

# Towards Evaluating Human-Instructable Software Agents

Robert D. Grant, David DeAngelis, Dan Luu, Dewayne E. Perry

Empirical Software Engineering Laboratory  
Center for Advanced Research in Software Engineering  
University of Texas at Austin

Kathy Ryall

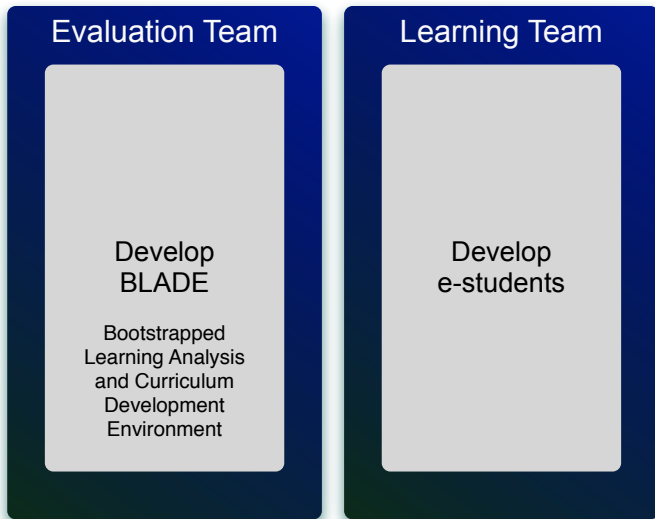
Advanced Information Technologies  
BAE Systems

25 June 2011

## Definition

**Bootstrapped Learning** is a novel approach to machine learning whose goal is to produce computational agents (**e-students**) that can be instructed by human teachers in the same ways that humans instruct one another.

# Background



# Method: Exploratory Case Studies

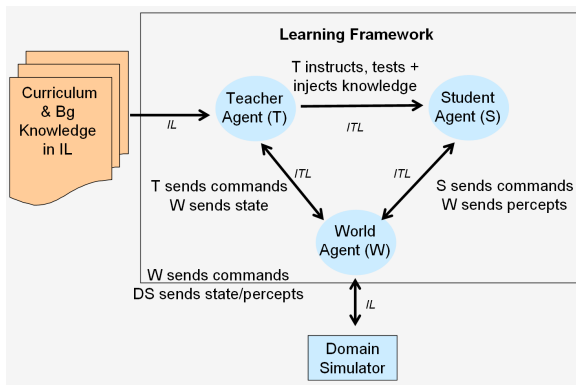
...leading to future controlled benchmarking experiments

- |           |  |  |
|-----------|--|--|
| Phase I   |  | What instruction methods would human teachers try to use to teach an e-student?  |
| Phase II  |  | Can we use human students to develop a benchmark set of lessons and tests (a curriculum) by which to evaluate e-students' success in learning by natural instruction?        |
| Phase III |  | Can we use human students to develop a curriculum covering a broader range of instruction methods and increasing the difficulty and complexity of tasks given to e-students? |

# Hidden Domain

- Must be kept secret until after e-student testing
- Lessons were presented in three forms:
  - by telling,
  - by example,
  - by feedback

# E-student Testing

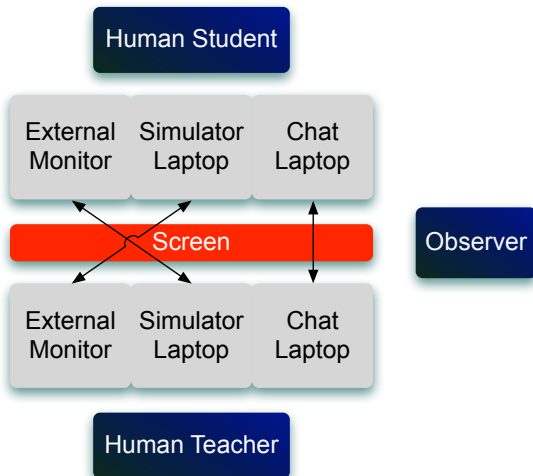


# Human Student Testing

- 15 minutes background knowledge
  - Students interacted with the knowledge domain through a simulator
- 15 minutes pre-test
- 3 hours lessons and quizzes
- 30 minutes post-tests

# Phase II: Initial Study Design

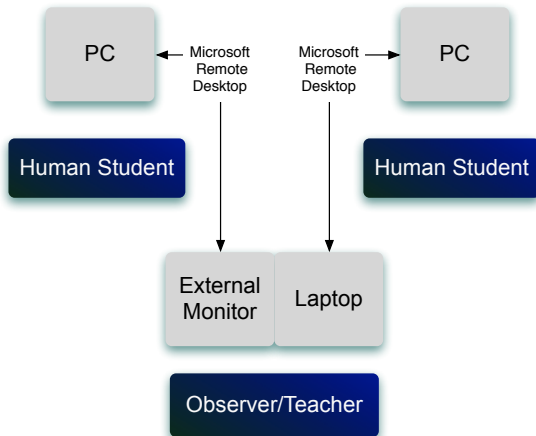
- Direct analog e-student/e-teacher interaction
- Teacher typed transliterations of e-student curriculum into Google Chat, performed examples on simulator laptop
- Took far too long
- Tedious and error prone
- No protocol for reporting or correcting mistakes
- No way for student to ack





# Phase II: Final Study Design

- Self-paced curriculum: PDF and PPT lessons given directly to students
- Example lessons used screenshots
- Feedback lessons in “choose-your-own-adventure” style
- Faster, more scalable, more consistent, more enjoyable





# Qualitative Results

- Human- and e-students differ in fundamental ways
- Small semantic details can be critically important
- Self-pacing and automation were key
- Training versus education was an unexpected issue

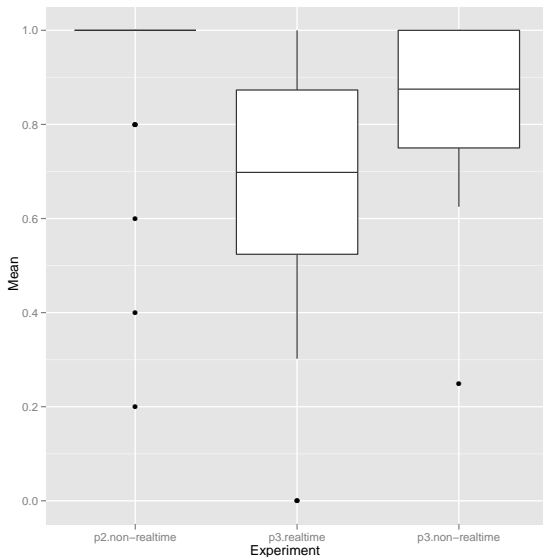
# Questions?

# Phase II Quantitative Results

Pre-Test:	0		000000000000000000000000000000
	10		0

Post-Test:	2		0
	4		0
	6		0
	8		000
	10		000000000000000000000000000000

# Quantitative Results



# Quantitative Results

