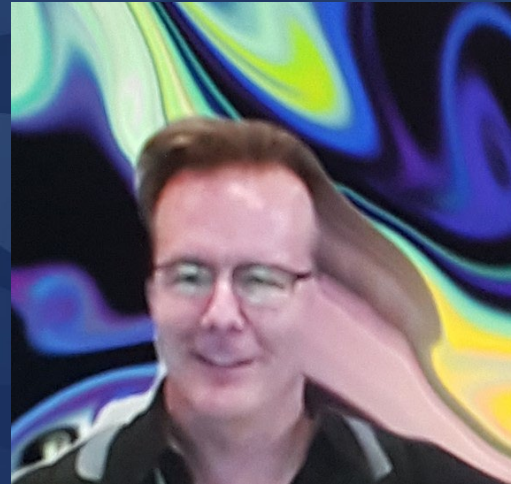


*Conversational Companions: A Challenge Problem**

Mark “Mak” Roberts – Code 5514
Naval Center for Applied Research in AI
4555 Overlook Ave. SW
Washington, DC 20375

Mark.C.Roberts20.civ@us.navy.mil (default email)
Mak.C.Roberts@gmail.com (just in case!)



Research Funded by NRL, ONR, and OUSD Basic Research Office.

AAAI 2025 Bridge Program: Bridging Planning and Reasoning in Natural Language with Foundation Models

Invited Talk: Planning and NL: Applications and Challenges

26 February 2025

**The contents of this talk are my opinion alone, and not an official position of the Navy or the US Government.*

Conversational Companion Example

[two colleagues surreptitiously meet at the top of the escalator]

Hi there ____, it's been too long. How are you and your partner ____ !?!

Great! We moved to Vancouver last fall.

Congratulations! New job or still at ____?

Still at ____; more leadership these days.
I saw you now work at ____?

Yes, though I still travel to Boston frequently.

How's your mother?

She's doing great; thanks for asking.

Hey, I'm eating with ____ for lunch today. Interested?

Sounds great.

Let's meet here at noon.

[escalator ends]

A Conversational Companion:

Tracks
History

Is
Personalized

Attends
to Context

Adapts
quickly

Uses
language

Creates
Plans

Anticipates
the Future(s)

Understands
Embodiment

...

Helpful Frames for Creating Conversational Companions

Logic

- Inductive
- Deductive
- Abductive

Representation

Cognitive Architectures

- Set of system modules
- Memory
- Processes
 - Reasoning
 - Planning
 - Learning

Context

Model Manipulation

- Really “symbol” manipulation
- Goes beyond language
- Deliberation “Functions”
 - Automated Planning
 - Symbolic Reasoning
 - Perceptual processing
 - Generative AI

Characterizing

Frame 1: Logic (aka Reasoning)

“The Myth of Artificial Intelligence” Erik Larson

- States the myth as the inevitability of super-human intelligence
- Posits there are three primary branches of intelligence

Inductive:

- Most of statistical ML fits here

Deductive

- Most “planning” or mathematical logic fits here

Abductive

- “Leap of reasoning”
- This is why he thinks the “myth” is false

Conversational Companions

I find Larson’s
three branches
helpful.

Companions will
probably need all
three branches.

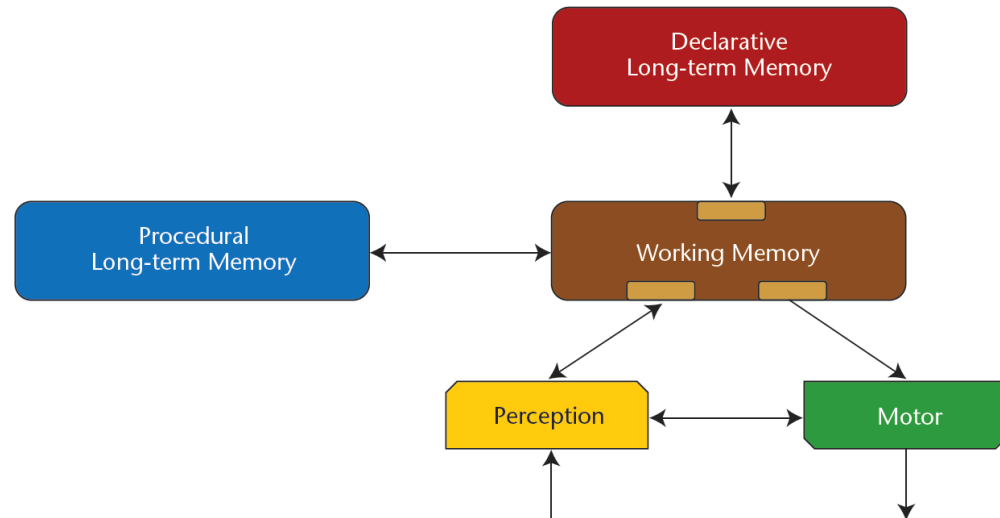
Representation is central!

Consider No Free Lunch Proofs wrt.
problems, algorithms, *and representation.*

Larson, Erik J. (2021). *The myth of artificial intelligence: why computers can't think the way we do*. Cambridge, Massachusetts London, England: The Belknap Press of Harvard University Press. https://en.wikipedia.org/wiki/Erik_J._Larson#The_Myth_of_Artificial_Intelligence

Frame 2: Cognitive Architectures

Standard Model of Mind



“cognitive architectures are an appropriate computational abstraction for defining a standard model, although the standard model is not itself such an architecture.”

- Long-term memory: how to do things, what they mean, consolidated experience
- Working memory: the agent “thinks” about
- Priming: improve recall or learning using context

Conversational Companions

Memory

Probably more important than we emphasize.

Working Memory:

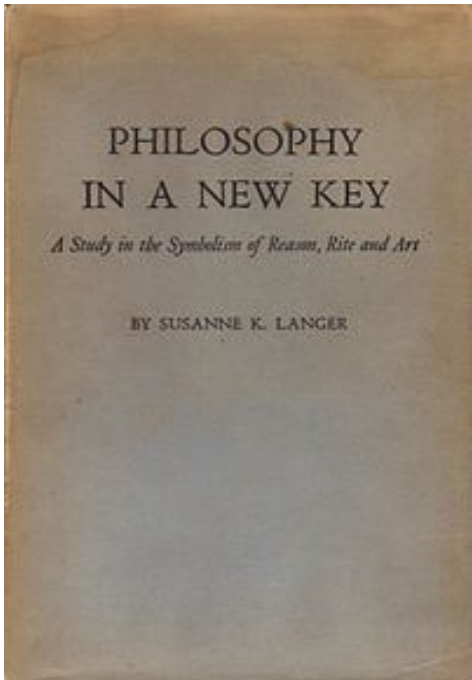
Kind of context similar to the LLM Prompt

Context is critical.

Goes beyond a prompt: involves internal and external “state” spanning multiple, simultaneous horizons of concern

J. E. Laird, C. Lebiere, and P. S. Rosenbloom, “A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics”, *AI Mag*, vol. 38, no. 4, pp. 13-26, Dec. 2017. <https://doi.org/10.1609/aimag.v38i4.2744>

Frame 3: Model Symbol Manipulation



- “Symbol” in the linguistic sense
 - The mind is a symbol-making machine: language, ritual, art
- Most thought is really “symbol” manipulation
 - Language is central, but..
 - Thinking goes beyond language
- Deliberation “Functions”
 - Learning appropriate abstractions*
 - Automated Planning
 - Symbolic Reasoning
 - Perceptual processing
 - Generative AI
 - Other “modules” of cognitive architectures

Conversational Companions

Hybrid approaches are likely more general.

Intelligence is multi-faceted:

IQ, EQ, Fine motor skill, etc.

Let’s focus on characterizing *when* and *why* an approach is appropriate.

From that, we can build descriptive, predictive, proscriptive and prescriptive models specific to the application.

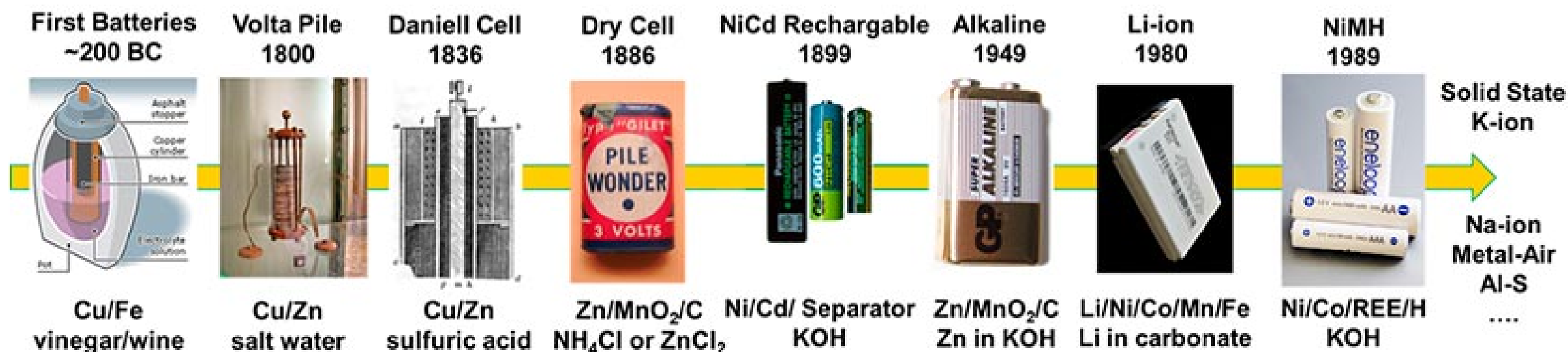
S. K. Langer, *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art, Third Edition*. Harvard University Press, 2009.

G. Konidaris, “On the necessity of abstraction,” *Current Opinion in Behavioral Sciences*, vol. 29, pp. 1–7, Oct. 2019, doi:[10.1016/j.cobeha.2018.11.005](https://doi.org/10.1016/j.cobeha.2018.11.005).

Some Gotchas

- Generative AI Hallucination
 - All models hallucinate, some just do it better than others.
- Bias
 - “I just want to eliminate human bias in the decision making.”
 - AI Mistakes are very different from human mistakes
 - <https://www.schneier.com/blog/archives/2025/01/ai-mistakes-are-very-different-from-human-mistakes.html>
- Ethical or Societal Concerns
 - Is it time to work towards certification of AI systems, as done in other engineering?
- Transparency
- Provenance of data or decisions
- Computational cost of LLMs vs. edge computing
 - Are we creating yet another haves and have nots situation?

AI is a little bit like early battery technology



<https://www.helios-h2020project.eu/news/batteries-long-history-powerful-future>

Morgan Fine-Morris

NRC Postdoc thru July

PhD: Lehigh

Hector Munoz-Avilla

Topics:

- Learning Hierarchical Planning Models
- CNN Approaches
- LLM Approaches
- Plan4LLM Workshop

Vincent Hsiao

NRC Postdoc thru Dec.

PhD: UMD

Dana Nau, Rina Dechter

Topics:

- Game Theory
- Curriculum Learning
 - AAMAS Paper
- All things LLM
- Plan4LLM Workshop

David Porfirio

NRC Postdoc -> NRL

PhD: Wisconsin-Madison

Bilge Mutlu

Topics:

- HRI
- User Interfaces for Expressing Robot Planning
- AAMAS and HRI papers

Working Hypothesis

Conversational Companions must employ, at a minimum:

- Long-term and short-term **memory representations**
 - Learning or adjustment of memory over time
 - Knowledge that captures **abstractions, skills, and experience**
 - some of this will certainly be hierarchical and domain-specific
- Retrieval based on internal and external **context**
- **Adaptable reasoning** from first principles or reflexes as appropriate
 - Exploration of alternative “**future**” **simulations**
- Understanding of others’ **preferences and history**
- **Social Norms**
- **NLP** and **Story Understanding**

There may be more here...

Conversational Companions: Some possible requirements?

- **Planning**
 - Over multiple horizons
 - Multiple agents
 - Exogenous events
 - Abstraction techniques like Hierarchy
 - Domain-specific Planning
 - Rules; methods; Compiled plans
 - Beyond “symbolic planning”
 - Novel or domain-specific heuristics
- **Reasoning**
 - “What if” reasoning
 - Epistemic / Theory of Mind / Analogy
 - Model reconciliation
 - Causality
 - Recognition & PAIR approaches
- **Memory**
 - Long-term
 - Episodic (agent-centric experience)
 - Procedural
 - Social and Cultural Norms
 - Short term / Current state
 - Context-sensitive retrieval
 - “Learning” into long-term memory
- **NLP**
 - Interaction
 - Story Understanding
 - Dialog
 - Norms

Use non-FM approaches too!



Conversational Companions: A Challenge Problem

How do we support long-running agents that converse with their human collaborators?

What is the role of memory in explainable intelligence?

The gift of memory's an awful curse.
With age, it just get much worse.
But I won't mind...

Stable Song - Death Cab for Cutie



Dr. Mark "Mak" Roberts – Code 5514
Naval Center for Applied Research in AI
Mark.C.Roberts20.civ@us.navy.mil
Mak.C.Roberts@gmail.com