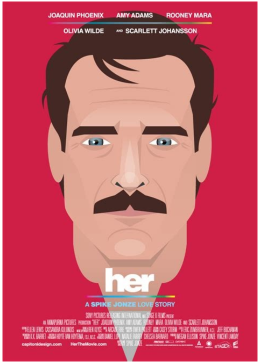
Humans of AI

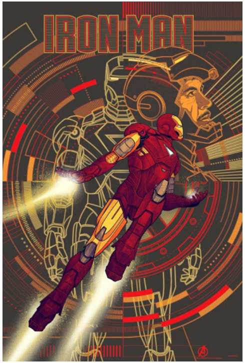
Modeling Humans for Designing Effective Collaborative AI Systems

Shiwali Mohan

February 24, 2021

Senior Member of Research Staff, Palo Alto Research Center

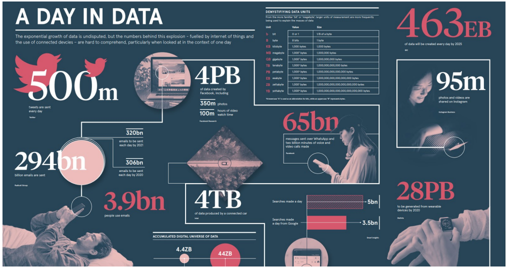


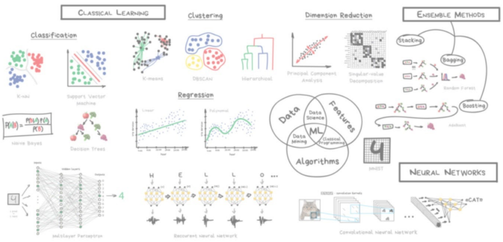


Intelligent collaborators:

independent, long-living entities with
goal-driven, problem-solving behavior who
interact and communicate with humans
learning from their experience

AI Algorithmic Research





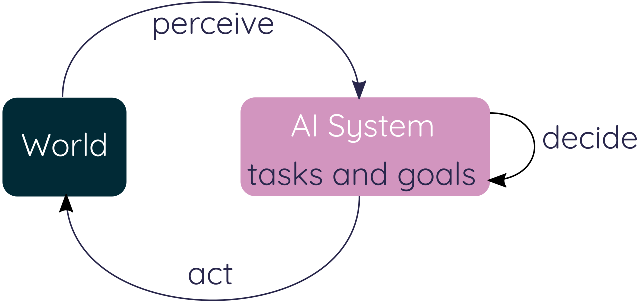
Show the proposed method is better than SOTA on a task-agnostic metric.

Will algorithmic research by itself will lead to intelligent collaborators?

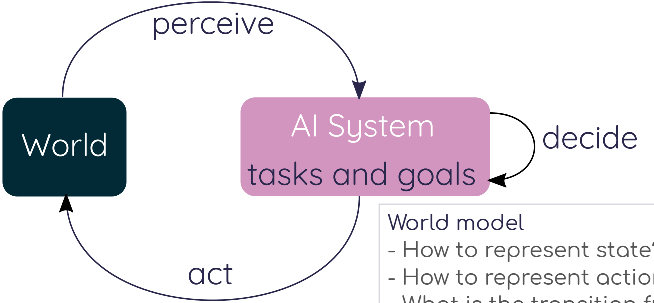
AI Systems Research: Allen Newell, John Laird, Ken Forbus, Yolanda Gil, Ashok Goel, Milind Tambe, and several others

Unified Theories of Cognition: from models to architecture

AI as a System View

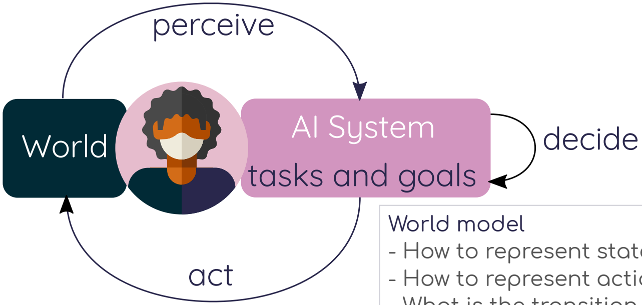


AI as a System View



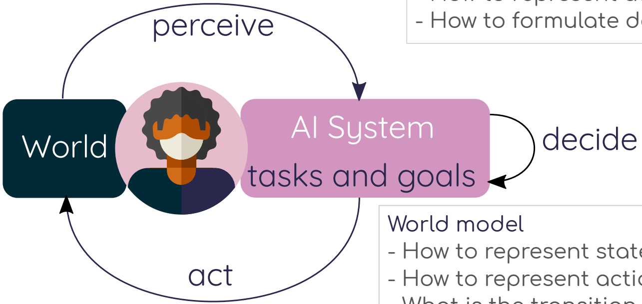
World model
- How to represent state? (s)
- How to represent action? (a)
- What is the transition function?
- How are decisions evaluated?

AI as a System View



World model
- How to represent state? (s)
- How to represent action? (a)
- What is the transition function?
- How are decisions evaluated?

AI as a System View

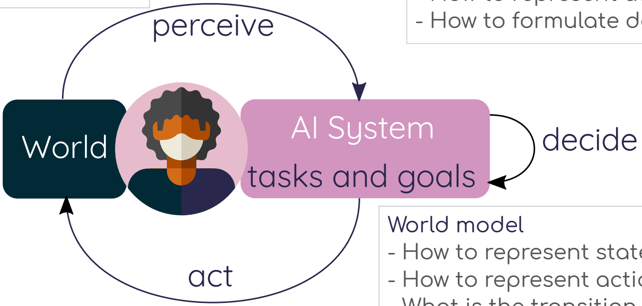


Human model?
- How to represent state?
- How to represent action?
- How to formulate decisions?

World model
- How to represent state? (s)
- How to represent action? (a)
- What is the transition function?
- How are decisions evaluated?

AI as a System View

Artificial Intelligence Journal 2018
Albrecht, S. V., & Stone, P. 2018. *Autonomous Agents Modelling Other Agents: A Comprehensive Survey and Open Problems.* Artificial Intelligence, 258, 66-95.



Human model?
- How to represent state?
- How to represent action?
- How to formulate decisions?

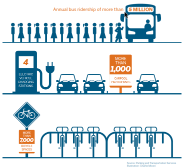
World model
- How to represent state? (s)
- How to represent action? (a)
- What is the transition function?
- How are decisions evaluated?

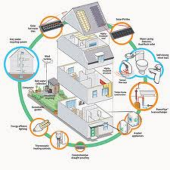
AAAI Presidential Address 2018
Khambhampat, S. 2018. *Challenges of Human-Aware AI Systems*

How do we build intelligent collaborators?

- that are designed to support the goals of a human partner
- that model the human partner explicitly
- that have effective performance on human tasks

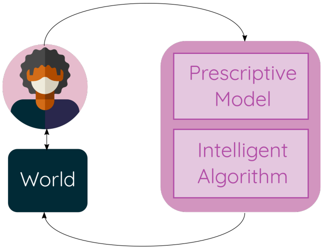
user programmable robots, augmented reality task assistant, mHealth, sustainable communities





A Constrained Approach

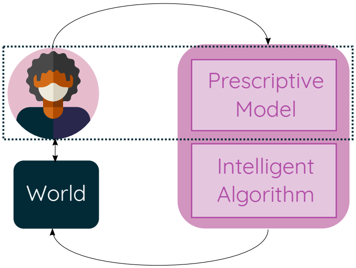
1. Define a real-world problem where success crucially depends on modeling the human collaborator
2. Develop human-centered desiderata, metrics, and evaluation
3. Design prescriptive models from insights of human-centered sciences
    - Indexical Model (Glenberg and Robertson 1999) is not computational
    - Choice theory (Tversky and Kahnemann 1987) is descriptive
4. Adapt AI algorithms to work in collaborative settings
    - Human-centered sciences provide useful desiderata for system behavior
5. Embody in end-to-end interactive systems, demonstrate efficacy
6. Refine and iterate



Collaborative Human-AI System

1. Why collaborative human-AI systems?
2. Sustainable transportation: modeling the human collaborator
3. Interactive task learning: designing AI systems for collaborative settings
4. Health behavior change: evaluating intelligent collaborators in ecological settings
5. Humans of AI: A Research Agenda

Sustainable Transportation

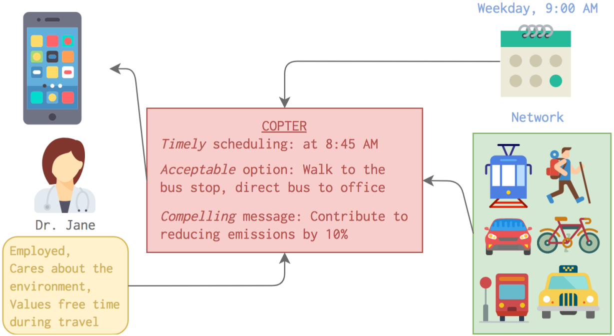


Energy Consumption in Transportation

- Transportation is one of the largest consumers of energy - 29% of energy in US in 2016
- It is far from efficient - both under and over utilization of networks
- Congestion wastes 6.1 billion hours and 3.1 billion gallons fuel per year (Schrank et al. 2015)
- ARPA-E TransNet: energy efficient transportation is an important technology and policy problem
- Success depends on understanding how humans are influenced

David Schrank, Bill Eisele, Tim Lomax, & Jim Bak. *2015 Urban Mobility Scorecard.* Annual Urban Mobility Scorecard (2015)

The Influence Problem

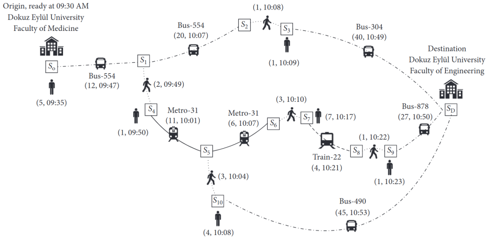


Shiwali Mohan, Hesham Rakha, and Matt Klenk. *Acceptable Planning: Influencing Individual Behavior to Reduce Transportation Energy Expenditure of a City.* Journal of Artificial Intelligence Research 66 (2019)
Shiwali Mohan, Frances Yan, Victoria Bellotti, Ahmed Elbery, Hesham Rakha, and Matt Klenk. *On Influencing Individual Behavior for Reducing Transportation Energy Expenditure in a Large Population.* Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. (2019)
Shiwali Mohan, Matt Klenk, and Victoria Bellotti. *Exploring How to Personalize Travel Mode Recommendations For Urban Transportation* ACM Intelligent User Interfaces Workshops'19, (2019)

Multi-modal Transportation Planning

AI planning theory; Bast et al. (2016), Dvorak et al. (2018)

- Multiple edges between nodes
    - $G = (V, E), lbl : E \rightarrow \Sigma$
- Regular expression language for plausible plans for every user
    - $L(u) = w * (d + |b+)w*$
- Solve $\pi^* = \arg\min_{\pi \in \Pi} \sum_{e \in \pi} cost(e_i)$
    - cost as a function of energy, money, and duration



Filip Dvorak, **Shiwali Mohan**, Victoria Bellotti, and Matthew Klenk. *Collaborative Optimization and Planning for Transportation Energy Reduction.* In ICAPS Proceedings of the 6th Workshop on Distributed and Multi-Agent Planning (2018).
Bast, H.; Delling, D.; Goldberg, A.; Müller-Hannemann, M.;Pajor, T.; Sanders, P.; Wagner, D.; and Werneck, R. F. *Route Planning in Transportation Networks.* In Algorithm Engineering. (2016)

Understanding Choice

Rational choice theory from behavioral economics; Domencich & McFadden (1975), Tversky and Kahnemann (1985)

- Determine a set of available alternatives
  - car, walking, bus, train
- Evaluate utility using attributes relevant to the decision
  - mode dependent: cost, distance, time
  - person dependent: income, education

$$\mathrm{val}(x_i, p) = \gamma_1 \times x_{i1} + \ldots + \gamma_k \times x_{ik} + \lambda_1 \times f_{p1} + \ldots + \lambda_l \times f_{pl}$$

- Probabilistic utility maximization; multinomial logit assumption

$$\Pr(i, p) = \frac{e^{\mathrm{val}(x_i, f_p)}}{\sum_{j \in C} e^{\mathrm{val}(x_j, f_p)}}$$

Amos Tversky & Daniel Kahneman. *Rational Choice and the Framing of Decisions.* Journal of Business. (1986)
Domencich, T. A., & McFadden, D. Urban Travel Demand - A Behavioral Analysis. Transport and Road Research Laboratory (1975)

Defining Acceptability

- Dr. Jane's utility of usual route, $\mathrm{val}(x_u, f_p)$
- Dr. Jane's utility of recommended route, $\mathrm{val}(x_r, f_p)$
- Dr. Jane's switching cost (- switching gain)
- Higher switching gain, more acceptable plan, better adoption

$$\frac{e^{\mathrm{val}(x_u, f_p)}}{e^{\mathrm{val}(x_r, f_p)}} = \frac{\Pr(u,p)}{\Pr(r,p)}$$

$$e^{\mathrm{val}(x_u, f_p) - \mathrm{val}(x_r, f_p)} = \frac{\Pr(u,p)}{\Pr(r,p)}$$

$$\Delta_{u,r} = \mathrm{val}(x_u, f_p) - \mathrm{val}(x_r, f_p) = \ln \frac{\Pr(u,p)}{\Pr(r,p)}$$

$$\Delta_{r,u} = -\Delta_{u,r} = \ln \frac{\Pr(r,p)}{\Pr(u,p)}$$

Estimating Acceptability

Machine learning methods; random forest and multi-layer perceptron

- Problem: multi-class prediction
- Dataset: trip data (CHTS) from CalTrans 2012 - 2012
- Features: trip related (distance), person-related (demographics), network-related (transit pass, license), experience (bike trips in the past week)
- Hypothesis: Dr. Jane's utility function is close to others' who are similar to

Table 1: F1 scores on 20% test set

Mode	Baseline 1	Baseline 2	RF	MLP
Walk	0.00	0.12	0.82*	0.62
Cycle	0.00	0.00	0.81*	0.28
Bus	0.00	0.02	0.78*	0.38
Subway/train	0.00	0.00	0.58*	0.05
Drive	0.72	0.56	0.93*	0.86
Ride	0.00	0.28	0.84*	0.65
Motorcycle	0.00	0.00	0.80*	0.00
Total	0.68	0.40	0.88*	0.74
Category				
Non-motorized	0.00	0.05	0.83*	0.60
Public transit	0.00	0.14	0.79*	0.43
Motorized	0.90	0.82	0.97*	0.93
Total	0.68	0.70	0.94*	0.86

Evaluating Acceptability

Stated preference choice experiments from behavioral economics

49 (27 female, 22 male) drivers in LA

1. Profiler survey:
   - Classifier features
   - Regular weekly trips
2. Choice experiments
   - 10 per participant

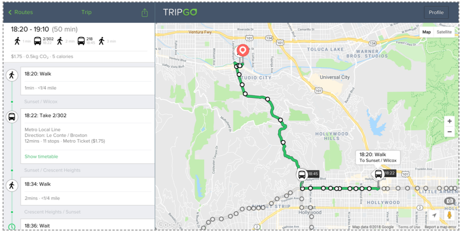
Imagine you are making your following usual trip:

Film Networking Events: from 1475 Wilcox Avenue, los angeles, CA 90038 to 4024 Radford Ave, Studio City, CA 91604
Day: Thursday
Usual departure time: 07:00 PM
Usual arrival time: 07:20 PM

Half an hour before you have to leave, a commuter app on your phone makes the following recommendation to make your travel more environment-friendly. Would you consider taking the recommendation to reach your destination?

EcoTripTip: Take public transit today
Departure time: 06:20 PM
Expected arrival time: 07:10 PM

Please wait for the route to load.



Extremely unlikely	Moderately unlikely	Slightly unlikely	Neither likely nor unlikely	Slightly likely	Moderately likely	Extremely likely
-3	-2	-1	0	1	2	3

I will take the recommended route for my travel.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., . . . Scarpa, R., et al. *Contemporary Guidance for Stated Preference Studies.* Journal of the Association of Environmental and Resource Economists (2017)

Modeling Impact of Acceptability on Adoption

Mixed-effects linear for ordinal adoption

$y = \alpha + \beta x + \gamma z + \epsilon$

Mixed-effects logit for binary adoption

$\Pr(y) = 1/(1 + e^{-(\alpha+\beta x+\gamma z+\epsilon)})$

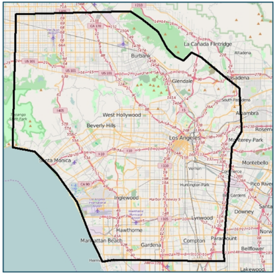
Dependent variables→ **Independent variables↓**	Adoption (ordinal)	Adoption (binary)
(intercept)	-0.017	-0.185
switching gain, $\Delta_{r,u}$	0.108*	0.104*
R^2m	0.034	0.035
R^2c	0.347	0.270
(intercept)	-0.949	-1.065
odds, $e^{\Delta_{r,u}}$	2.386***	2.159*
R^2m	0.075	0.064
R^2c	0.379	0.300
(intercept)	-0.964	-1.080
probability, $\Pr(r, p)$	3.623***	3.317*
R^2m	0.066	0.058
R^2c	0.369	0.293

Acceptable Planning

1. Generate a mode candidate set for Dr. Jane; $M = \{walk(w), bus(b), subway(s)\}$.
2. Determine regex language for valid plans; $L_p = \{w*, w*b+w*, w*s+w*\}$
3. Generate the most time-efficient plan for each element; Π_p.
4. Compute the energy reduction in each plan using energy models (Elbery et al 2018)
5. Evaluate the likelihood of adoption
6. Select a plan that has maximal expected energy savings

Potential Impact of Acceptable Planning

- Agent-based models from complexity sciences
  - LA transportation network: 170,000 roadway links, 1 million daily trips
  - State-of-the-art simulation model of the LA region (Elbery et al. 2018)
- Influence experiment: 10% influenced population
- Expected outcome: 4% energy and 20% time savings in LA, mode shift in influenced population



AM (7am-12pm)	Baseline	Influence	Change (CI)
Total Fuel (l)	3,195,637	3,048,278	-4.6% (-3.6% -5.6%)
Total Delay (hr)	249,221	199,395	-20% (-13.6% -26.4%)
PM (4pm-9pm)	**Baseline**	**Influence**	**Change (CI)**
Total Fuel (l)	3,487,982	3,367,675	-3.5% (-2.6% -4.3%)
Total Delay (hr)	375,137	322,228	-14.1% (-10% -18%)

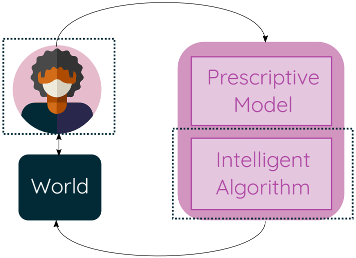
Mode	AM Share	PM Share
Car	54%	53%
Walk	42.7%	42.8%
Bike	3.6%	3.8%
Bus	38.9%	39.1%
Train	14.4%	14%

Ahmed Elbery, Filip Dvorak, Jin Du, Hesham Rakha, & Matt Klenk. *Large-scale Agent-based Multi-modal Modeling of Transportation Networks - System Model and Preliminary Results.* 4th Conference on Vehicle Technology and Intelligent Transport Systems (2018)

Publications

1. **Shiwali Mohan**, Hesham Rakha, and Matt Klenk. *Acceptable Planning: Influencing Individual Behavior to Reduce Transportation Energy Expenditure of a City.* Journal of Artificial Intelligence Research 66 (2019)
2. **Shiwali Mohan**, Matthew Klenk, and Victoria Bellotti. *Exploring How to Personalize Travel Mode Recommendations For Urban Transportation." IUI Workshops* (2019)
3. **Shiwali Mohan**, Frances Yan, Victoria Bellotti, Ahmed Elbery, Hesham Rakha, Matt Klenk *On Influencing Individual Behavior for Reducing Transportation Energy Expenditure in a Large Population.* Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (2019)
4. Filip Dvorak, **Shiwali Mohan**, Victoria Bellotti, Matt Klenk. *Collaborative Optimization and Planning for Transportation Energy Reduction.* ICAPS Proceedings of the 6th Workshop on Distributed and Multi-Agent Systems (2018)
5. Matt Klenk, Victoria Bellotti, Filip Dvorak and **Shiwali Mohan**. Palo Alto Research Center Inc, 2021. *Generating Collaboratively Optimal Transport Plans.* U.S. Patent 10,885,783.
6. Matt Klenk, Victoria Bellotti, and **Shiwali Mohan**. Palo Alto Research Center Inc, 2020. *User Behavior Influence in Transportation Systems.* U.S. Patent Application 16/181,152.

Interactive Task Learning

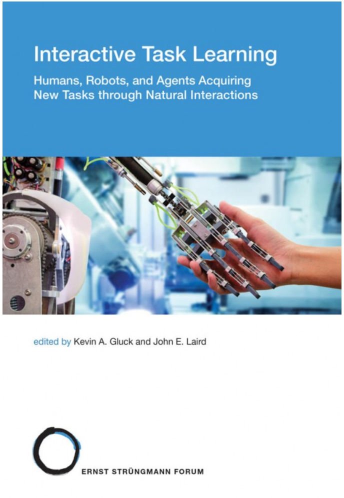


Deployment in Dynamic Environments

- AI designers cannot predict all deployment usecases
- AI should be designed to enable non-expert human programming
- Success depends on understanding how do humans teach and learn

Interactive Task Learning - ESF 2017

Machine learning: Tom Mitchell
Cognitive architectures: John Laird, Ken Forbus, Christian LeBiere, Paul Rosenbloom, Shiwali Mohan
Robotics: Andrea Thomaz, Julie Shah, Maya Cakmak, Peter Stone, Matthias Scheutz
Psychology: Ken Koedinger, Suzanne Stevenson, Andrea Stocco, Peter Pirolli
Computational Linguistics: Joyce Chai, Parisa Kordjamshidi, Candy Sidner



ITL is a Different Paradigm

Classical machine learning

- Batch: dataset -> model
- Phased: training -> testing
- Passive: learn when asked
- Big data
- Data: confounding

Interactive learning

- Incremental: experience -> knowledge
- Online
- Active: learn when failed
- Small data
- Teacher: benevolent

Ishaan

Soar Family of ITL Systems

Built with Soar (Laird 2012); enhanced with computer vision and control. Michigan Rosie - http://soargroup.github.io/rosie/

2012 First demonstration of Rosie an end-to-end interactive learning system. **Mohan** et al. ACS 2012.

2013 Cognitively plausible model of learning from instruction. **Mohan** et al. ICCM 2013.

2014 Defined indexical language comprehension for embodied agents. **Mohan** et al. ACS 2014

2014 First demonstration of interactive explanation-based learning for robots. **Mohan** and Laird AAAI 2014.

2014 Rosie learns over 10 table-top games from interactions. Kirk and Laird 2014, Kirk, Mininger, and Laird 2016.

2016 Rosie learns with perceptual untertainty. Mininger and Laird ACS 2016.

2017 Interactive Task Learning defined at the Ernst Strungmann Forum. Laird et al. IEEE Intelligent Systems 2017.

2018 First demonstration of learning goal-oriented and procedural tasks with interactive EBL. Mininger and Laird AAAI 2018

2018 Learning Fast and Slow: Levels of Learning in General Autonomous Intelligent Agents. Laird and **Mohan** AAAI 2018 (Blue Sky Award)

2018 Interactive Task Learning, MIT Press. Gluck and Laird, 2018.

2019 Rosie learns over 40 table-top games from interactions. Kirk and Laird IJCAI 2019

2020 First demonstration of joint concept and language learning with analogical processing. **Mohan** et al. ACS 2020

2021 First observational analysis of human teaching. Ramaraj et al. [arXiv preprint arXiv:2102.06755]

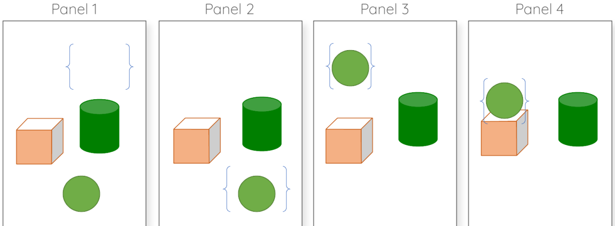
Embodied Language Processing

DARPA GAILA: Where does meaning come from?

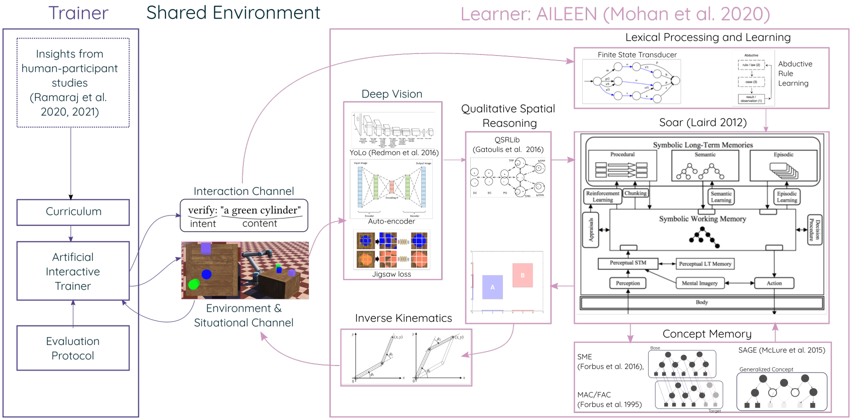
- NLP (BERT, GPT) derives meaning from statistical patterns in word usage
- GLP (VQA task) derives meaning from paired visual/linguistic stimuli

The Indexical Hypothesis (Glenberg and Robertson, Discourse Processes 1999)
Language - a mechanism to guide attention to relevant seen and unseen elements on the environment and compose them for useful action.
The Indexical Model of Comprehension (Mohan et al. ACS 2014); Similar in philosophy to DMAP (Livingston and Reisbeck 2009)

"Move the green sphere onto the red cube."

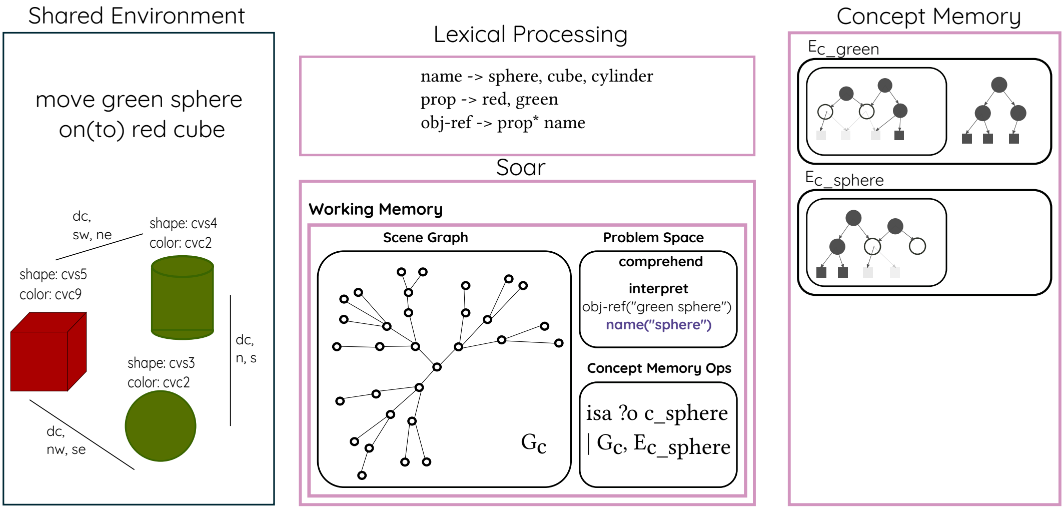


AILEEN



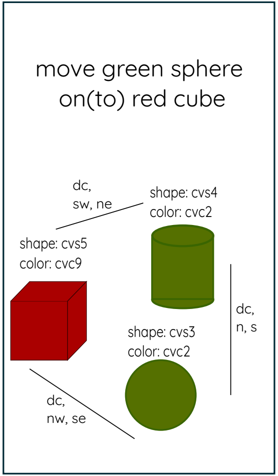
Shiwali Mohan, Matt Klenk, Matthew Shreve, Kent Evans, Aaron Ang, John Maxwell. *Characterizing an Analogical Concept Memory for Newellian Cognitive Architectures.* Proceedings of the Eighth Annual Conference on Advances in Cognitive Systems (ACS). 2020

Interactive Concept Learning

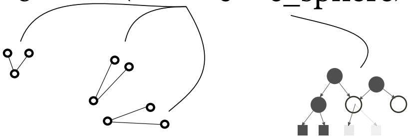


Interactive Concept Learning

Shared Environment



$$M_o = \mathrm{sim}(o \text{ in } G_c, E_{c\_sphere})$$



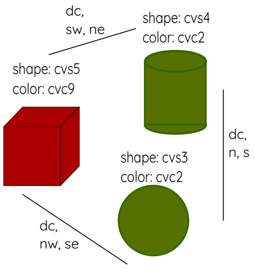
$$\text{if } M_o > t,\ \mathrm{assert}(G_c,\ o \text{ isa } c\_sphere)$$

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). *Extending SME to Handle Large-Scale Cognitive Modeling*. Cognitive Science, 41, 1152–1201

Interactive Concept Learning

Shared Environment

move green sphere on(to) red cube

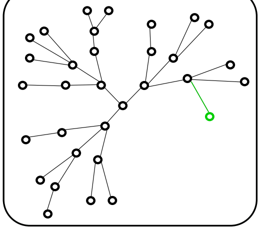


Lexical Processing

name -> sphere, cube, cylinder
prop -> red, green
obj-ref -> prop* name

Working Memory

Scene Graph



Problem Space

comprehend

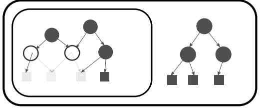
interpret
obj-ref("green sphere")
name("sphere")

Concept Memory Ops

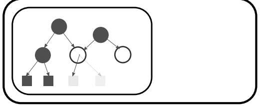
isa ?o c_sphere | G_c, E_c_sphere

Concept Memory

E_c_green



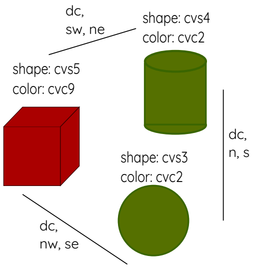
E_c_sphere



Interactive Concept Learning

Shared Environment

move green sphere
on(to) red cube

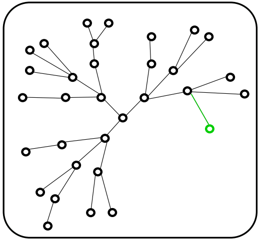


Lexical Processing

name -> sphere, cube, cylinder
prop -> red, green
obj-ref -> prop* name
abduce (action-word -> "move"
action-ref -> action-word np config np)

Working Memory

Scene Graph



Problem Space

comprehend

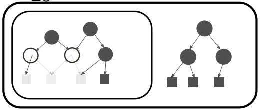
interpret
act-ref("move...")
act-word("move")

Concept Memory Ops

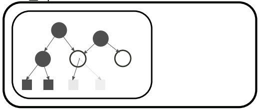
FAIL

Concept Memory

Ec_green

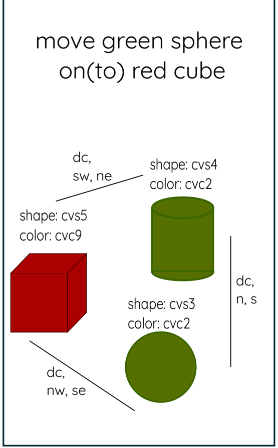


Ec_sphere



Interactive Concept Learning

Shared Environment

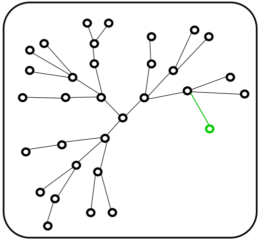


Lexical Processing

name -> sphere, cube, cylinder
prop -> red, green
obj-ref -> prop* name
abduce (action-word -> "move"
action-ref -> action-word np config np)

Working Memory

Scene Graph



Problem Space

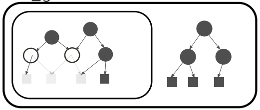
learn
from-demo-trace

Concept Memory Ops

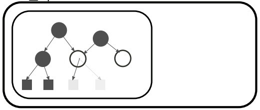
isa T_c,c_move
store (T_c,
E_c_sphere)

Concept Memory

E_c_green



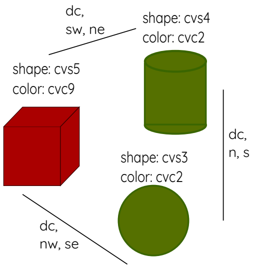
E_c_sphere

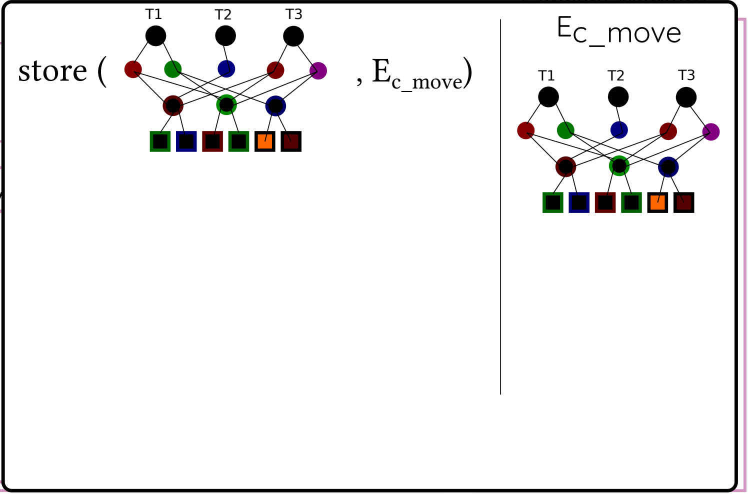


Interactive Concept Learning

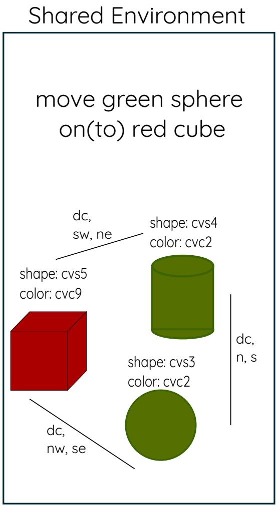
Shared Environment

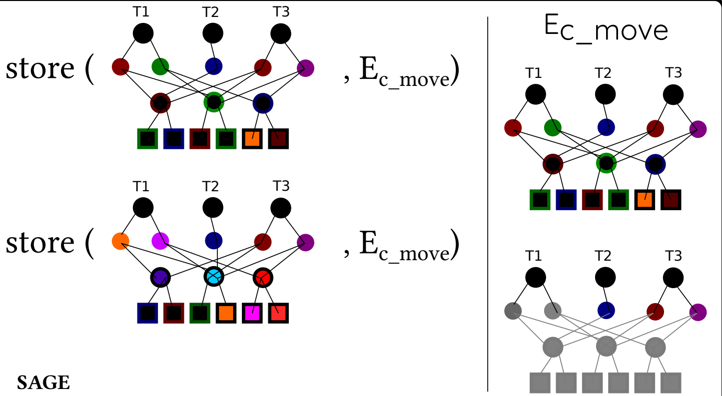
move green sphere on(to) red cube





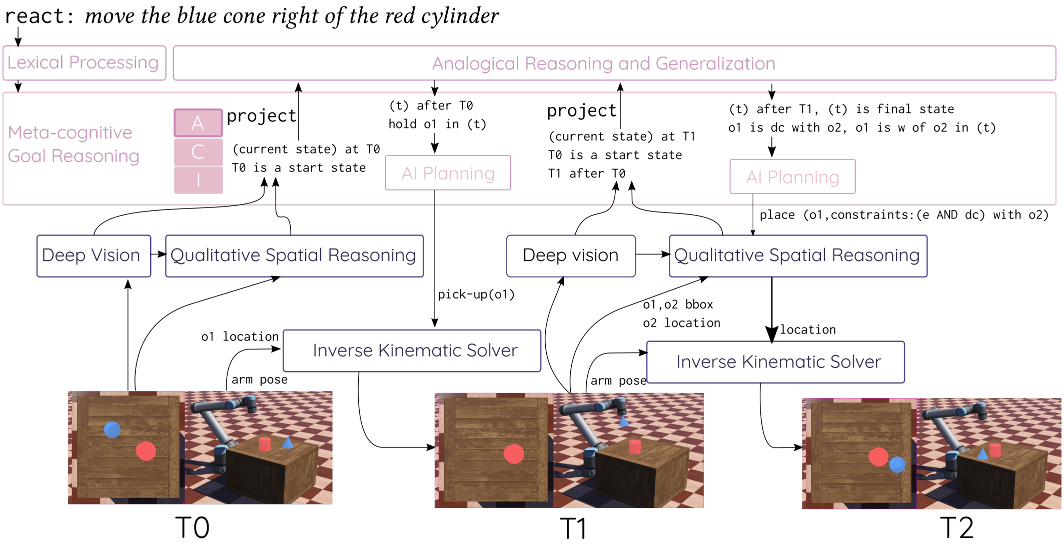
Interactive Concept Learning





SAGE
McLure, M. D., Friedman, S. E., & Forbus, K. D. (2015).
Extending Analogical Generalizationwith Near-Misses.
Twenty-Ninth AAAI Conference on Artificial Intelligence

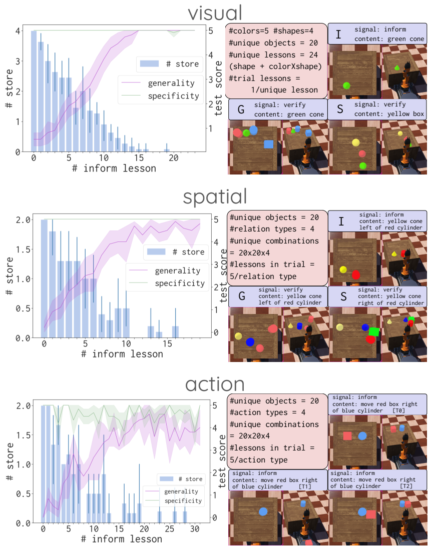
A Neuro-Cognitive Architecture



Evaluating Interactive Concept Learning

- Evaluating online, interactive, learning systems is a challenge
- New experimental scheme inspired by reinforcement learning
- A trial: N lessons of [inform, generality exam, specificity exam]
- Results:
    - Can learn a diverse types of concepts (and language)
    - Learns quickly, generalizes rapidly
    - Is smart about when to learn

Shiwali Mohan, Matt Klenk, Matthew Shreve, Kent Evans, Aaron Ang, John Maxwell. *Characterizing an Analogical Concept Memory for Newellian Cognitive Architectures.* Proceedings of the Eighth Annual Conference on Advances in Cognitive Systems (ACS). 2020



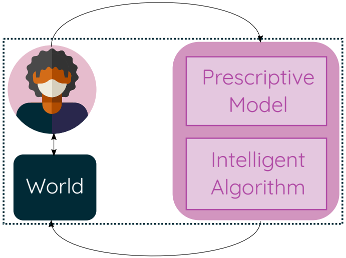
Demonstration

Publications

1. Preeti Ramaraj, Charlie Ortiz, Matt Klenk, **Shiwali Mohan**. *Unpacking Human Teachers' Intentions for Natural Interactive Task Learning*.[arXiv preprint arXiv:2102.06755]
2. **Shiwali Mohan**, Matt Klenk, Matthew Shreve, Kent Evans, Aaron Ang, John Maxwell. *Characterizing an Analogical Concept Memory for Architectures Implementing the Common Model of Cognition.* In the Proceedings of the Eigth Annual Conference on Advances in Cognitive Systems (ACS). 2020
3. John Laird and **Shiwali Mohan**. *Learning Fast and Slow: Levels of Learning in General Autonomous Intelligent Agents.* In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI/Blue Sky Award). 2018.

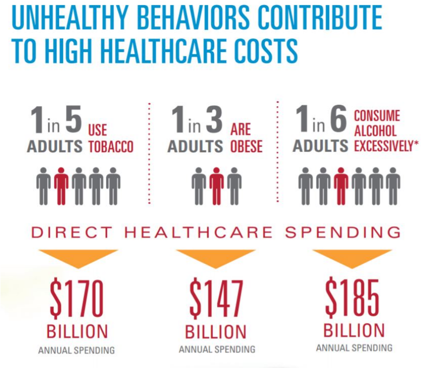
More at http://soargroup.github.io/rosie/

Health Behavior Change



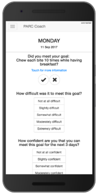
Healthcare Costs from Unhealthy Behaviors

- Behaviors rooted in sedentary lifestyles imposes major costs on healthcare
- NSF/NIH Smart and Connected Health: healthy behaviors is a critical technology and society problem
- Success depends on understanding how do humans learn new behaviors

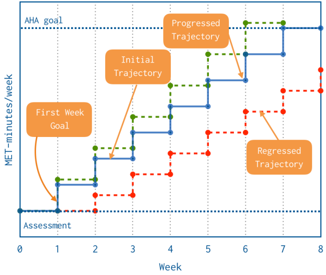


Interactive Coaching Agent for Walking

- Motivation
  - Individual coaching is highly effective
  - Mobile computing devices are pervasive
- Design goal
  - AHA recommended - 30 minutes of moderate activity 5 times a week
- Computational Approach
  - Prescriptive model from physical therapists guidelines (Bushman 2014)
  - Heuristics scheduling to design adaptive goal setting based on goal setting theory (Shilts and Townsend 2000)



User's Weekly Goal Trajectory



Shiwali Mohan, Anusha Venkatakrishnan, Michael Silva, and Peter Pirolli. *On Designing a Social Coach to Promote Regular Aerobic Exercise.* In the Proceedings of the 29th IAAI/AAAI. 2017

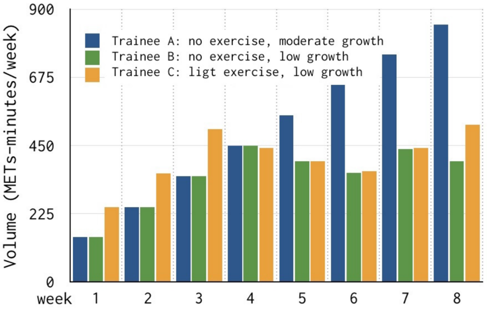
Evaluating an Interactive Coaching Agent

1. Is the AI system's behavior productive and beneficial?
2. Can humans trainees provide relevant information to the AI system?
    - Novel exercise goals interface
    - Interactive measurement of rate of perceived exertion (RPE), self-efficacy
3. Is the AI system effective in promoting regular walking?

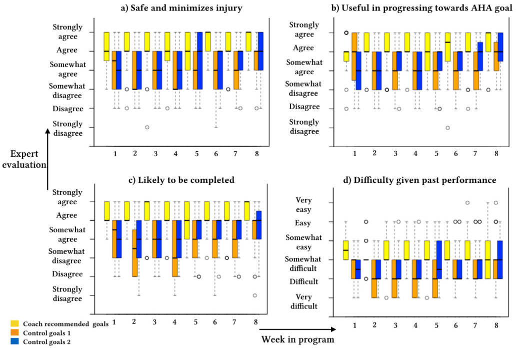
Shiwali Mohan, Anusha Venkatakrishnan, Andrea Hartzler. *Designing an AI Health Coach and Studying its Utility in Promoting Regular Aerobic Exercise.* In ACM Transactions on Interactive Intelligent Systems. 2020

Is It Productive?

1. Is it adaptive?



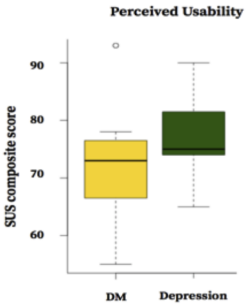
2. Do it experts agree with it?

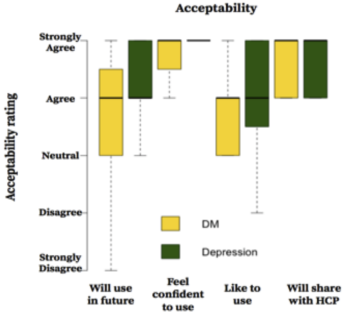


Shiwali Mohan, Anusha Venkatakrishnan, Michael Silva, and Peter Pirolli. *On Designing a Social Coach to Promote Regular Aerobic Exercise.* In the Proceedings of the 29th IAAI/AAAI. 2017

Do Human Trainees Understand It?

8 participants managing diabetes, 7 managing depression. Participants' perception of adaptive goal setting : could provide users with control (P9), help you take responsibility (P1), with more choice (P7), and allow you to set goals that you can strive for (P8).





Andrea Hartzler*, Anusha Venkatakrishnan*, **Shiwali Mohan**, Paula Lozano, James D Ralston, …,Les Nelson, Peter Pirolli. *Acceptability of a Team-Based Mobile Health Application for Lifestyle Self-Management in Individuals in Chronic Illnesses* In 38th Annual International Conference of the Engineering in Medicine and Biology Society. 2016

Is it Effective?

An ecological, observational study with 21 participants managing diabetes used our interactive coach for 6 weeks.

1. Increased exercise volume [✓]
2. Over-optimistic with self-assessment [X] [✓]
3. Personalized goals + collaborative selection led to more successful completion [✓]
4. Rate of perceived measurement scale provides informative feedback for adaptation [✓]

Independent Variables↓	(1) Goal Volume	(2) Performed Exercise	(3) Performed Exercise
Week	9.608^{*}	12.392^{*}	-0.487^{*}
	(5.166)	(12.202)	(12.007)
Goal Volume			0.618^{***}
			(0.119)
Mean Dependent Variable	601.098	392.250	392.250
	(23.138)	(24.830)	(24.830)
Random effect	✓	✓	✓
Marginal R^2	0.004	0.005	0.378
Conditional R^2	0.868	0.662	0.639

Table 2. Mixed-effect linear regression models for goal volume (column 1) and performed exercise volume (column 2). Volume is measured in MET-mins/week. The numbers in parentheses are standard errors. *** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$

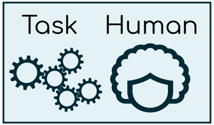
Publications

1. **Shiwali Mohan**. *Exploring the Role of Common Model of Cognition in Designing Adaptive Coaching Interactions for Health Behavior Change.* (in press). In ACM Transactions on Interactive Intelligent Systems. 2020.
2. **Shiwali Mohan**, Anusha Venkatakrishnan, Andrea Hartzler. *Designing an AI Health Coach and Studying its Utility in Promoting Regular Aerobic Exercise.* In ACM Transactions on Interactive Intelligent Systems. 2020
3. Aaron Springer, Anusha Venkatakrishnan, **Shiwali Mohan**, Les Nelson, Michael Silva, Peter Pirolli. *Leveraging Self-Affirmation to Increase mHealth Behavior Change.* In Journal of Medical Information Research. 2018
4. Peter Pirolli, **Shiwali Mohan**, Anusha Venkatakrishnan, Les Nelson, Michael Silva, Aaron Springer. *Implementation Intention and Reminder Effects on Behavior Change in a Mobile Health System: A Predictive Cognitive Model.* In Journal of Medical Information Research. 2017
5. **Shiwali Mohan**, Anusha Venkatakrishnan, Michael Silva, and Peter Pirolli. *On Designing a Social Coach to Promote Regular Aerobic Exercise.* In the Proceedings of the 29th IAAI/AAAI. 2017
6. Andrea Hartzler*, Anusha Venkatakrishnan*, **Shiwali Mohan**, Paula Lozano, James D Ralston, ...,Les Nelson, Peter Pirolli. *Acceptability of a Team-Based Mobile Health Application for Lifestyle Self-Management in Individuals in Chronic Illnesses* In 38th Annual International Conference of the Engineering in Medicine and Biology Society. 2016

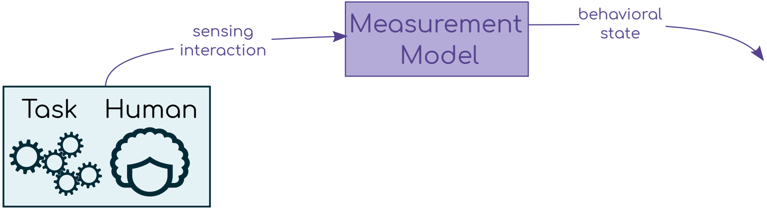
Humans of AI: A Research Agenda

Science of Collaborative Human-AI Systems

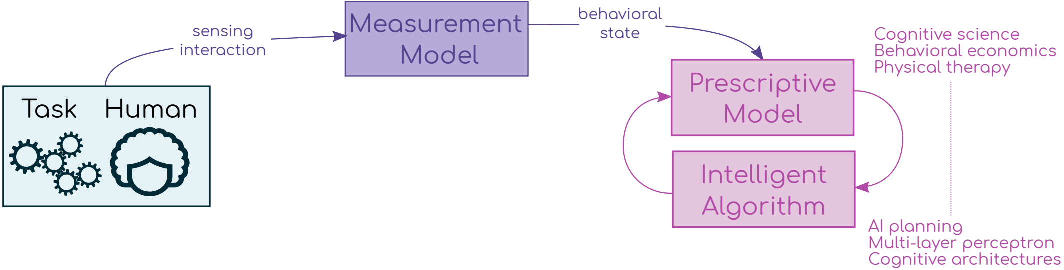
Science of Collaborative Human-AI Systems



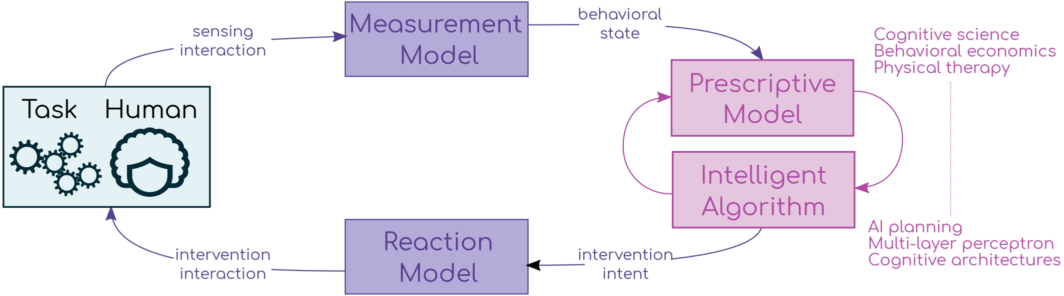
Science of Collaborative Human-AI Systems



Science of Collaborative Human-AI Systems

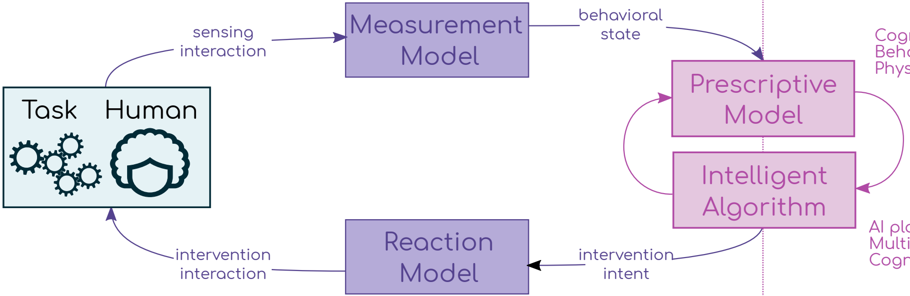


Science of Collaborative Human-AI Systems



Science of Collaborative Human-AI Systems

Q1: Intelligent system design

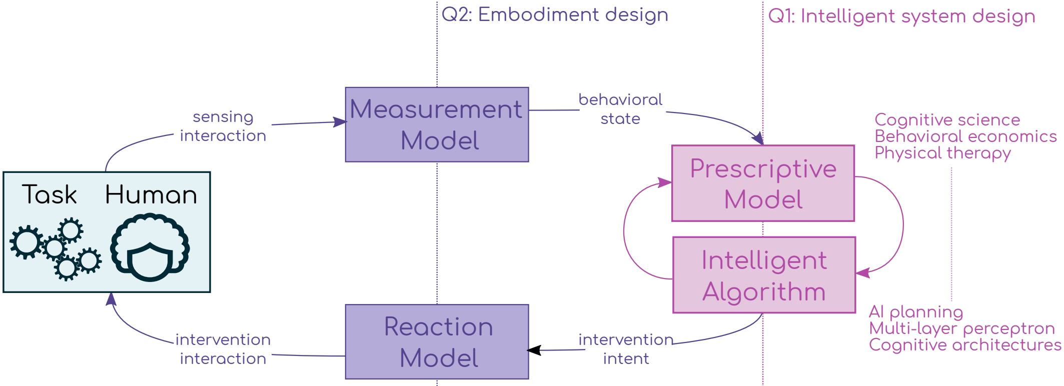


Cognitive science
Behavioral economics
Physical therapy

AI planning
Multi-layer perceptron
Cognitive architectures

Mohan et al. ACS 2020
Mohan et al. JAIR 2019
Laird and Mohan AAAI 2018 (Blue Sky Award)
Mohan et al. IAAI 2017
Mohan and Laird AAAI 2014
Mohan et al. ACS 2014
Mohan et al. ACS 2012

Science of Collaborative Human-AI Systems



Mohan ACM TiiS 2021 (forthcoming)
Ramaraj et al. IEEE RO-MAN (in prep)
Mohan et al. ACM TiiS 2020
Mohan et al. ACM IUI Workshops 2019
Mohan et al. AAAI FSS AI-HRI 2015
Mohan and Laird HRI Pioneer 2014

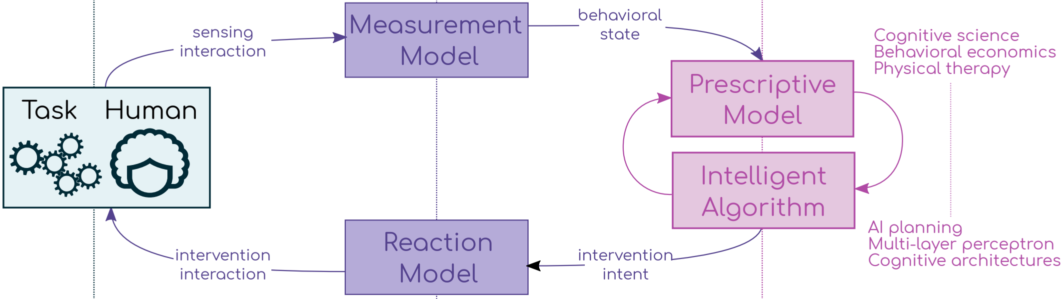
Mohan et al. ACS 2020
Mohan et al. JAIR 2019
Laird and Mohan AAAI 2018 (Blue Sky Award)
Mohan et al. IAAI 2017
Mohan and Laird AAAI 2014
Mohan et al. ACS 2014
Mohan et al. ACS 2012

Science of Collaborative Human-AI Systems

Q3: Efficacy and impact

Q2: Embodiment design

Q1: Intelligent system design

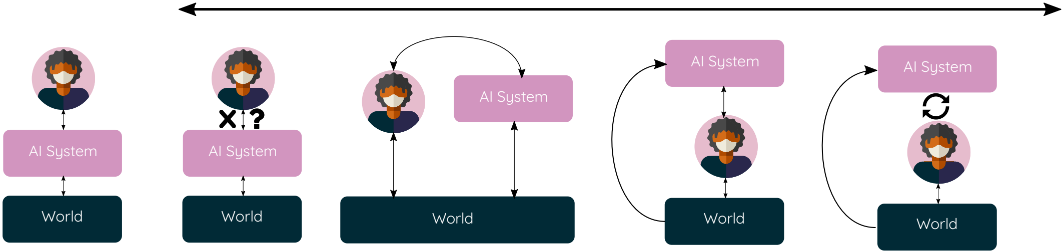


Mohan et al. AIES 2019
Springer et al. JMIR 2018
Pirolli, Mohan et al. JMIR 2017
Hartzler et al. EMBC 2016

Mohan ACM TiiS 2021 (forthcoming)
Ramaraj et al. IEEE RO-MAN (in prep)
Mohan et al. ACM TiiS 2020
Mohan et al. ACM IUI Workshops 2019
Mohan et al. AAAI FSS AI-HRI 2015
Mohan and Laird HRI Pioneer 2014

Mohan et al. ACS 2020
Mohan et al. JAIR 2019
Laird and Mohan AAAI 2018 (Blue Sky Award)
Mohan et al. IAAI 2017
Mohan and Laird AAAI 2014
Mohan et al. ACS 2014
Mohan et al. ACS 2012

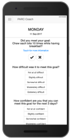
Impact of Collaborative Human-AI Systems



Interactive Task Learning
(DARPA, AFOSR, ONR)

Improving Triage
with Human-AI Collaboration
(USAID)

Sustainable Living
(DOE, ARPA-E)



Health Behavior
(NSF/NIH)

AR Trainer
(ONR, Xerox)

Thanks!

Colleagues:

Xerox PARC: Kalai Ramea, Matt Klenk, Victoria Bellotti, Charlie Ortiz, Matthew Shreve, Anusha Venkatakrishnan, Peter Pirolli, Aaron Ang, Kent Evans, Bob Price, Les Nelson.

University of Michigan: John Laird, Preeti Ramaraj, Aaron Mininger

SoarTech: James Kirk, Mike van Lent

Virginia Tech: Hesham Rakha, Hoda Eldardiry

University of Washington: Andrea Hartzler, Andrea Stocco

Funders:

DARPA

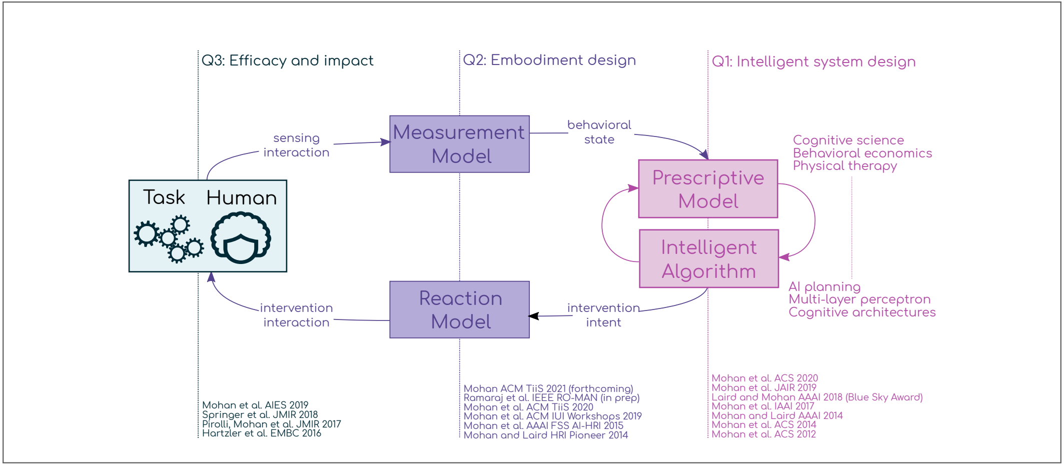
arpa·e CHANGING WHAT'S POSSIBLE

NSF

NIH National Institutes of Health

xerox

Collaborative Human-AI Systems



@shiwalimohan
shiwali.mohan@parc.com, shiwali.mohan@gmail.com