Human-CPS Cooperation

- ➢ Evolution of Human-Machine Interaction
  - ➢ Phase 1: Human-Physical Systems (HPS)
    - ➢ Humans in full control (perception, analysis, decision)
    - ➢ Limitation: Productivity & Quality ceiling
  - ➢ Phase 2: Human-Cyber-Physical Systems (HCPS)
    - ➢ Digital tech aids in data acquisition, processing, decision-making
    - ➢ Benefit: Enhanced efficiency
  - ➢ Phase 3: AI-Integrated HCPS (Present/Future)
    - ➢ AI enables automation of complex tasks, learning from experience, and true cooperation.

Liang Cheng, Ph.D., EECS Professor and Department Chair, https://liangcheng.info/

THE UNIVERSITY OF TOLEDO

Flipping the Perspective

➢ It's not just about humans trusting AI

➢ How can an AI-powered machine assess a human, build appropriate trust, and interact optimally to achieve a shared goal?

   ➢ Know-How-to-Cooperate (KHC)

      ➢ Machines must assess teammates (humans)

      ➢ They must identify and resolve discrepancies

      ➢ They must manage responsibilities and activities

Framework

➤ Flipping the Perspective

  ➤ It's not just about human trusting AI

Liang Cheng, Ph.D., EECS Professor and Department Chair, https://liangcheng.info/
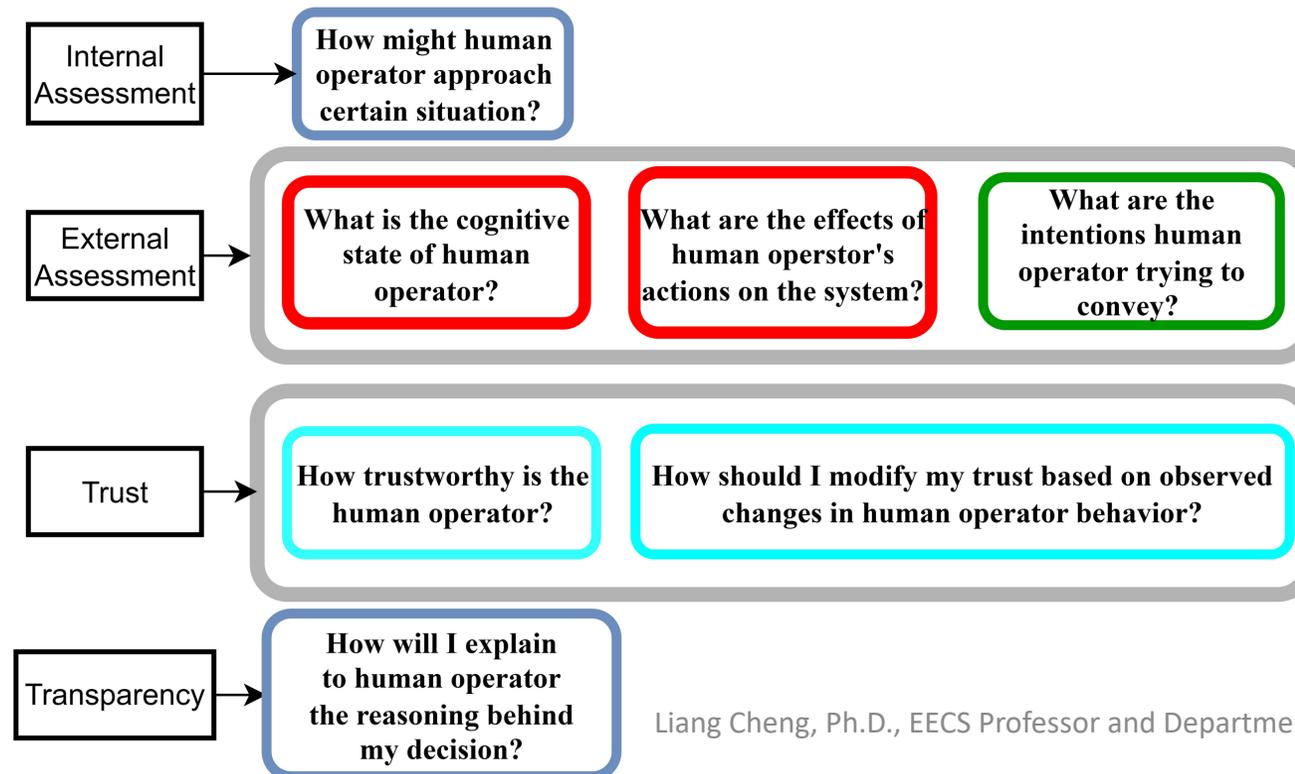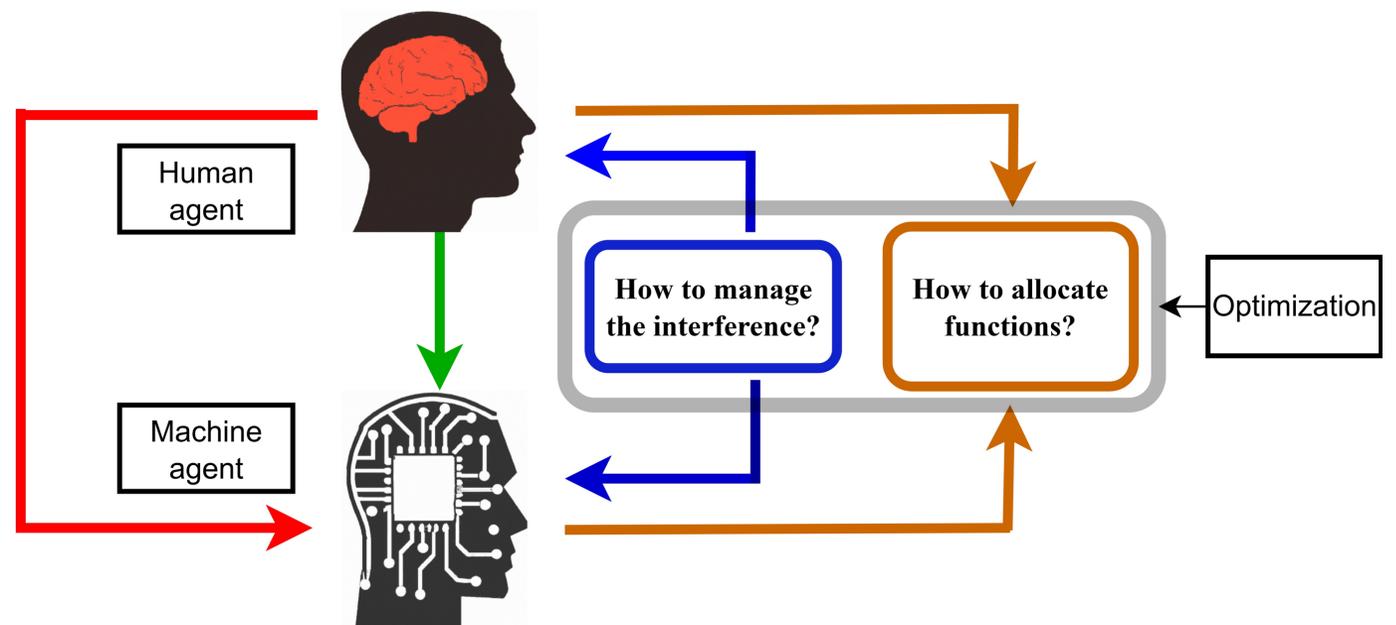
How Machines Assess Humans

➢ External Assessment (Observing the "What")
  ➢ Using sensors to interpret observable cues
  ➢ Cognitive States: Emotion, attention, cognitive load (via facial recognition, eye-tracking, biometrics)
  ➢ Actions & Intentions: Gesture recognition, predicting next move.
  ➢ Communication: Natural Language Processing (NLP), speech recognition

➢ Internal Assessment (Simulating the "Why")
  ➢ Using cognitive models (like ACT-R, Soar) to simulate human problem-solving and decision-making
  ➢ Creating a "digital representation" of the human's thought process

Assessment in Action

➤ Manufacturing applications
  ➤ Gesture and Intent Recognition
    ➤ AI interprets gestures (e.g. hand signals) for seamless collaboration
    ➤ AI processes sensor data (e.g. haptic signals) and sync's w/ intended motion
  ➤ Emotion Detection
    ➤ Using Convolutional Neural Networks (CNNs) to analyze a worker's facial expression, detecting fatigue or frustration to prevent errors
    ➤ A Smart Safety Helmet tracks head movement and brain activity to monitor for fatigue and ensure worker safety, e.g. using COGNET/iGEN model
  ➤ NLP, LLMs, and Multimodal Prediction
    ➤ Workers use natural voice commands for complex, hands-free operations in noisy industrial environments
    ➤ Workers state a high-level goal, and AI (e.g. Fusion-based Spiking Neural Networks) generates a collaborative plan

Crucial Role of Machine Trust

➢ Core Principle: A machine's trust in a human should mirror the human's actual trustworthiness.

  ➢ Dimensions of Trustworthiness (ABI Model):

    ➢ Ability: Is the human skilled and capable?

    ➢ Benevolence: Does the human have good intentions for the team?

    ➢ Integrity: Is the human reliable and consistent?

  ➢ The Goal: Trust Calibration

    ➢ Avoiding over-trust (e.g., relying on a fatigued human operator)

    ➢ Avoiding under-trust (e.g., ignoring an expert human's input)

AI's Role in Calibrating Machine Trust

- ➤ Pattern Recognition & Prediction
  - ➤ AI analyzes patterns in human behavior, decision consistency, and performance to assess Ability & Integrity and predicts errors
- ➤ Continuous Learning
  - ➤ AI (RL) agent learns and adapts its trust model over time based on the success of interactions
- ➤ Personalization
  - ➤ AI builds individualized models of human partners, tailoring trust levels to a specific person's ABI
- ➤ Trust Calibration
  - ➤ AI adjusts its own autonomy level—deciding when to rely on the human and when to take the lead

Optimizing Human-Machine Cooperation

➢ Intelligent Function Allocation ("Who Does What?")

  ➢ Manual/Rule-Based: Simple, but rigid.

  ➢ AI-Driven Approaches

    ➢ Optimization-Based: AI calculates the most efficient division of labor.

    ➢ Adaptive: AI re-assigns tasks in real-time based on workload and performance.

    ➢ Mixed-Initiative: Either human or AI can initiate a change in task allocation, allowing for maximum flexibility.

➢ Interference Management ("How Do We Stay in Sync?")

  ➢ AI can act as a Mediator to resolve goal conflicts, a Monitor to provide feedback, and a tool for Bias Detection to ensure fairness.

Key Challenges

➤ Assessment Challenges
  ➤ Robustness: real-world noise, poor lighting, occlusions etc.
  ➤ Privacy: How do we monitor workers ethically?

➤ Trust & Transparency Challenges
  ➤ Metrics: How do we measure "machine trust"?
  ➤ Explainability (XAI): Explaining an AI's decision can be complex and may overload the human collaborators.

➤ Interaction Challenges
  ➤ Adaptability vs. Stability: How can a system be flexible without becoming unpredictable?
  ➤ Edge Cases: How to handle unexpected scenarios?

Key Takeaways

- ➤ A Paradigm Shift
  - ➤ True human-machine cooperation requires a focus on the machine's perspective, not just the human's.

- ➤ Beyond Execution
  - ➤ Machines must evolve from simple task-doers to active partners that can assess, trust, and adapt.

- ➤ Trust is Dynamic
  - ➤ The ability to dynamically calibrate trust in a human partner is the cornerstone of safe and effective collaboration.