encountered in theorems

- We will call elements (grains or items) of Knowledge “Knowlets”

- Definition not the same as Mons’ (2008) and knowlets are not just the smallest “piece” of Knowledge. They are hierarchical elements of Knowledge
Knowledge entropy (1/2)

- \[ H(K_L) = \alpha_n \log_2 \left[ 1 + \frac{K_{in}^2 - K_{in} \cdot K_S}{K_{in}^2} \right] \quad (1) \]
  or \[ H(K_L) = \alpha_n \log_2 \left[ 2 - \frac{K_{in} \cdot K_S}{K_{in}^2} \right] \]

- \[ K_{in} \cdot K_S = \text{Sim}(K_{in}, K_{in} \cap K_S) \quad (2) \]

- Special case:
  - \[ K_{in} \cdot K_{in} = \text{Sim}(K_{in}, K_{in}) = K_{in}^2 \].
Knowledge entropy (2/2)

- $H(K_L) = \alpha_n \log_2 \left[ 1 + \frac{(K_{in}^2 - K_{in}K_S)}{K_{in}^2} \right]$ \hspace{1cm} (1)

- $K_{in} \subset K_S$: $H(K_L) = 0$

- $K_{in} \not\subset K_{in}$: $H(K_L) = \alpha_n$ (i.e. max value)

- In general:
  - $K_{in} \cap K_S \neq 0$: $0 < H(K_L) < \alpha_n$
  - $\alpha_n = \log_2 2^{n/2} = n/2$ ($\alpha_n = n\alpha_1$ and $\alpha_1 = \log_2 2^{1/2} = \frac{1}{2}$), $n = 1, 2, 3, 4$.  

$\Box$
A simple Similarity metric:

\[ \text{Sim}( K_1 \ast K_2 ) = \frac{K_1 \cap K_2}{K_1 \cup K_2} \]

An example:

If \( K_1 = \{ \text{father} \equiv \text{man} \wedge \text{parent} \} \)

and \( K_2 = \{ \text{mother} \equiv \text{woman} \wedge \text{parent} \} \), then:

\[ \text{Sim}(K_1, K_2) = (\wedge + \text{parent})/(\text{man} + \wedge + \text{parent} + \text{woman}) \]
\[ = 2/4 = 0.5 \]

Check with Gestalt theory!
Knowlets in practice
(math and science)

- **Concepts:** in the form of one or more words.
- **Theorems:** generally in the form of IF-THEN or IF-THEN-ELSE.
- **Equations:** definitions or a series of derivations.
- **Examples:** applications of theorems and equations for specific values and conditions.
The analysis of Shannon’s paper (without the appendices) reveals at least

- 16 concepts
- 36 relations/properties (these two numbers can only be more or less subjective)
- 9 equations
- 12 theorems: 7 are of the form IF-THEN (equivalent to equations) and the rest 5 are of the form IF-THEN-ELSE
- and 17 examples.

According to our metrics, we have:

\[ H(\text{Sh1948}) = (16 + 36 \cdot 2 + 9 \cdot 3 + 7 \cdot 3 + 5 \cdot 4) \cdot a_1 + H(\text{examples}). \]
Case study (cont’d)

- $H(\text{examples}) = ?$

- Use the fact that:
  - $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + (1/2)^r \ldots \approx 1$ for $r \gg 1$
  - $H(r \text{ examples}) = \alpha_n \log_2[1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + (1/2)^r ]$

- 17 examples:
  - 07 examples are for concepts (distributed as 1, 1, 1, 1, 1, 2)
  - 05 for equations (distributed as 1, 1, 3) and
  - 05 for theorems (one If-Then and 2, 1, 1 If-Then-Else).

- $H(\text{examples}) = 19.6\alpha_1 = 9.8 \text{ bits}$.

- We have finally:
  - $H(Sh1948) = (16+36 \cdot 2+9 \cdot 3+7 \cdot 3+5 \cdot 4) \cdot \alpha_1 + H(\text{examples})$
  - $H(Sh1948) = 175.6\alpha_1 = 87.8 \text{ bits}$. 
Case study (cont’d)

- Exams for the last three years:
  - an average of 10 concepts and 4 equations per exam
  - \( H(\text{exam}) = 22 \cdot a_1 = 11 \text{ bits} \)
    - i.e. **12% of Shannon’s paper**.

- **Note**: analysis was done manually due to lack of appropriate tools; would have been carried out ideally with automatic techniques.
Related work

- To the best of my knowledge, the only works that grade exams and course contents using quantitative metrics are based on Bloom’s six-level learning taxonomy; e.g. Oliver et al. (2004) and Zheng (2008)

- Also work of Walther Umstätter on knowledge measurement: [here](#)

- We therefore believe that we have presented original ideas to assess quantitatively our exams with respect to course contents.
Conclusion:
Future directions

Epistemological questions
Triad Data-Information-Knowledge

Knowledge

Data

Information

Memory recall and/or prediction

Modeling-Learning-Sensing

Validation

Interpretation

Learning

Memorization-learning
The Degrees of Knowledge according to J. Locke

Three degrees or levels of knowledge (like Descartes and Spinoza):

- **Intuitive knowledge:** the most secure kind of knowledge; it corresponds to Descartes’ notion of things we perceive by the natural light, and which are so obviously true that we can’t doubt them.

- **Demonstrative knowledge:** the kind of knowledge we have of most mathematical truths: we can’t just see that they are correct, but we can prove them on the basis of axioms of which we have intuitive knowledge, by means of very small and simple steps which we can intuitively perceive must be valid.

- **Sensitive knowledge:** the knowledge of the existence of things that we acquire by sensing them. (Here Locke needs to respond to Cartesian worries; he does so jokingly in 4.2.14, and more seriously in 4.11.)
Thank you
Resources

- Learning theories: http://www.learning-theories.com
- Walther Umstätter web site
Instructional Phases

- Problem
- Demonstration
- Activation
- Integration
- Application
- Demonstration