HuMaINs: Human-Machine Inference Networks

Aditya Vempaty, IBM T. J. Watson Research Center
Bhavya Kailkhura, Lawrence Livermore National Laboratories
Pramod K. Varshney, Syracuse University

HuMaINs

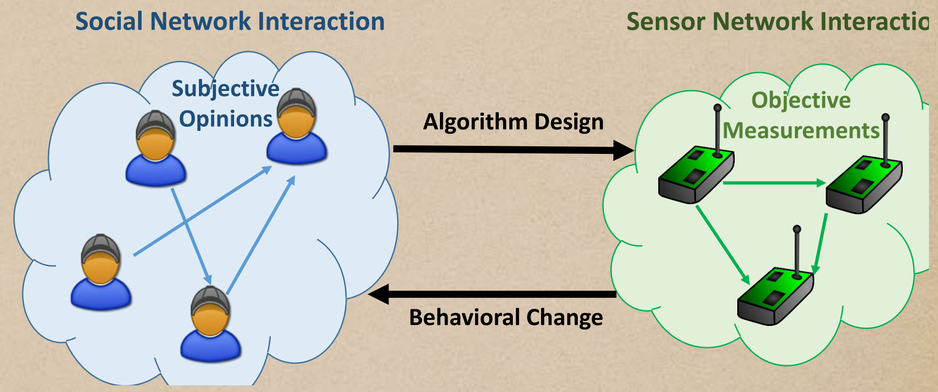
- Why?
- What?
- How?
- Where?

Motivation

- Problem solving has two steps: identify solution-space, and evaluate solutions
- Bounded-rational humans are cognitively limited and have limited time, information, and resources
  - Not fully aware of all possible choices
  - Sub-optimally evaluate consequences of choices
- Machines are rational
  - Stronger/larger memory for storing alternatives
  - Computational capability to accurately evaluate consequences
- Machines can aid humans in fast and accurate problem-solving ⟶ Human-Machine collaborative framework

Architecture

- HuMaIN = Social Network + Sensor/Machine Network
- Intelligent collaboration to exploit strengths of humans and machines
- Architecture governs interactions between agents in the networks
- Three architecture types:
  - Human controls autonomous system
  - Autonomous system monitor humans
  - Hybrid
- Relevant research problems:
  - Quantify human representation in decision-making tasks under uncertainty
  - Develop an estimator for model parameters
  - Provide common ontology for all agents to share information



Notional HuMaINs architecture

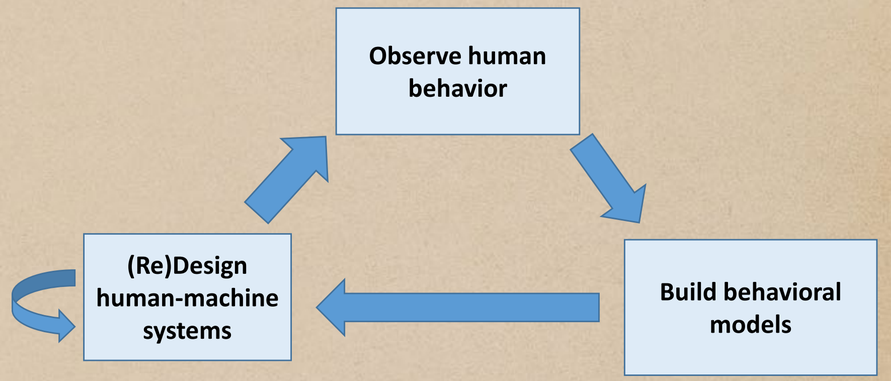
Recent Research

- W. Li, D. Sadigh, S. S. Sastry, and S. A. Seshia, "Synthesis for human-in-the-loop control systems," in TACAS2014 , Apr. 2014, pp. 470–484.
- P. B. Reverdy, "Human-inspired algorithms for search: A framework for human-machine multi-armed bandit problems," Ph.D. dissertation, Princeton University, 2014.
- P. B. Reverdy, V. Srivastava, and N. E. Leonard, "Modeling human decision making in generalized Gaussian multiarmed bandits," Proceedings of the IEEE, vol. 102, iss. 4, pp. 544-571, 2014.
- P. Reverdy and N. E. Leonard, "Satisficing in Gaussian bandit problems," in IEEE 53rd Annual Conference on Decision and Control (CDC), 2014, pp. 5718-5723.

Human-in-the-loop Cyber-Physical System (HiLCPS), Control Systems, Human-Computer Interface

Algorithms

- Development of new algorithms dealing with human behavioral data
- Collaborative integration of human expertise with automated processing
- Timely decision making and intervention
- Support but not supplant humans
- Relevant research problems:
  - Develop mathematical models of human decision making
  - Design robust fusion algorithms that handle potentially unreliable data



General approach for the design and analysis of HuMaINs

Recent Research

- S. Narayanan and P. G. Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language," Proc. IEEE , vol. 101, no. 5, pp. 1203–1233, Feb. 2013.
- A. Vempaty, L. R. Varshney, G. J. Koop, A. H. Criss, and P. K. Varshney, "Experiments and Models for Decision Fusion by Humans in Inference Networks," IEEE Trans. Signal Process., 2018.
- A. Vempaty, L. R. Varshney, and P. K. Varshney, "Reliable crowdsourcing for multi-class labeling using coding theory," IEEE J. Sel. Topics Signal Process. , vol. 8, no. 4, pp. 667–679, Aug. 2014.

Behavioral Signal Processing, Cognitive Psychology, Mathematical Modeling, Robust Data Fusion

Applications

- Major driver of HuMaINs research
- Specific application nuances drive relevant architectural and algorithmic research
- Timely and important application areas:
  - Education
  - Autonomous Vehicles
  - Healthcare
  - Scientific Discovery

Education: Intelligent Tutoring Systems

- Provide immediate and customized instruction/ feedback to learners with teacher intervention
- Complement a human teacher
- Ensure personalized and adaptive learning at scale

ITS	HuMaIN
Domain model	Task
Student model	Human
Tutoring model	Machine/Robust fusion
User interface	Architecture

Autonomous Vehicles

- Sense and interact with physical world via sensors and actuators
- Fully autonomous makes them vulnerable to cyber-attacks
- Semi-autonomous vehicles: human-in-the-loop for safe and intelligent operation
- Require joint environment-driver state sensing, inference, and shared control and new metrics to characterize safety
- Communication among multiple self-driving cars can enable collective intelligence
- Requires design of robust communication protocols.

Healthcare

- Fully automated inference holds enormous potential to increase quality, efficacy, and efficiency of treatment and care
- Certain tasks with a small number of data sets or rare events, ML-approaches suffer from insufficient training samples
- Health decisions have serious consequences and necessitate human experts' domain knowledge
- Growing trend of litigation requires doctor-in-the-loop
- Need interpretable models that are transparent and earn experts' trust before being adopted in their workflow

Scientific Discovery

- Scientific research spans from screening of novel materials in material science, to tracking of extreme weather phenomena in climate science.
- Currently, machine role is limited to solving a well-defined task where the data and techniques are given by the scientists.
- Limits ability to tackle problems where data or task complexity itself challenge human capabilities to make discoveries.
- Advancing machine learning techniques for independent inquiry, proactive learning, and deliberative reasoning in the presence of hypotheses, domain knowledge, and insights provided by the scientists.

Session Details

- ARCHITECTURE:
  - P. Reverdy and V. Srivastava, “Multi-armed bandits for human-machine decision making”
- ALGORITHMS:
  - D. Seo, R. Raman, and L. R. Varshney, “Probability reweighting in social learning: Optimality and suboptimality”
  - J. Rhim and V. Goyal, “Team decision making with social learning: Human subject experiments”
  - S. Mourad and A. Tewfik, “Machine assisted human decision making”
- APPLICATIONS:
  - R. Kokku, S. Sundararajan, P. Dey, R. Sindhgatta, S. Nitta, and B. Sengupta, “Augmenting classrooms with AI for personalized education”