Human-Computer Interaction

Human-AI Interaction

Professor Bilge Mutlu

Questions

To ask questions during class:

» Go to slido.com and use code **#2938904** or direct link or scan QR code

» Anonymous

» I will monitor during class

Today's Agenda

- Topic overview: *Human-AI Interaction*
- Discussion

What are we talking about when we are talking about "AI?"

AI-infused systems: ... systems that have features harnessing AI capabilities that are directly exposed to the end user.[1]

[1] Amershi et al. (2019). Guidelines for human-AI interaction. *CHI 2019.*

Recap: Week 02: History of HCI

1960s: *Man-Computer Symbiosis*, 1960, Joseph Licklider, ARPA[2] [3]

"Men will set the goals, formulate the hypotheses, determine the criteria, and perform the evaluations. Computing machines will do the routinizable work that must be done to prepare the way for insights and decisions in technical and scientific thinking."

[2] Licklider (1960). Man-computer symbiosis. *IRE transactions on human factors in electronics*, (1), 4-11.

[3] Image source

But how can computers do this?

Attempts at Establishing Human-Computer Symbiosis

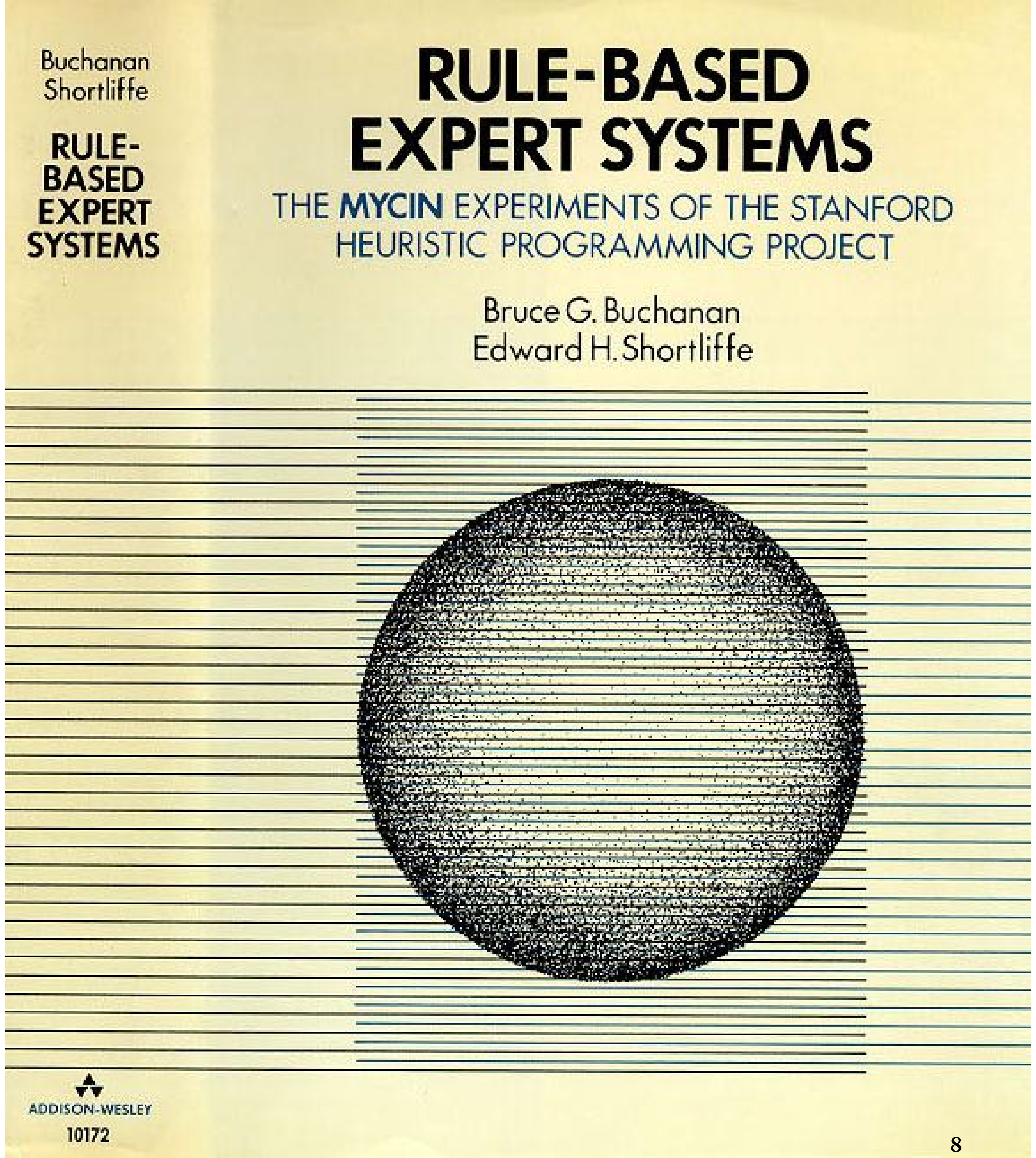
1960s-1970s: The rise of **expert systems**

Definition: Expert system, a computer program that uses artificial-intelligence methods to solve problems within a specialized domain that ordinarily requires human expertise. The first expert system was developed in 1965 by Edward Feigenbaum and Joshua Lederberg at Stanford. [4]

[4] Britannica: expert system

The Problem with Expert Systems[12]

Production rules are a popular representation for encoding heuristic knowledge in programs for scientific and medical problem solving. However, experience with one of these programs, MYCIN, indicates that the representation has serious limitations: people other than the original rule authors find it difficult to modify the rule set, and the rules are unsuitable for use in other settings, such as for application to teaching. These problems are rooted in fundamental limitations in MYCIN's original rule representation: the view thatexpert knowledge can be encoded as a uniform, weakly-structured set of If/then associations is found to be wanting.[5]



[12] Image Source

[5] Clancey (1983). The epistemology of a rule-based expert system—a framework for explanation.
*Artificial intelligence
*, 20(3), 215-251.

Overcoming these Limitations: The Neural Net Revolution

- **Milestone 1:** Neural Nets — initial inspiration from the human brain introduced in 1940s;[6] practical applications started in late 1980s[7]
- **Milestone 2:** Deep Learning — GPUs in the late 1990s, large datasets (e.g., ImageNet), unsupervised learning, GANs in mid-2010s[8]
- **Milestone 3:** Transformers — transformer as a specific type of NN in 2017;[9] BERT in 2018;[10] commercial development of GPTs (2018–2023)[11]

---

[6] McCulloch & Pitts (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*, 115-133.

[7] LeCun et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, *1*(4), 541-551.

[8] A brief history of deep learning

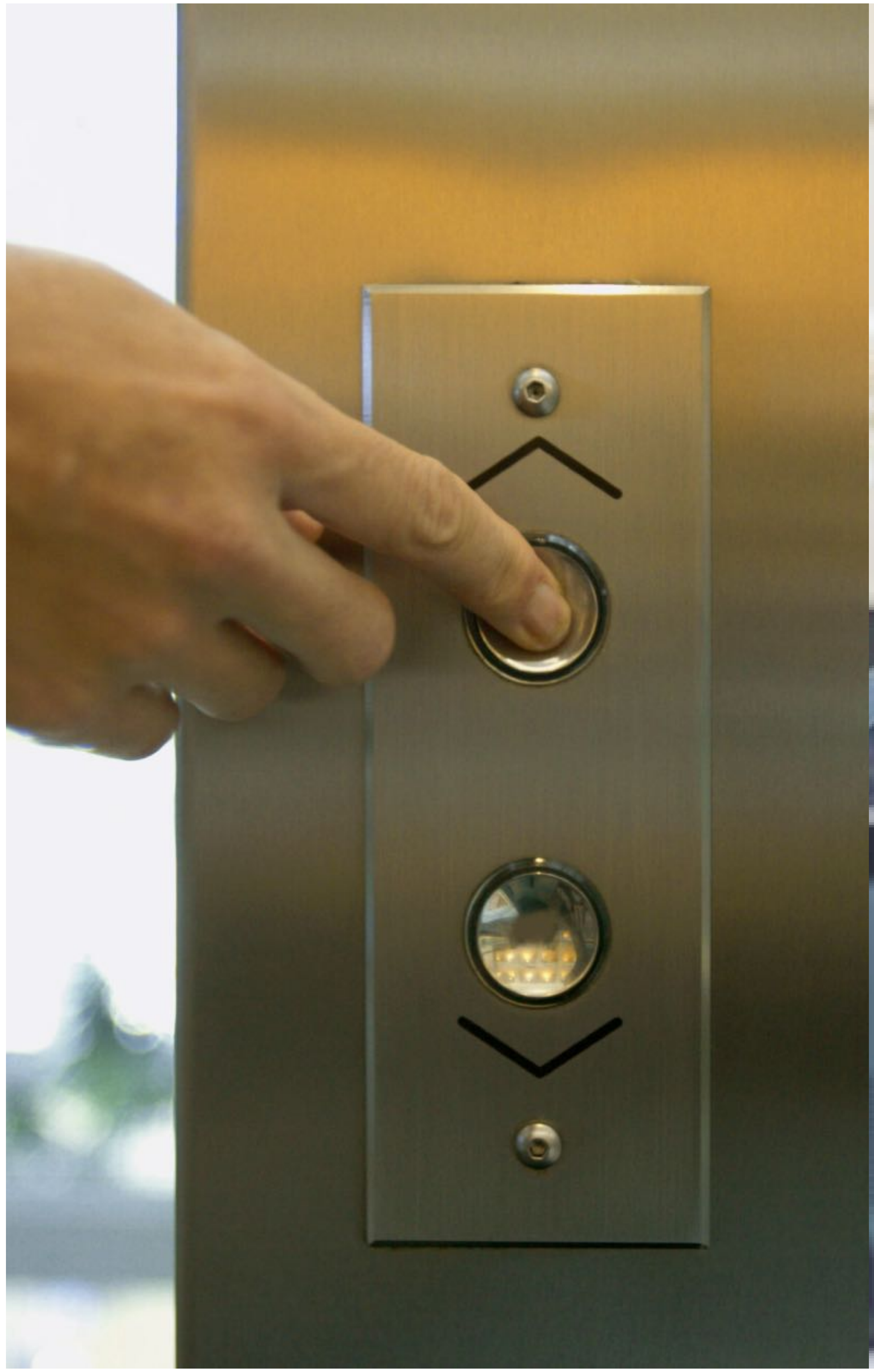
[9] Vaswani et al. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

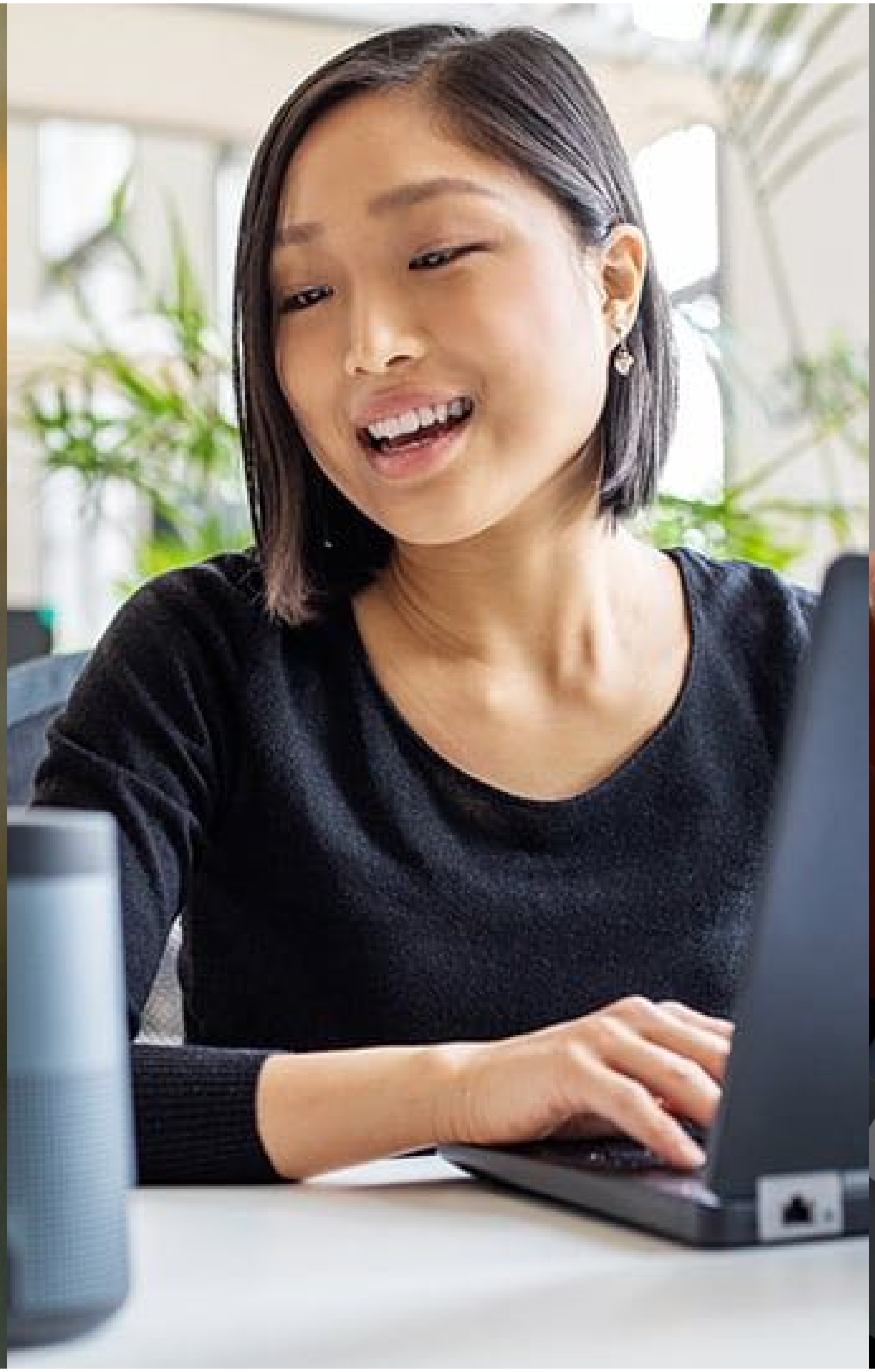
[10] Devlin et al. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

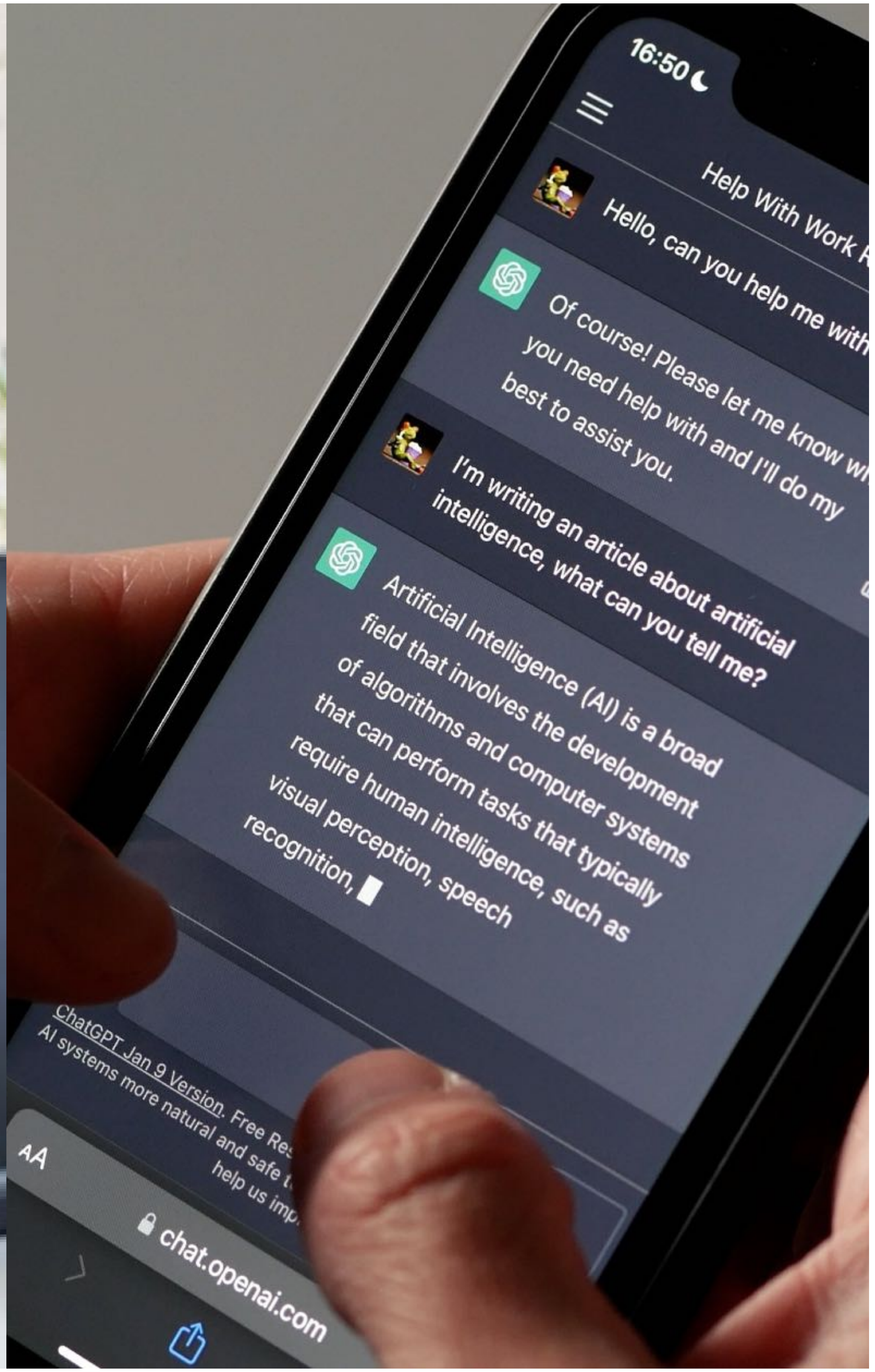
[11] GPT-4: History, Freatures, and Predictions

So we now have "machines" that can support human-machine "symbiosis." What do we do now?[13]

[13] Images on next slide: Left | Center | Right






16:50
Help With Work
Hello, can you help me with
Of course! Please let me know w
you need help with and I'll do my
best to assist you.
I'm writing an article about artificial
intelligence, what can you tell me?
Artificial Intelligence (AI) is a broad
field that involves the development
of algorithms and computer systems
that can perform tasks that typically
require human intelligence, such as
visual perception, speech
recognition,
ChatGPT Jan 9 Version
chat.openai.com

These interactions represent **intermittent** interaction scenarios, in which there is a clear line between the human initiator of an interaction and an almost immediate system response.[14] [15]

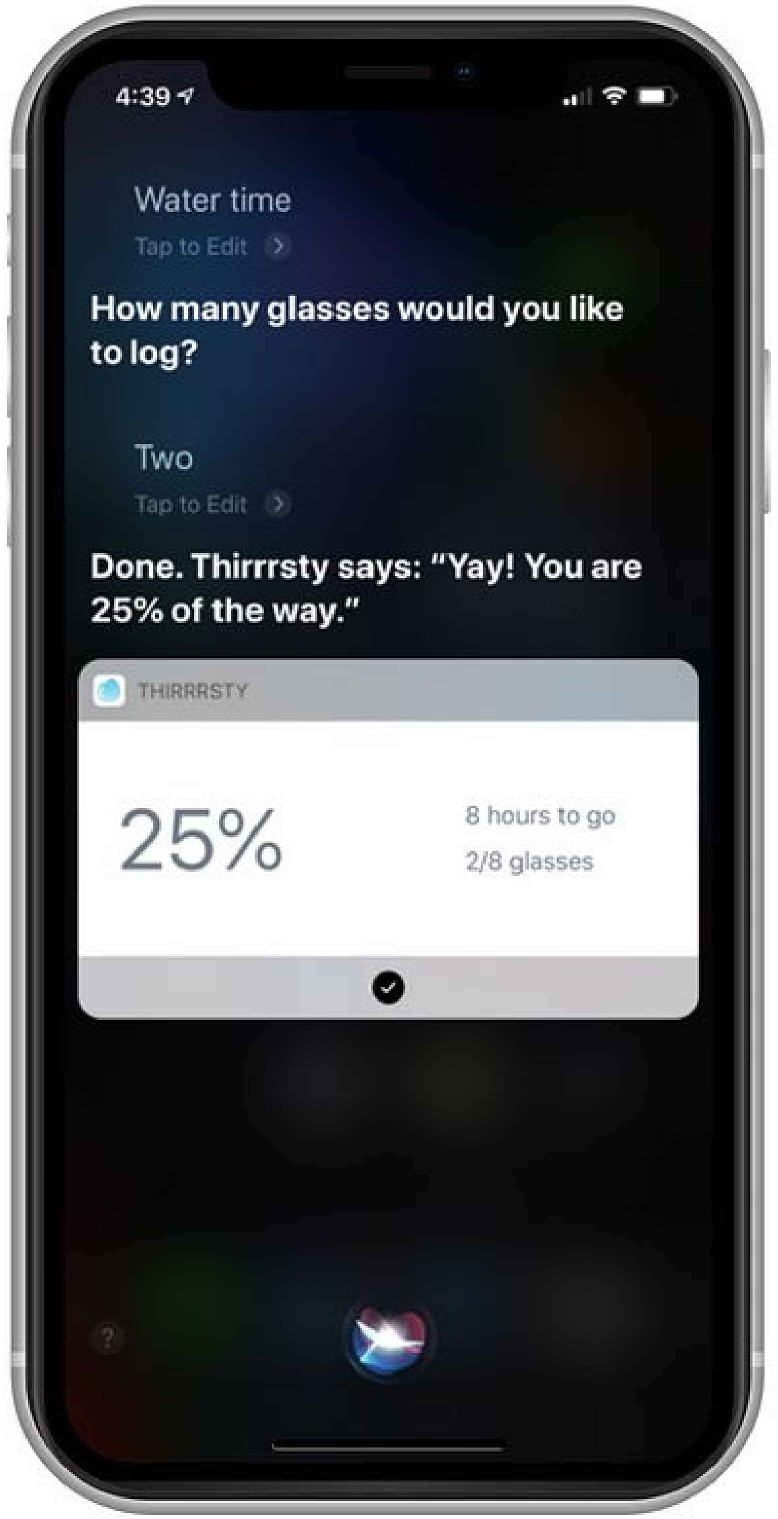
AI enables other forms of interaction:

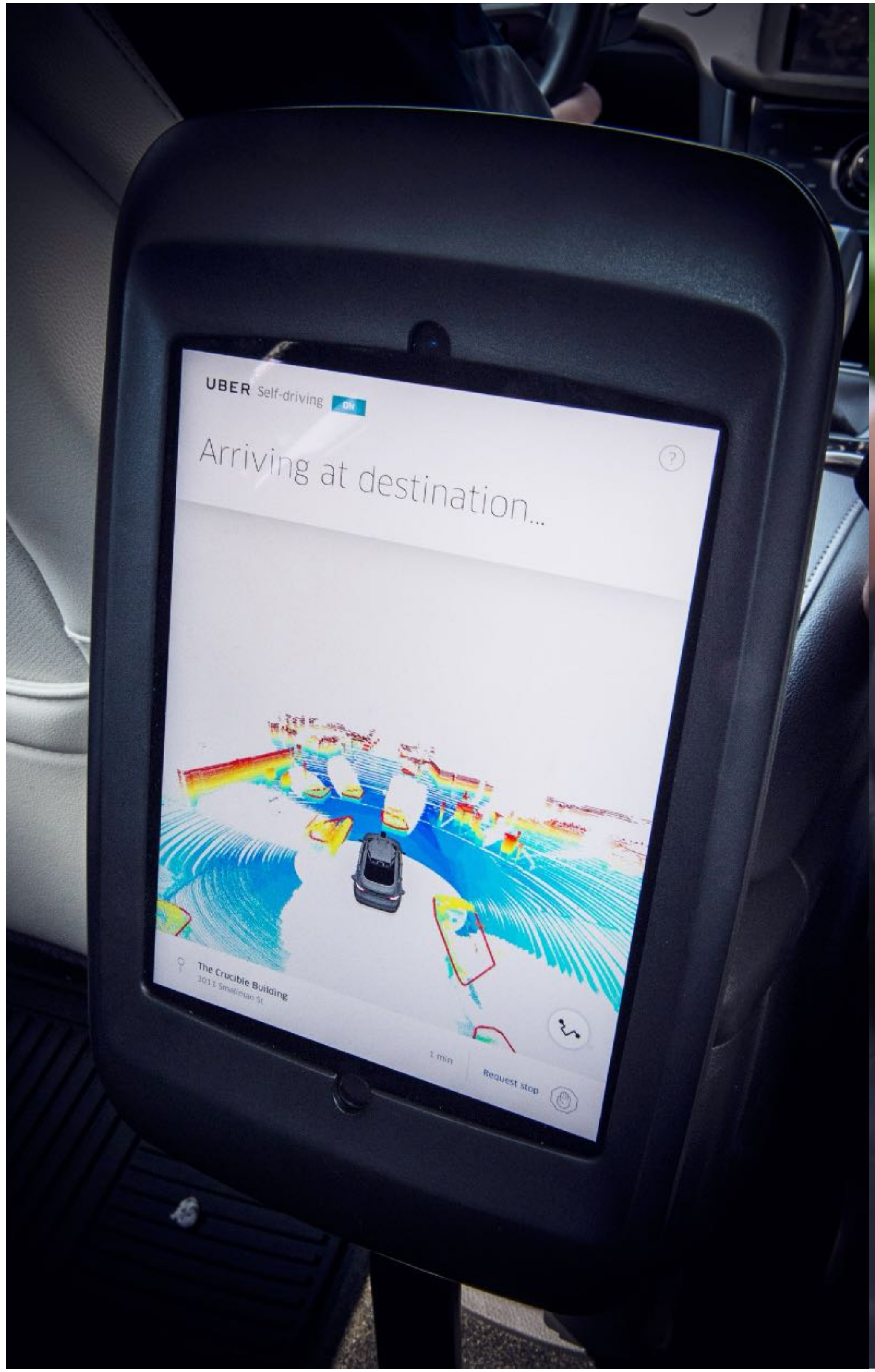
- **Intermittent human-AI interaction** describes systems that follow the notion of interaction between user and system as a turn-taking process.
- **Continuous human-AI interaction** describes systems that "listen" to a stream of uninterrupted user input rather than individual instructions and can respond to this input throughout the duration of the interaction.
- **Proactive human-AI interaction** describes systems that do not wait for user input but instead actively initiate and complete tasks based on, for example, sensor readings.

---

[14] Van Berkel et al. (2021). Human-AI interaction: intermittent, continuous, and proactive. *Interactions*, *28*(6), 67-71.

[15] Images on next slide: Left | Center | Right


4:39
Water time
Tap to Edit
How many glasses would you like to log?
Two
Tap to Edit
Done. Thirrrsty says: "Yay! You are 25% of the way."
THIRRRSTY
25%
8 hours to go
2/8 glasses


UBER Self-driving
ON
Arriving at destination...


Relax Reminder
Higher than normal stress detected. Start a breathing exercise to relax?
GARMIN

	Intermittent[14]	Continuous	Proactive
Initiator of interaction	User	User/AI	AI
Trigger for AI input	Explicit and predefined cue. For example, "Hey, Google."	Implicit, as part of constant input. For example, a video feed on which each frame is analyzed.	Change in system state matching a predefined condition. For example, movement captured through a sensor.
Resulting AI response	System response to user in predefined format. For example, "Temperature in Berlin is 18 degrees."	System response to user alongside the continuing user input. For example, a visual marker is introduced alongside a video feed.	System response to change in system state by changing the system. For example, turning on the lights and music when user comes home.
User response	User interprets the presented information. May be followed up by a new information request.	User decides to either ignore or react to the presented information, possibly adjusting the presented input stream.	User response not required; user might not be aware of AI intervention.

[14] Van Berkel et al. (2021). Human-AI interaction: intermittent, continuous, and proactive. *Interactions*, *28*(6), 67–71.

Wait, have we achieved "human-machine symbiosis"?

Current Approaches to Human-Machine Symbiosis

Mixed-initiative interaction: Mixed-initiative interaction refers to a flexible interaction strategy in which each agent (human or computer) contributes what it is best suited at the most appropriate time.[16]

Human-AI collaboration: … collaboration [is] an evolving, interactive process whereby two or more parties actively and reciprocally engage in joint activities aimed at achieving one or more shared goals… Human-AI collaboration then refers to the collaboration between single or multiple humans and AI systems.[17]

Human-machine teaming: …a purposeful combination of human and cyber-physical elements that collaboratively pursue goals that are unachievable by either individually.[18]

---

[16] Allen et al. (1999). Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications.*

[17] Lai et al. (2021). Human-AI collaboration in healthcare: A review and research agenda. *HICSS 2021.*

[18] Madni & Madni (2018). Architectural framework for exploring adaptive human-machine teaming options in simulated dynamic environments. *Systems.*

Definition (cont'd)

autonomous /ɔˈtɑnəməs/ *adj.* Of a process: self-initiating or self-sustaining; occurring or developing spontaneously.

automated /ˈɔdəˌmeɪdɨd/ *adj.* Converted so as to operate automatically; that has been automated; automatic.

initiative /ɨˈnɪʃədɪv/ *n.* That which initiates, begins, or originates; the first step in some process or enterprise; hence the act, or action, of initiating or taking the first step or lead; beginning, commencement, origination.

Bringing These Ideas Together

- Users directly interact with AI-driven features (i.e., AI-driven features are not only doing automated work in the background withour the user's knowledge)
- The user's goal is modeled into the system
- The user and the system know each other's responsibilities (at any given time)
- Who does what is based on capabilities

Human-autonomy[19] collaboration: Human collaboration with autonomous features of an interactive system where the human and the system work toward a shared goal and perform tasks that best fit their capabilities.

[19] Substitute with AI, AI-based system, AI-based capabilities, autonomous capabilities depending on context

An Example: Human-Robot Teaming in Assembly[20]

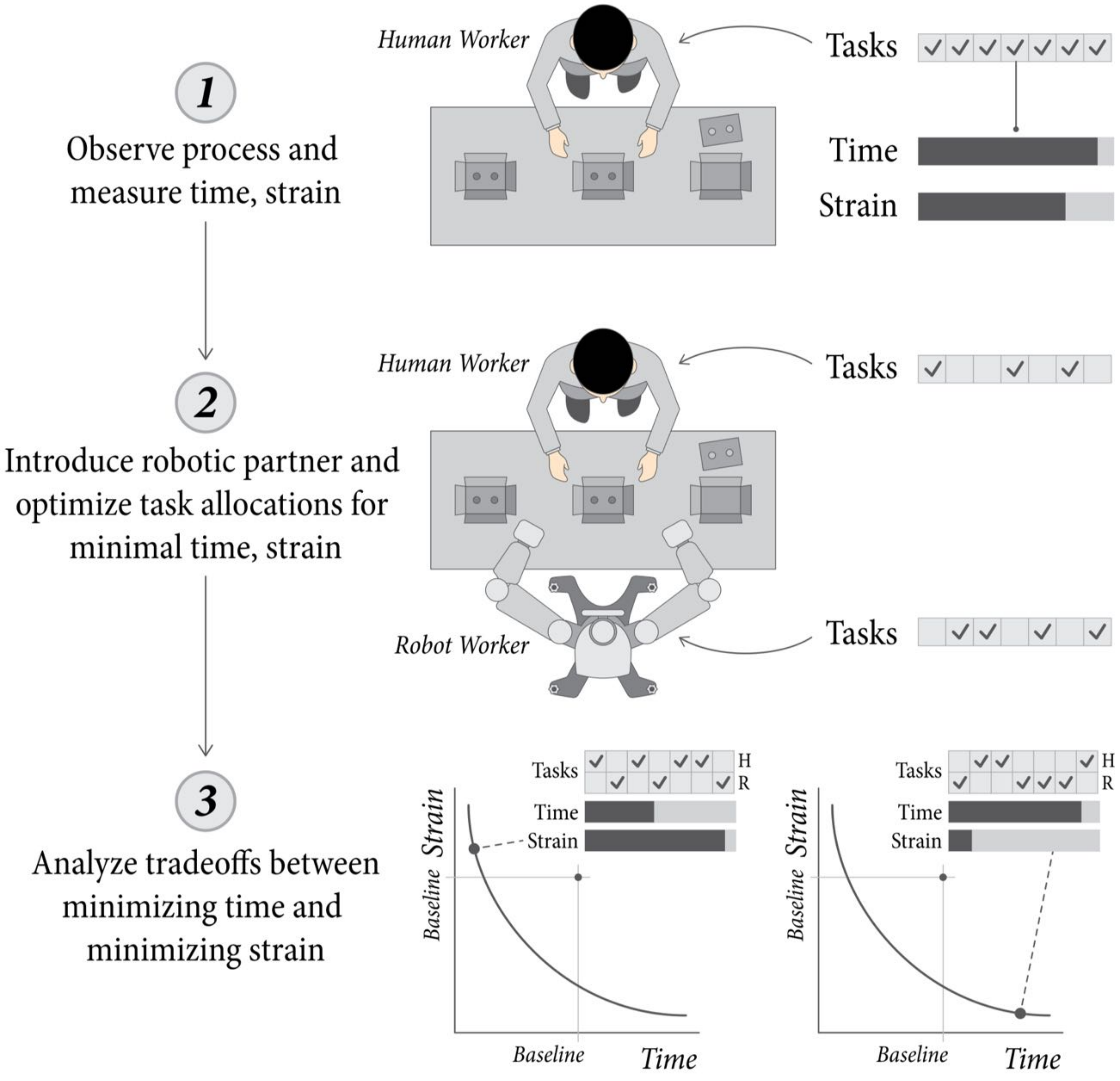
[20] ErgoPack via YouTube

How do know who (human/AI) does what best?

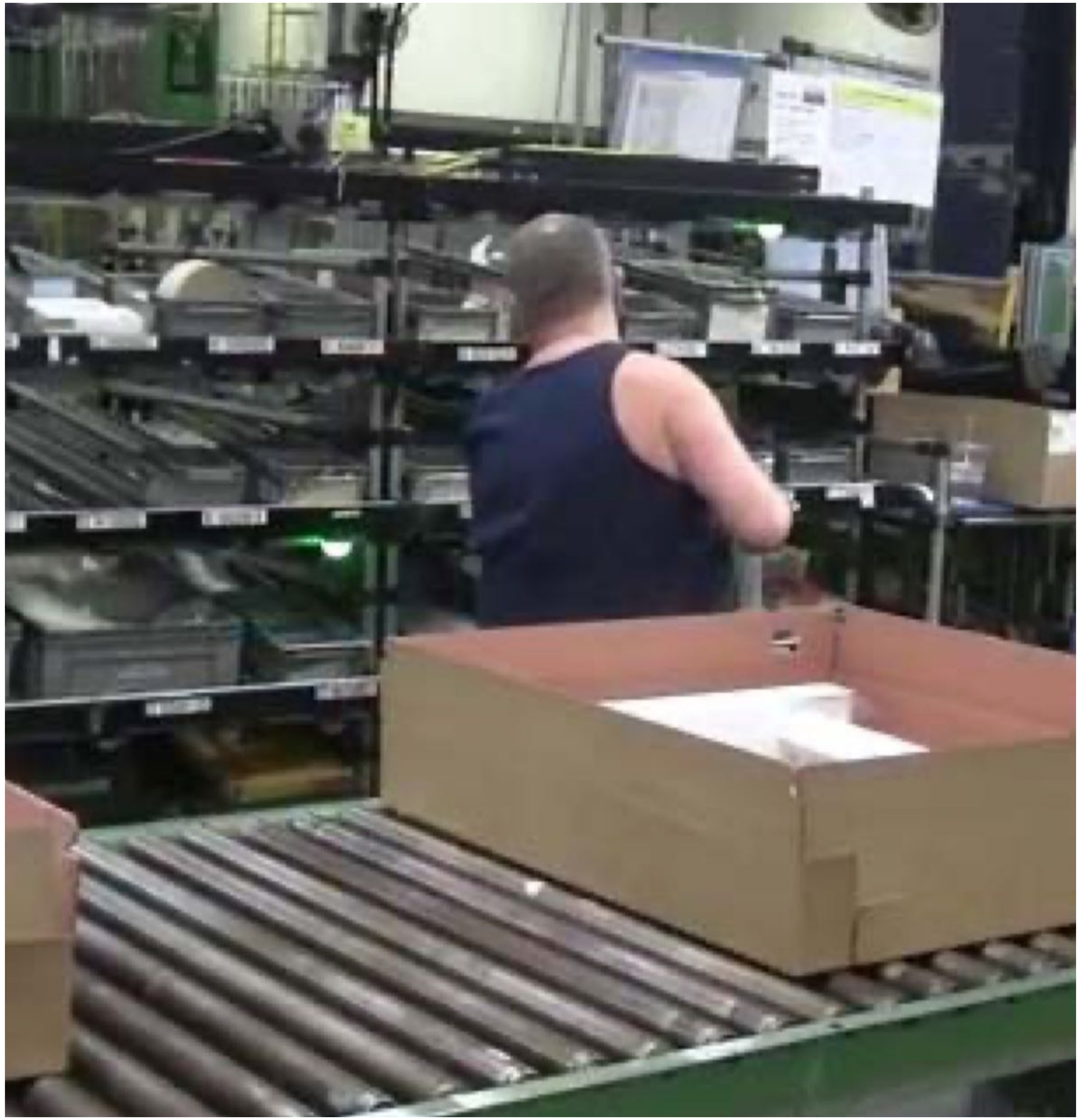
An Example Approach

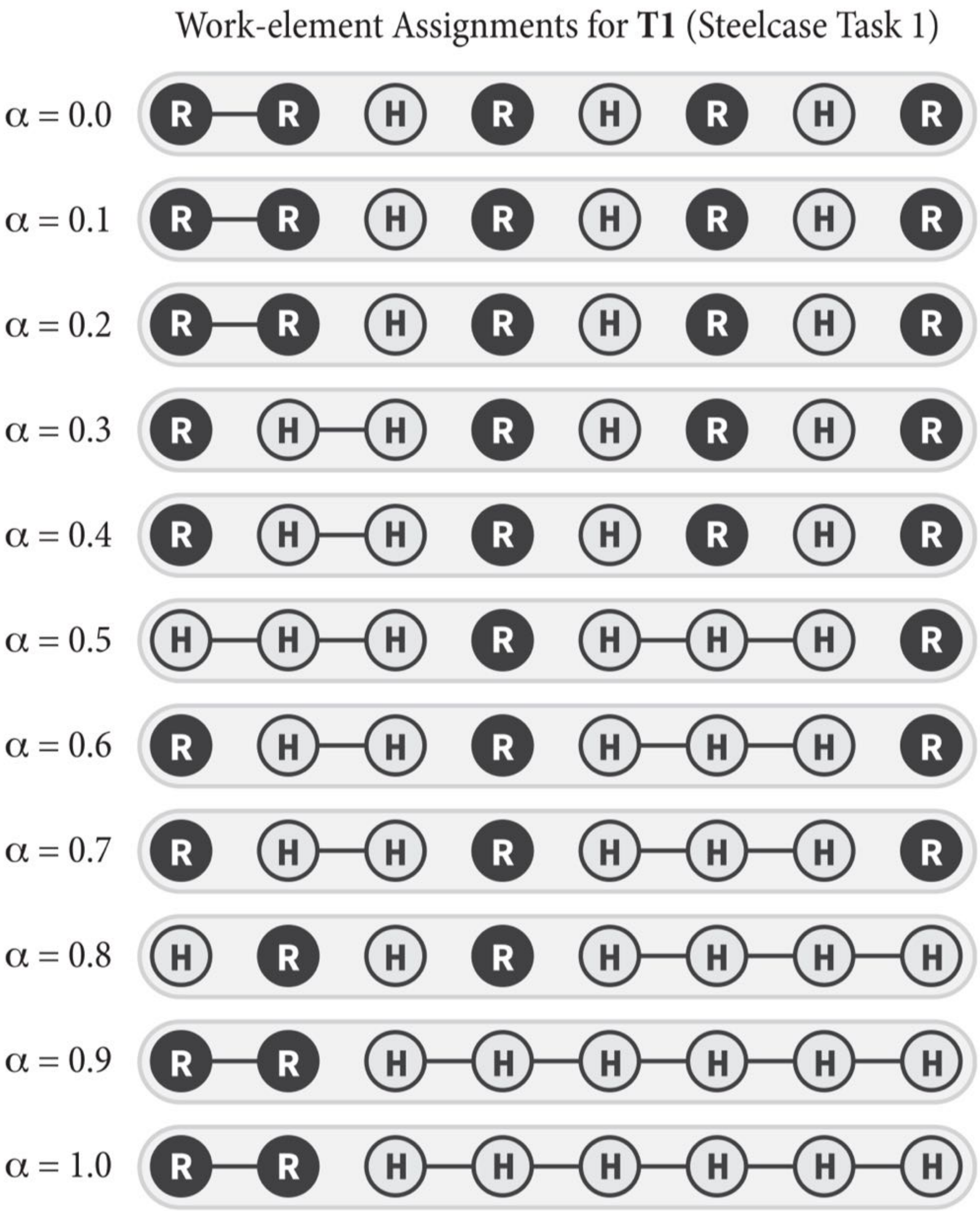
Research Question: *How do we translate a manual process into a human-autonomy process?*[21]

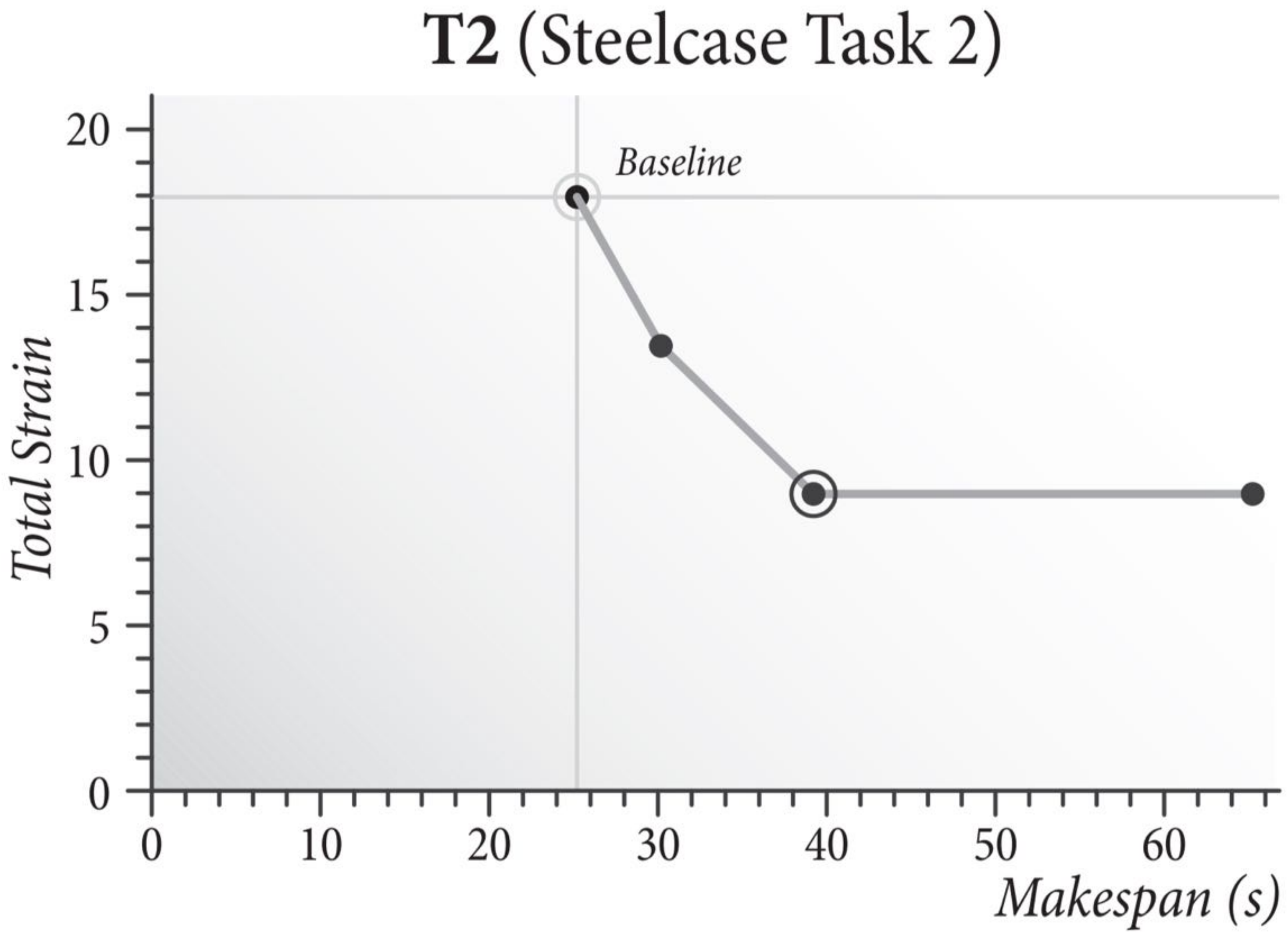
1. Model manual process
2. Allocate tasks to user + autonomy
3. Identify effects to identify optimal allocation

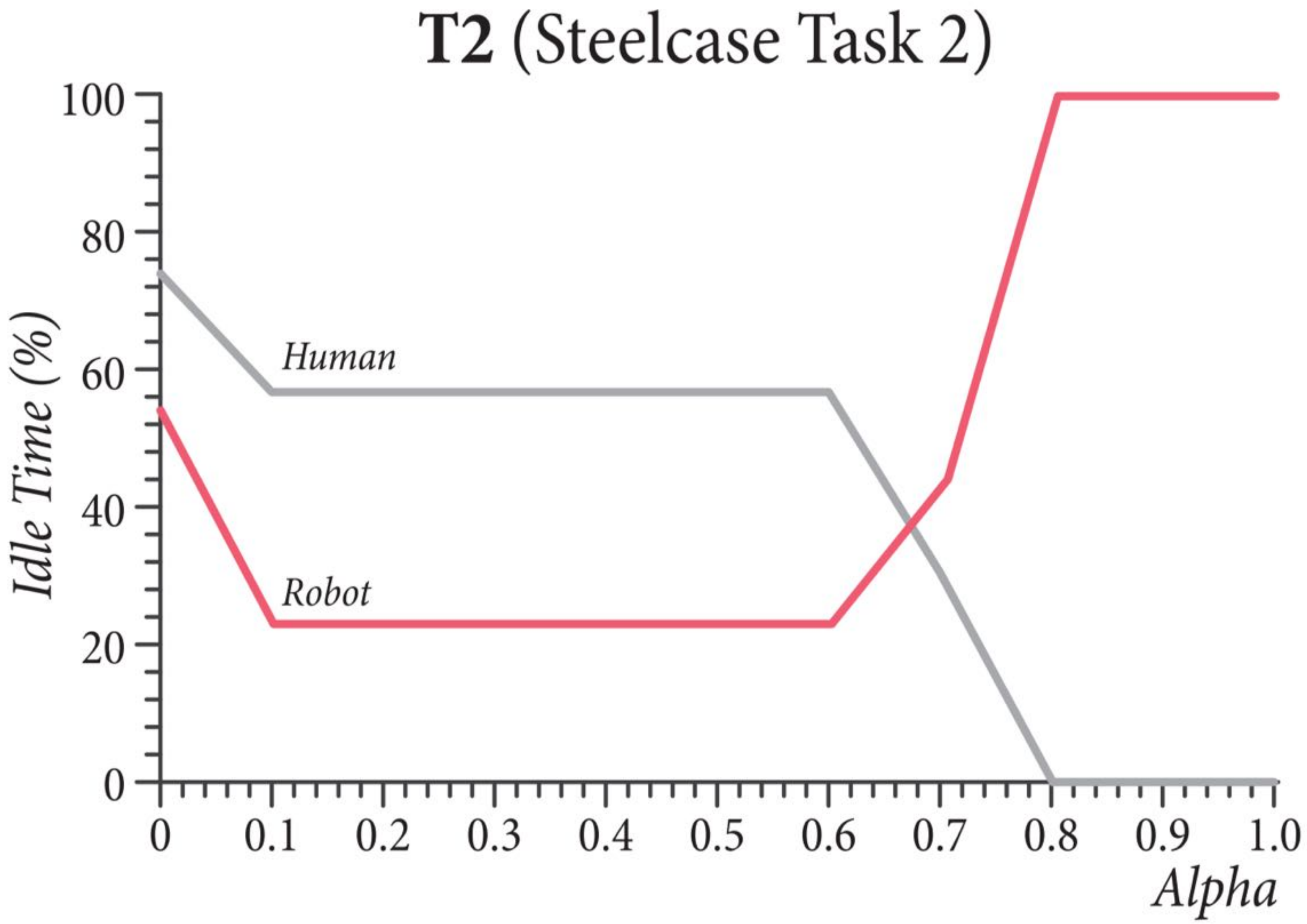


[21] Pearce et al. (2018). Optimizing makespan and ergonomics in integrating collaborative robots into manufacturing processes. *IEEE TASE*.




Work-element Assignments for **T1** (Steelcase Task 1)
α = 0.0 R R H R H R H R
α = 0.1 R R H R H R H R
α = 0.2 R R H R H R H R
α = 0.3 R H H R H R H R
α = 0.4 R H H R H R H R
α = 0.5 H H H R H H H R
α = 0.6 R H H R H H H R
α = 0.7 R H H R H H H R
α = 0.8 H R H R H H H H
α = 0.9 R R H H H H H H
α = 1.0 R R H H H H H H


T2 (Steelcase Task 2)
Baseline
Total Strain
Makespan (s)
0
5
10
15
20
0
10
20
30
40
50
60


T2 (Steelcase Task 2)
Human
Robot
Idle Time (%)
Alpha
0
20
40
60
80
100
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

How do we make sense of different forms of human-AI interaction?

Mixed-Initiative Levels[16]

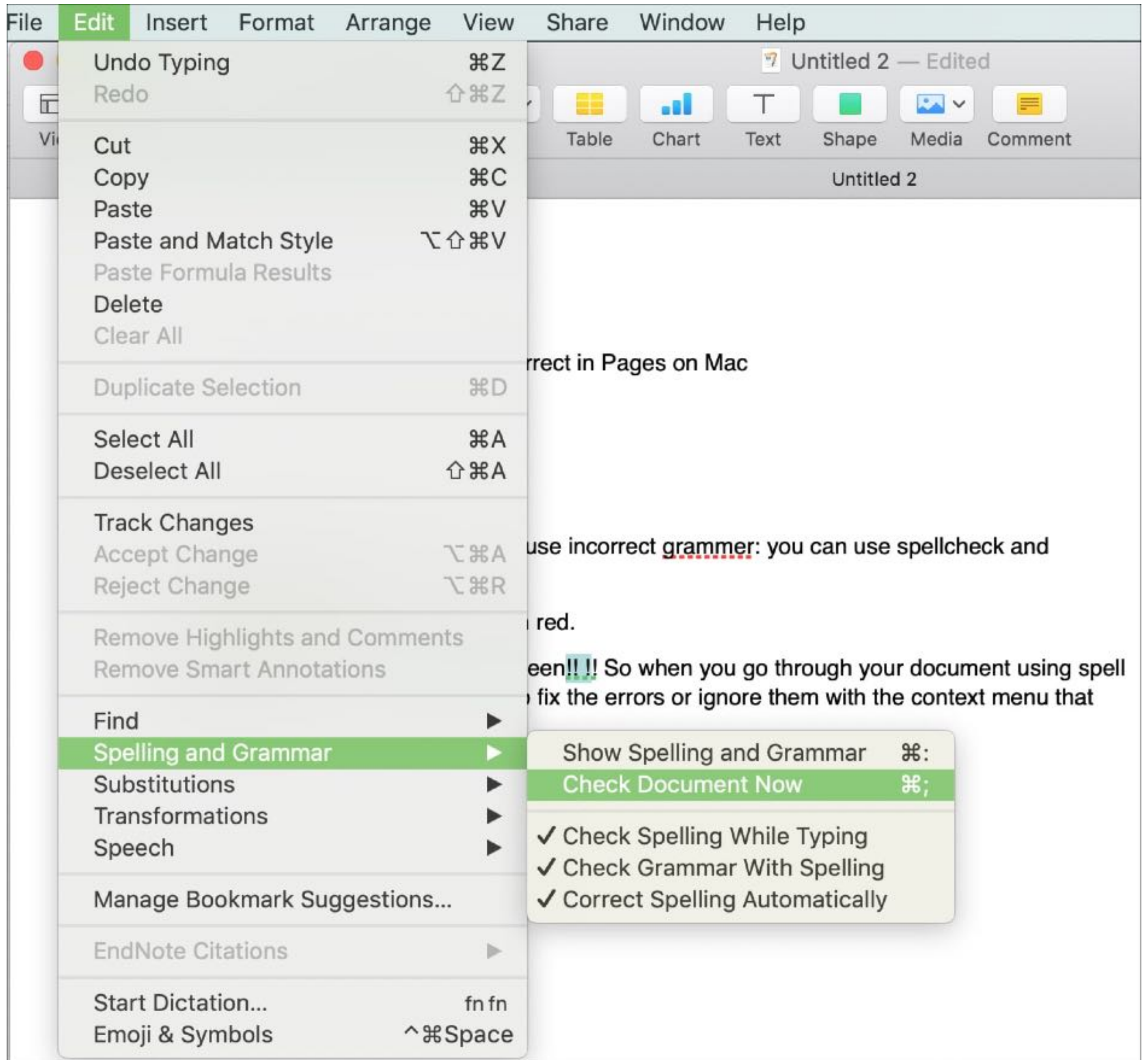
Mixed-initiative levels	Capabilities
Unsolicited reporting	Agent may notify others of critical information as it arises
Subdialogue initiation	Agent may initiate subdialogues to clarify, correct, and so on
Fixed subtask initiative	Agent takes initiative to solve predefined subtasks
Negotiated mixed initiative	Agents coordinate and negotiate with other agents to determine initiative

[16] Allen et al. (1999). Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications.*

Unsolicited Reporting[22]

System continuously checks human work and reports on findings

Example: Automatic spellcheck



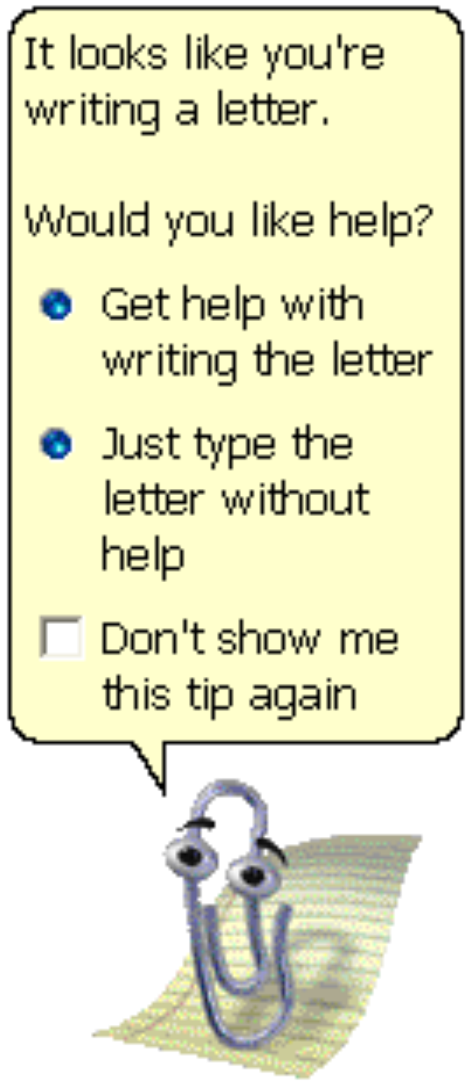
---

[22] Image

Subdialogue Initiation[23]

System engages in dialogue to get clarification from the user on whether help from autonomous features is needed

Example: Writing assistants

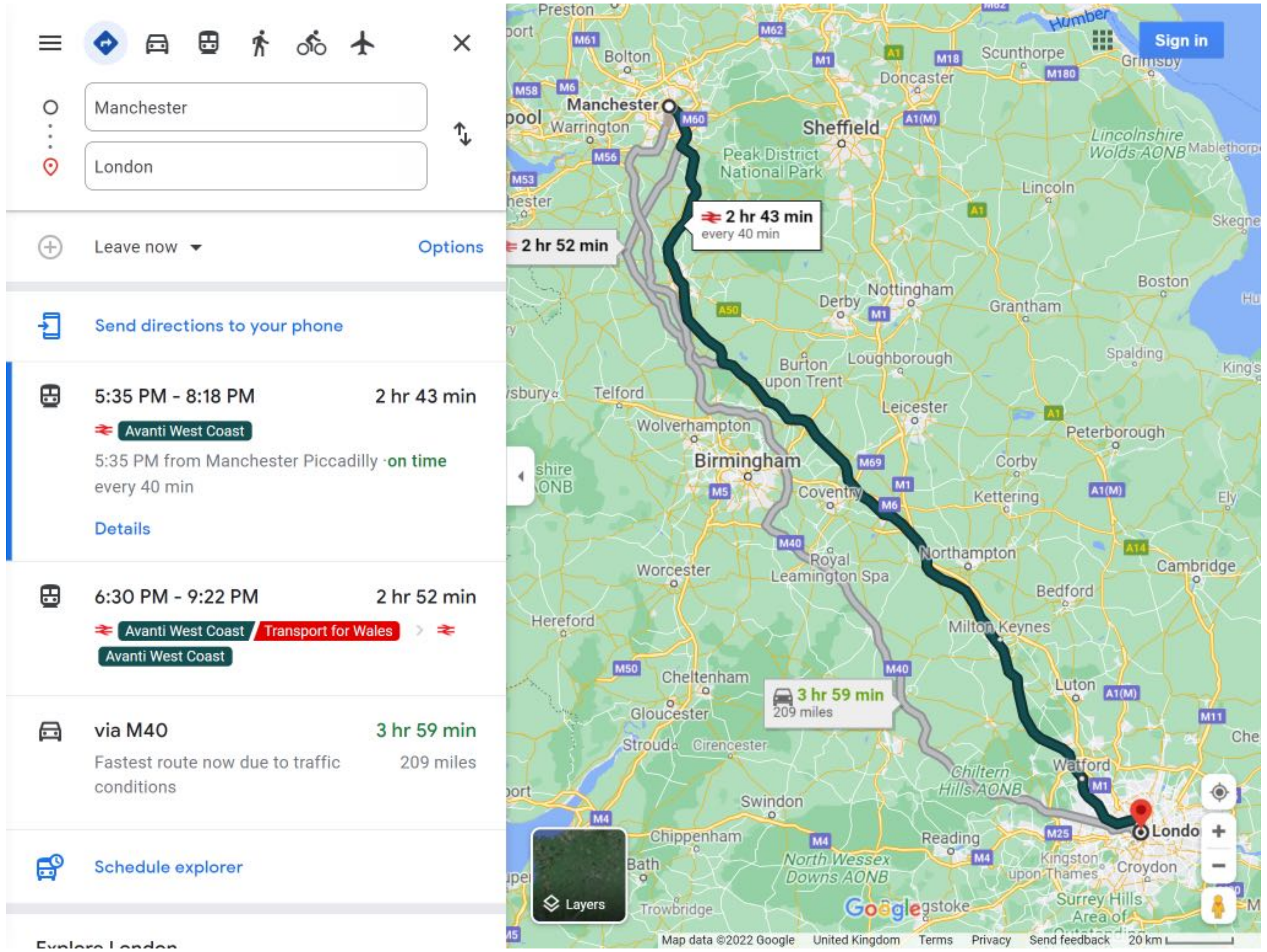


[23] Image

Fixed Subtask Initiative[24]

System engages in dialogue to get clarification from the user on whether help from autonomous features is needed

Example: Route planner



---

[24] Image

[29] Tesla via YouTube

Negotiated Mixed Initiative[25]

System constantly monitors, negotiates action, and/or takes initiative

Example: Autonomous breaking

[25] IIHS via YouTube

Other Classifications of Human-Autonomy Collaboration

Society of Automobile Engineers classification:[26]

» **Level 0:** No automation (e.g., manual driving, cruise control)

» **Level 1:** Driver assistance (e.g., adaptive cruise control, lane-keep assist)

» **Level 2:** Partial automation (automation of rudimentary tasks; e.g., lane centering)

» **Level 3:** Conditional automation (automation under ideal conditions; e.g., autopark)

» **Level 4:** High automation (automation except in poor conditions; e.g., snow)

» **Level 5:** Full automation (driver is not required)

[26] SAE Standard J3016_202104

How do we design AI-infused systems?

Approaches to designing human-AI interaction

Two key approaches:

1. Inspiration from human collaboration (e.g., joint action[27])
2. Design guidelines (e.g., Amershi et al.[1])

---

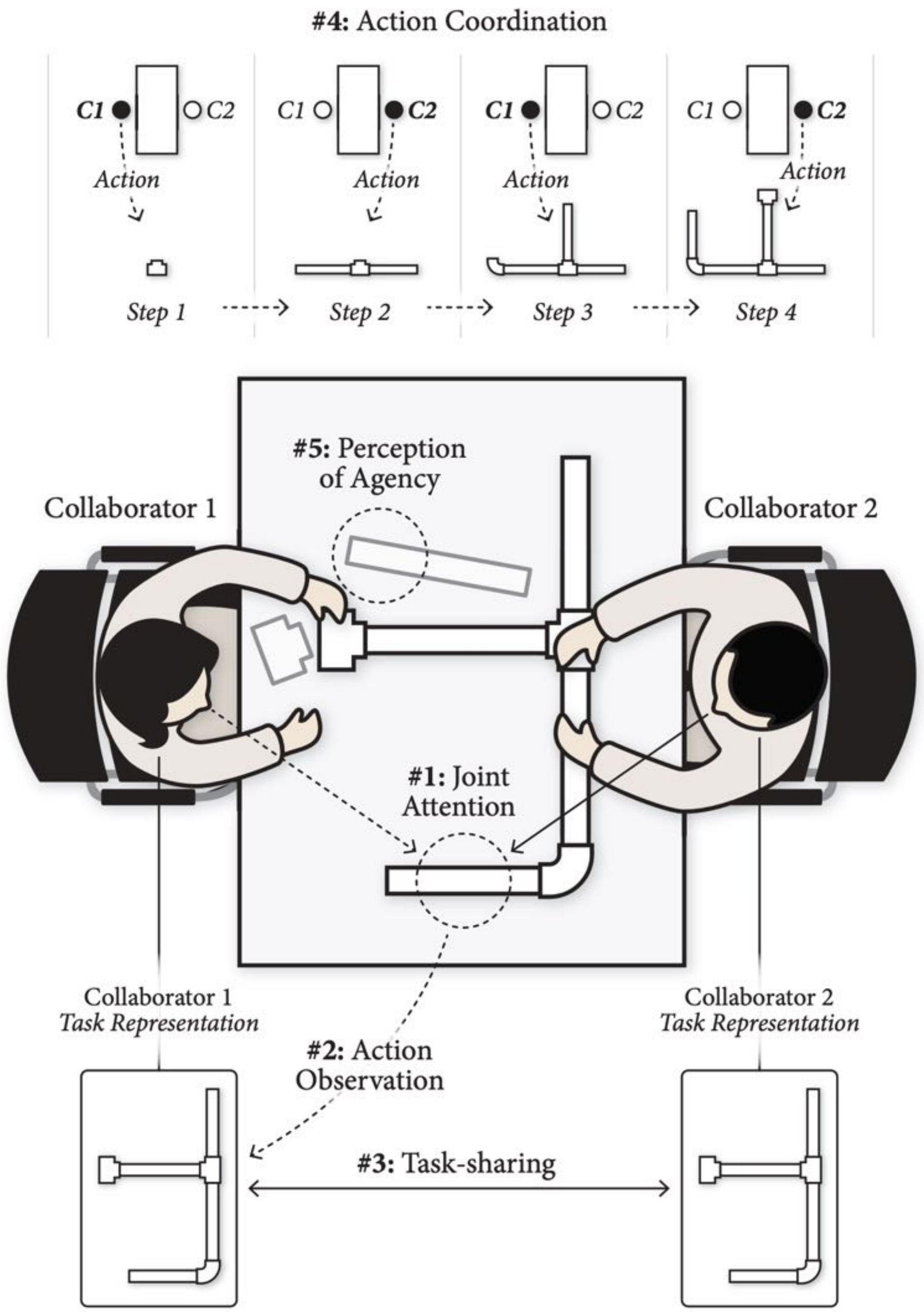
[27] Sebanz et al. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*.

[1] Amershi et al. (2019). Guidelines for human-AI interaction. *CHI 2019.*

Coordination Mechanisms

Human collaboration mechanisms:[28]

1. Joint attention
2. Action observation
3. Task sharing
4. Action coordination
5. Perception of agency



[28] Mutlu et al. (2013). Coordination mechanisms in human-robot collaboration. *Workshop on Collaborative Manipulation.*

Let's see this again:[29]

[29] Tesla via YouTube

Design Guidelines

Guidelines by Amerishi et al.[1]:

1. Initially (G1–G2)
2. During interaction (G3–G6)
3. When wrong (G7–G11)
4. Over time (G12–G18)

[1] Amershi et al. (2019). Guidelines for human-AI interaction. *CHI 2019.*

Example Guidelines[1]

G9 Support efficient correction: *Make it easy to edit, refne, or recover when the AI system is wrong.*

[Voice Assistants, Product #2] "Once my request for a reminder was processed I saw the ability to edit my reminder in the UI that was displayed. Small text underneath stated 'Tap to Edit' with a chevron indicating something would happen if I selected this text."

G13 Learn from user behavior: *Personalize the user's experience by learning from their actions over time.*

[Music Recommenders, Product #2] "I think this is applied because every action to add a song to the list triggers new recommendations."

[1] Amershi et al. (2019). Guidelines for human-AI interaction. *CHI 2019.*

Tradeoffs of Approaches

Model-based

- Top-down
- Higher generalizability
- Limited coverage of corner cases

Design guidelines

- Bottom-up
- Case-based applicability
- Limited use in conceptual design

Discussion Format

- Group discussion ~15 minutes
  - Separate to 9 groups randomly
  - Discuss with your group members
  - Take notes in the shared doc– pick your group number
- Summary from each group & discussion ~10 minutes

Discussion Questions

- What AI-infused systems do you use in your day-to-day lives?
- What form of interaction do these systems offer (intermittent/continuous/proactive)?
- What are challenges and limitations of interacting with these systems?
- What are real-world examples of human-AI collaboration that you know?
- Where else do you see potential for human-AI collaboration?
- What are promising design approaches?