Week 3 Video 5

Knowledge Engineering

Knowledge Engineering

- Where your model is created by a smart human being, rather than an exhaustive computer

Knowledge Engineering

- Also called
- Rational modeling
- Cognitive modeling

Knowledge Engineering at its best

- Knowledge engineering is the art of a human being
  - Becoming deeply familiar with the target construct
  - Carefully studying the data, including possibly process data (such as think-alouds)
  - Understanding the relevant theory and how it applies
  - Thoughtfully crafting an excellent model

Knowledge Engineering at its best

- In its classical version
  - A knowledge engineer and a domain expert work together to model the construct
  - Through an iterative process
  - Where the knowledge engineer interviews the expert, creates models, goes through the model and its implications with the expert, gets feedback, enhances the models, and repeats the process
  - Until both the knowledge engineer and domain expert believe the model has fully captured the expert's reasoning

Knowledge Engineering at its best

- Achieves higher construct validity than data mining
- Achieves comparable performance in data
- And can transfer better to new data in some cases, by capturing more general aspects of the construct
(Paquette et al., 2015)

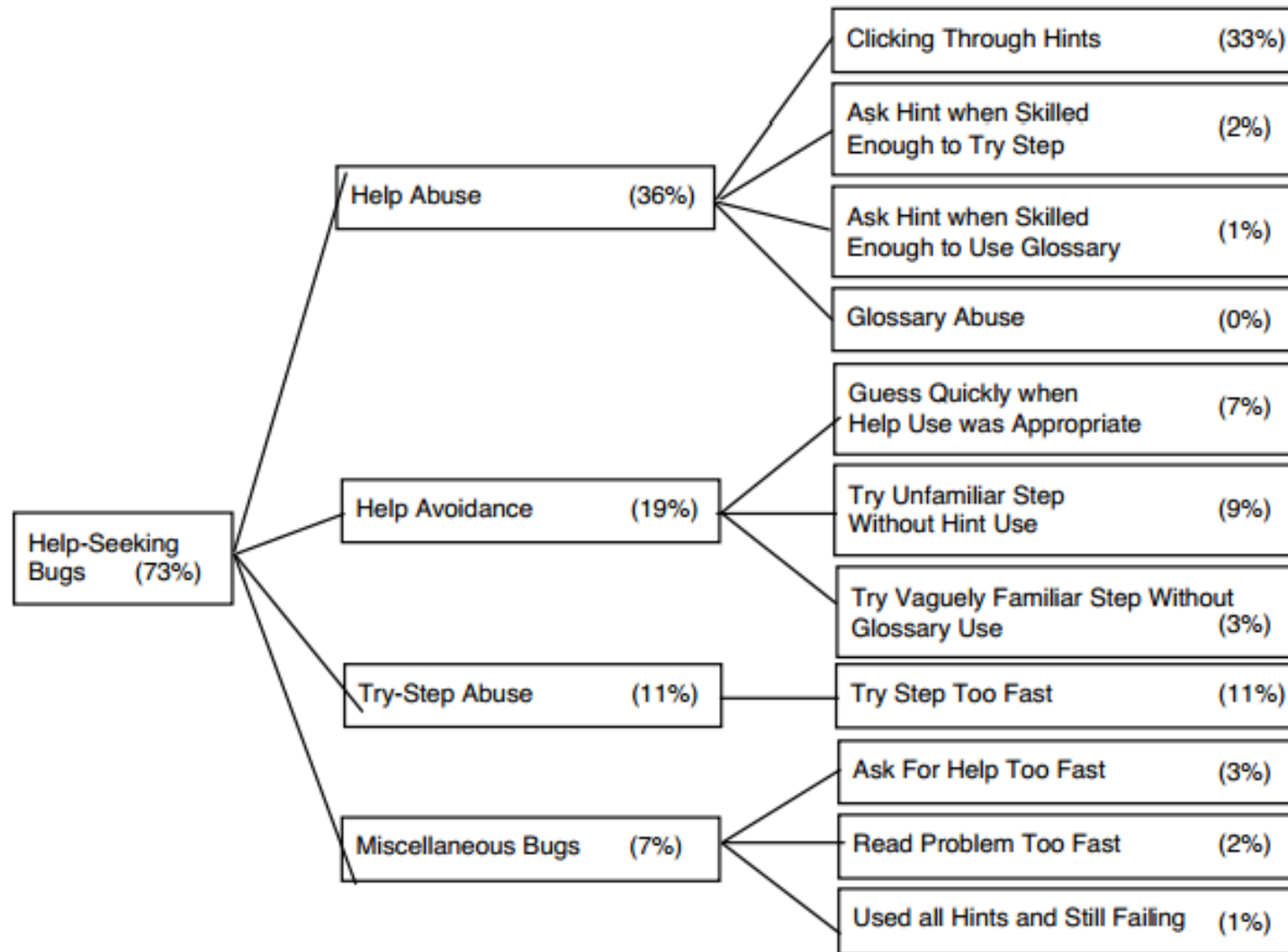# Example of excellent knowledge engineering

- Aleven et al.'s (2004, 2006) help-seeking model

A prescriptive model of good help-seeking behavior in an online tutor

With a taxonomy of errors in student help-seeking

		Clicking Through Hints (33%)
	Help Abuse (36%)	Ask Hint when Skilled Enough to Try Step (2%)
		Ask Hint when Skilled Enough to Use Glossary (1%)
		Glossary Abuse (0%)
Help-Seeking Bugs (73%)	Help Avoidance (19%)	Guess Quickly when Help Use was Appropriate (7%)
		Try Unfamiliar Step Without Hint Use (9%)
		Try Vaguely Familiar Step Without Glossary Use (3%)
	Try-Step Abuse (11%)	Try Step Too Fast (11%)
	Miscellaneous Bugs (7%)	Ask For Help Too Fast (3%)
		Read Problem Too Fast (2%)
		Used all Hints and Still Failing (1%)

Developed based on

- Thorough study of dozens of scientific articles
- Years of experience in designing online learning environments
- Intensive study of log files of student interaction with learning system
- Plus experience watching kids use educational software in real classrooms

Resultant models

- Predictive of student learning (Aleven et al., 2004, 2006) and preparation for future learning (Baker et al., 2011)

- Specific aspects of model correlate to data-mined detectors of same constructs, and improve data-mined models if added to them (Roll et al., 2005)

Knowledge Engineering at its worst

- Knowledge engineering (and the other terms) are sometimes used to refer to
  - Someone making up a simple model very quickly
  - And then calling the resultant construct by a well-known name
  - And not testing on data in any way
  - And asserting that their model *is* the construct, despite having no evidence

Knowledge Engineering at its worst

- Achieves poorer construct validity than data mining
- Predicts desired constructs poorly, sometimes even worse than chance
  - Due to over-simplifying a complex construct
  - Or even failing to match it
- Can slow scientific progress by introducing false results
- Can hurt student outcomes by intervening at the wrong times

How can you tell if knowledge engineering is bad

- If a data mining model is bad
  - It's usually relatively easy to identify, from the features, the validation procedure, or the goodness metrics

- Telling top-notch knowledge engineering from junk is a little harder

- The hard work is in the researcher's brain, and the process is usually invisible

- But… look for very simple models of complex constructs

Whether You Use Knowledge Engineering or Data Mining…

- You should be testing your models on data in some fashion

- Even if you can't get a direct measure (training labels)

- You can usually get some kind of indirect measure (predicting student learning, for example)

It's not an either-or…

- Feature engineering is very closely related to knowledge engineering
  - Careful study of a construct will lead to better features and ultimately to better models

It's not an either-or…

- Using knowledge-engineered models as features in data mining models can be a powerful tool

It's not an either-or…

- Some research has used knowledge engineering to discover what basic operators domain experts think in terms of

- And then re-combine those operators in a broader range of ways
  (Paquette et al., 2014)

It's not an either-or…

- Knowledge-engineered models sometimes depend on cut-offs or numerical parameters that are hard to determine rationally

- These parameters can be empirically fit using data

- Some variants of Aleven et al.'s model do this
- As does Bayesian Knowledge Tracing (next week)

Next Up

- Week 4 – knowledge modeling