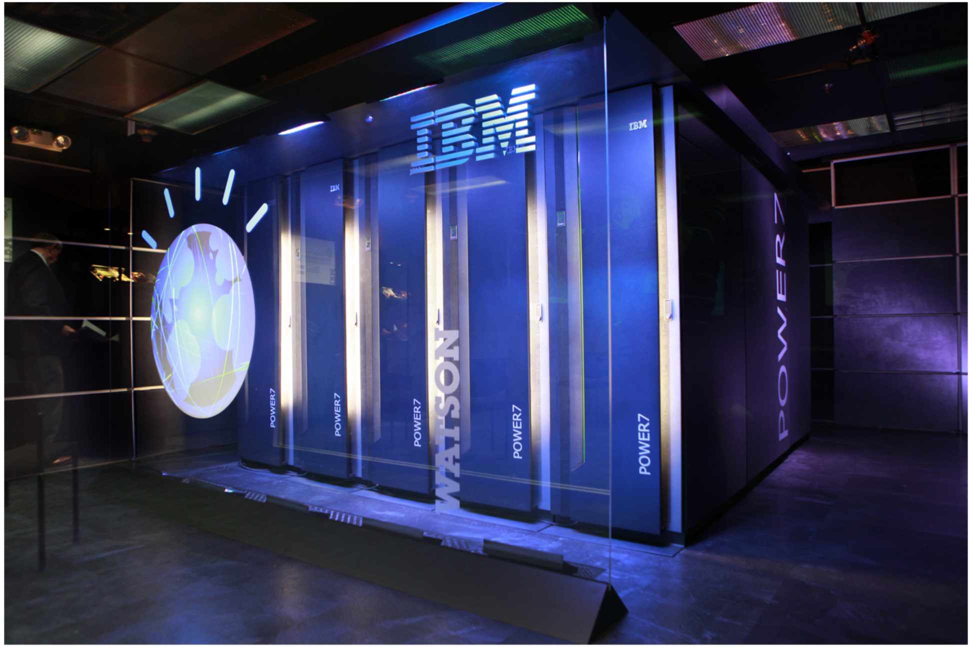
Towards Natural Cognitive System Training Interactions
A Preliminary Framework

Erik Harpstead & Christopher MacLellan

Human-Computer Interaction Institute

SOARTECH
Modeling human reasoning.
Enhancing human performance.

The AI Problem: Knowledge Acquisition / Engineering



Took over a person-century of AI expert development time to build

Laird, J. E., Gluck, K., Anderson, J., Forbus, K. D., Jenkins, O. C., Lebiere, C., … Kirk, J. R. (2017). Interactive Task Learning. IEEE Intelligent Systems, 32(4), 6–21. https://doi.org/10.1109/MIS.2017.3121552

Image by Carolyn Cole—LA Times via Getty Images

Not Realistically "Personalizable"

- Our daily lives are filled with a wide range of tasks across multiple domains
  - But, state-of-the-art AI systems are implemented to perform specific tasks in specific domains.
- Developing each new task requires substantial time and expertise
  - Watson required a century of AI expert man hours
  - Deep learning shifts the burden from programming to data curation and training, but still requires substantial time and expertise

Key AI Innovation: Interactive Task Learning

Aims to develop the **computational** and **cognitive theory** to support task learning that is:

general, *efficient*, *scalable*, and *natural*

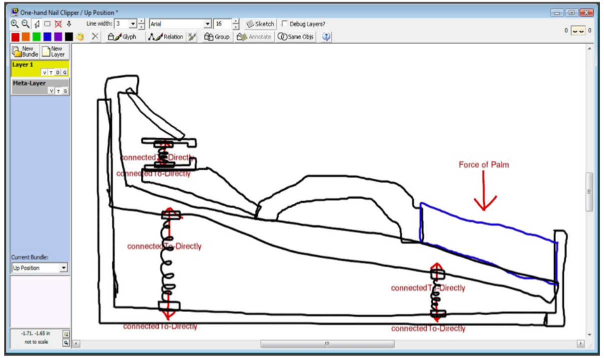
Laird, J. E., Gluck, K., Anderson, J., Forbus, K. D., Jenkins, O. C., Lebiere, C., … Kirk, J. R. (2017). Interactive Task Learning. IEEE Intelligent Systems, 32(4), 6–21. https://doi.org/10.1109/MIS.2017.3121552

Metaphor for *Personalization*

Monolithic Mainframes (1960s)

Personal Computers (1970-Now)

Monolithic Cognitive Systems



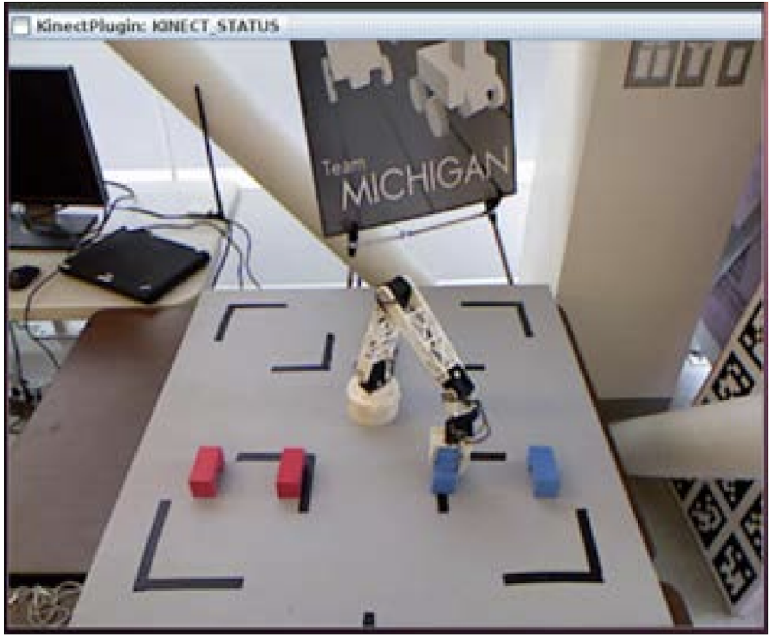
?

Personal Cognitive Systems

Three Examples of Task Learning Systems

Rosie System

	Dialog	Line
M:	The game is tic-tac-toe.	1
R:	I don't know that game, how many players are there?	2
M:	Two.	3
R:	Please start by teaching me the name of a legal action in the game.	4
M:	Place	5
R:	What are the verb and parameter arguments associated with this action?	6
M:	move 1 to 2	7
R:	Ok I know that verb, what kind of object is the first parameter? (block, location, or either)	8
M:	M: block	9
R:	Please list all constraints for this parameter (such as 'it is red' or 'it is on [parameter] 2') and then finished	10
M:	it is red	11
M:	it is not on a location	12
M:	Finished	13
R:	What is the type of the next parameter? (or finished if no more)	14
M:	Location	15
R:	What are the constraints?	16
M:	it is not under a block	17

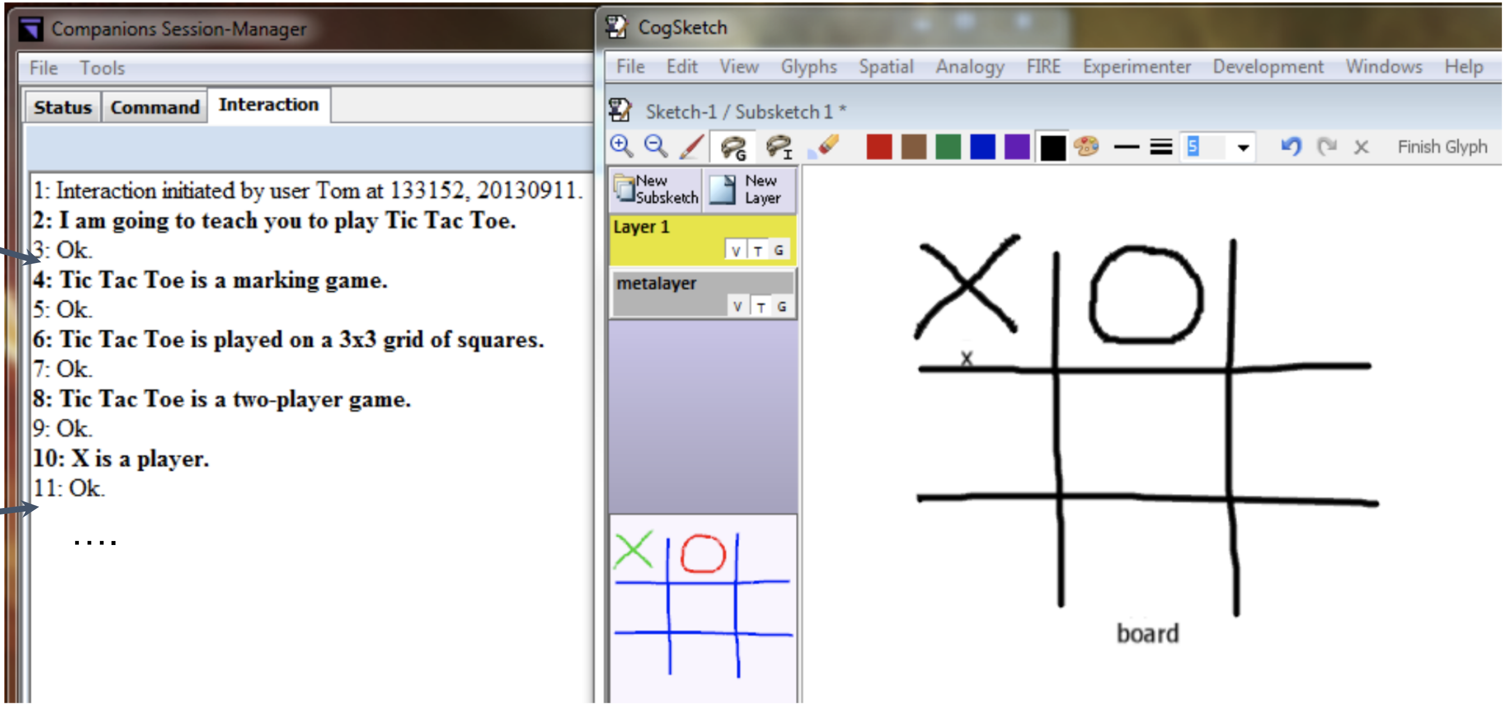


Kirk, J., & Laird, J. E. (2013). Learning Task Formulations through Situated Interactive Instruction. In Second Annual Conference on Advances in Cognitive Systems (pp. 219–236).

Companion Architecture

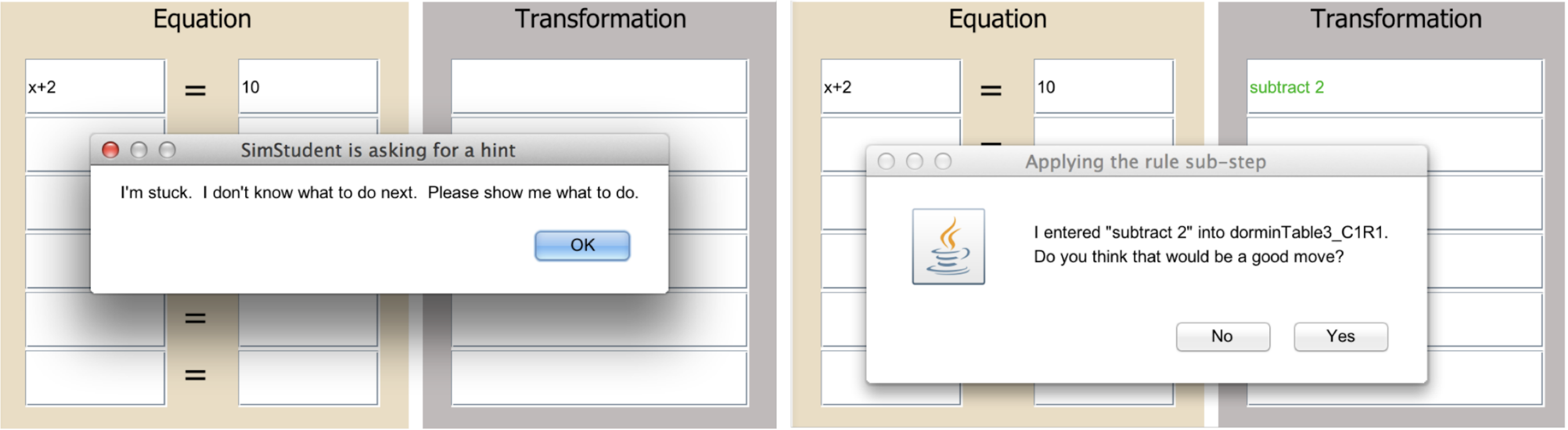
User draws board

User draws the X



Hinrichs, T. R., & Forbus, K. D. (2013). X Goes First: Teaching a Simple Game through Multimodal Interaction. In Second Annual Conference on Advances in Cognitive Systems (pp. 205–218).

Apprentice Learner Architecture



Maclellan, C. J., Harpstead, E., Patel, R., & Koedinger, K. R. (2016). The Apprentice Learner Architecture: Closing the loop between learning theory and educational data. In Proceedings of the 9th International Conference on Educational Data Mining.

An HCI Problem

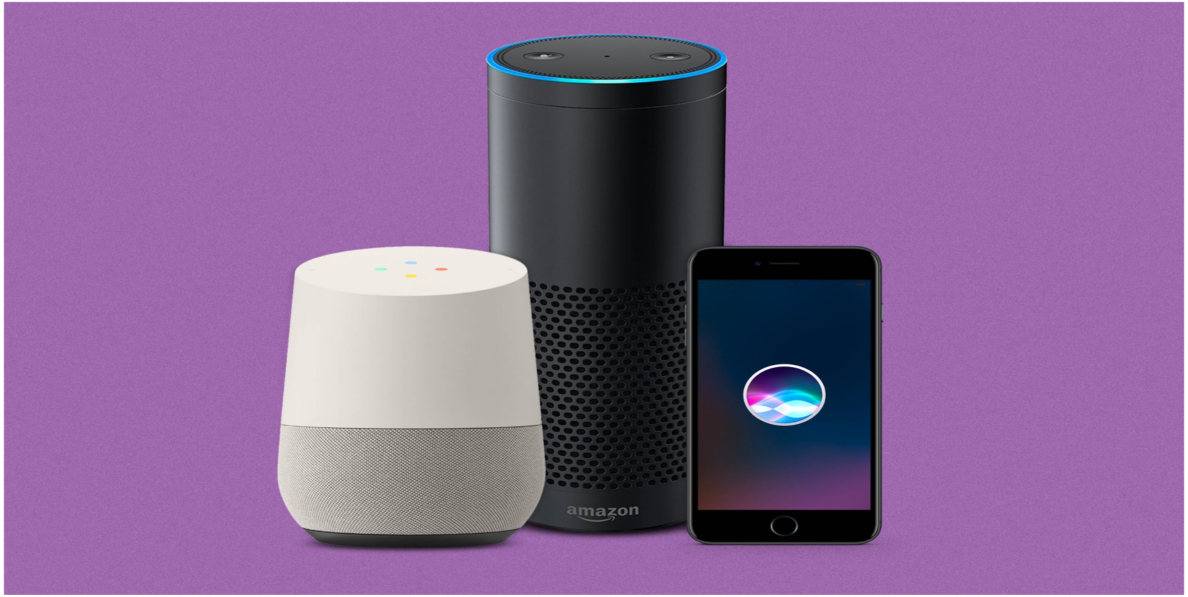
How can we make training interactions in these kinds of systems more *natural*?

Suggestions from HCI Literature

Identified four characteristics of systems that support naturalness from literature:

1. Support the goals of the user
2. Do what the user expects
3. Allow the user to work the way they want
4. Leverage user's prior experience to minimize necessary training

Support the Goals of the User



Do not overemphasize ease of use at the expense of limiting what users can achieve.

Myers, B., Hudson, S. E., & Pausch, R. (2000). Past, present, and future of user interface software tools. ACM Transactions on Computer-Human Interaction, 7(1), 3–28. https://doi.org/10.1145/344949.344959

Do What the User Expects

Humans adhere to patterns, scripts, or norms and systems should be aware of and align with these patterns.

Bicchieri, C. (2006). The grammar of society: the nature and origins of social norms. Cambridge University Press.

Work the way they want

Natural language is not always how work is done

Schön, D. A. (1982). The Reflective Practitioner: How Professionals Think in Action. New York, New York, USA: Basic Books.

Leverage Prior Experience

Systems should support modalities that enable “instant expertise”

Wigdor, D., & Wixon, D. (2011). Brave NUI World: Designing Natural User Interfaces for Touch and Gesture. Burlington, MA: Morgan Kaufmann.

Preliminary Framework

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	Command	Command-Line Interface
Beliefs	Operant Conditioning	Clarify	Control device
Concepts	Direct Instruction	Acknowledge	GUI
Experiences	Apprentice Learning	Inform	Sketch
Skills	After-Action Review	Spotlight	API
Dispositions	Socratic Learning	Annotate	Gesture
	Collaborative Learning	Reward	Speech
		Demonstrate	Text
		Request <type>	Multimodal

Non-Interactive Labeling System

Knowledge	Patterns	Types	Modalities
Goals	**Passive Learning**	Command	Command-Line Interface
Beliefs	Operant Conditioning	Clarify	Control device
Concepts	Direct Instruction	Acknowledge	GUI
Experiences	Apprentice Learning	**Inform**	Sketch
Skills	After-Action Review	Spotlight	**API**
Dispositions	Socratic Learning	Annotate	Gesture
	Collaborative Learning	Reward	Speech
		Demonstrate	Text
	Provide data labels	Request <type>	Multimodal

Interactive Labeling System

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	Command	Command-Line Interface
Beliefs	Operant Conditioning	Clarify	Control device
Concepts	Direct Instruction	Acknowledge	**GUI**
Experiences	**Apprentice Learning**	**Inform**	Sketch
Skills	After-Action Review	Spotlight	API
Dispositions	Socratic Learning	Annotate	Gesture
	Collaborative Learning	**Reward**	Speech
		Demonstrate	Text
	Provide data labels	**Request <type>**	Multimodal
	Mark labels correct/incorrect		
	Request data labels		

Rosie

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	**Command**	Command-Line Interface
Beliefs	Operant Conditioning	**Clarify**	Control device
Concepts	**Direct Instruction**	**Acknowledge**	GUI
Experiences	Apprentice Learning	**Inform**	Sketch
Skills	After-Action Review	Spotlight	API
Dispositions	Socratic Learning	Annotate	Gesture
	Collaborative Learning	Reward	Speech
		Demonstrate	**Text**
		Request <type>	Multimodal

Companion Architecture

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	**Command**	Command-Line Interface
Beliefs	Operant Conditioning	Clarify	Control device
Concepts	**Direct Instruction**	**Acknowledge**	GUI
Experiences	Apprentice Learning	**Inform**	**Sketch**
Skills	After-Action Review	Spotlight	API
Dispositions	Socratic Learning	Annotate	Gesture
	Collaborative Learning	Reward	Speech
		Demonstrate	**Text**
		Request <type>	**Multimodal (text and sketch)**

Apprentice Learner Architecture

Knowledge	Patterns	Types	Modalities
Goals	Passive Learning	Command	Command-Line Interface
Beliefs	Operant Conditioning	Clarify	Control device
Concepts	Direct Instruction	Acknowledge	**GUI**
Experiences	**Apprentice Learning**	Inform	Sketch
Skills	After-Action Review	Spotlight	API
Dispositions	Socratic Learning	**Annotate**	Gesture
	Collaborative Learning	**Reward**	Speech
		Demonstrate	Text
		Request <type>	Multimodal

An Implicit Hypothesis

H1: Similar to human–human teaching, we expect that different pattern–type–modality combinations will be better for transferring different kinds of knowledge into the system

E.g., Teaching *concepts* might be more natural via direct instruction whereas teaching *skills* might be more natural via apprentice learning.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction (KLI) framework: Toward bridging the science-practice chasm to enhance robust student learning . Cognitive Science, 36, 757–798.

Some open questions we have

Q1: What is the role for UX methods in trying to understand this space and inform AI / task learning system development?

Q2: Is a domain-general *natural* training interaction framework possible, or is naturalness of training interactions something that only makes sense on a domain by domain basis?

Q3: What dimensions are missing in the preliminary framework?

Q4: When should machine learning/teaching be put in the hands of the user vs. just having developers prebuild knowledge into systems (maybe using machine learning)?