Carnegie Mellon University Africa
Certificate I: Understanding AI and Machine Learning in Africa

Course AIML01: Artificial Intelligence – Past, Present, and Future

Module 4: Future Challenges
Lecture 1: Collaborating with Machines and Robots

Welcome to Lecture 1 of Module 4 in which we begin to explore the challenges that AI faces as it continues to develop and evolve. The future of AI will be devoted to overcoming these challenges. We start with the challenge of creating machines and robots that can collaborate with humans. In the remaining lectures, we will address the challenge of creating self-learning and self-programming machines, the challenge of dealing with the social and ethical aspects of AI, and the challenge of understanding the relationship between intelligence, brains, and consciousness.

In this lecture, we will look at the complexity and uncertainty of real-world interaction. We will explain why collaboration between people and robots is so challenging. We will look in particular at what is involved in reading intentions and establishing a theory of mind. We will discuss what is meant by variable autonomy and why it is needed when agents collaborate. We will explain why establishing trust is so important for collaboration and the challenges this poses for human-robot interaction. We will finish up by summarizing what we have covered and identifying the articles that you should read to consolidate what you have learned.

We have three learning objectives, so that, after studying the material covered in this lecture, you should be able to do the following.

1. Identify the challenges posed by designing robots that can interact and collaborate with people in natural, realistic, everyday settings.

2. Identify different levels and degrees of autonomy and provide examples of situations where machines and robots might need to exhibit them.

3. Identify the technical challenges of establishing long-term, trustworthy interactions between robots and people.

Slide 1 Welcome to Lecture 1 of Module 4 in which we begin to explore the challenges that AI faces as it continues to develop and evolve.

The future of AI will be devoted to overcoming these challenges.

We start with the challenge of creating machines and robots that can collaborate with humans,

In the remaining lectures, we will address the challenge of creating self-learning and self-programming machines,

the challenge of dealing with the social and ethical aspects of AI,

and the challenge of understanding the relationship between intelligence, brains, and consciousness.

Slide 2 AI has contributed significantly to the design of intelligent control systems and cognitive architectures for robots,

Both for sensorimotor behaviors - for example, perception, navigation, and manipulation

and cognitive capabilities - such as planning and language understanding.

We touched on some of these topics in Lecture AIML01-03-02. You might like to revisit it before proceeding.

But major challenges still remain for the reliable, robust realisation of these skills,

specifically to allow the robot to handle the uncertainty and complexity of real-word scenarios,

such as dealing with cluttered, dynamic, unpredictable environments
where objects to be grasped or obstacles to be avoided can be difficult to see or appear in different, unforeseen places.

Slide 3 A significant challenge for robots, and intelligent machines in general, is that of handling interaction with people for collaborative tasks, also known as human-robot interaction (HRI) and social robotics

Slide 4 This type of interaction includes a variety of scenarios,

such as joint action in flexible manufacturing setup between a worker and a cobot (collaborative robot)

Slide 5 Assistive robot companions for older and disabled people in hospitals and care homes,

Slide 6 And robot tutors for education or entertainment.

Slide 7 Within the field of HRI and collaborative human-machine interaction,

Some of the challenges for future research on the combination of AI and collaborative robots are

(i) the skills required for effective interaction

(ii) the mode of interaction with machines with a variable degree of autonomy, and

(iii) the quality of interaction for long-term, trustworthy interaction that fosters well-being.

Slide 8 The research challenge in using AI to achieve the social and cognitive skills for interaction includes, for example,

the ability of to read a person's intentions and the implementation of an artificial Theory of Mind (Vinanzi et al. 2019).

Slide 9 Intention reading is the ability of the robot to infer the person's intended goal when working together on some joint interaction.

For example, when a cobot is working with a person to assemble the parts of some object,

it must anticipate both the goal and the next action that the person expects the robot to perform.

Slide 10 Theory of Mind describes a more general view of intention reading,

as it concerns the robot's ability to understand and infer not just the person's goals, but also their beliefs and desires.

Slide 11 This sequence of pictures depicts a situation in which the iCub humanoid robot (www.icub.org) is interacting with a human, reading her intention to get her phone from her bag.

Note that this sequence has been staged to illustrate the desired capabilities of a cognitive robot and has not yet been implemented.

Slide 12 AI methods, such as Bayesian networks and deep learning, can be used to build artificial Theory of Mind skills in robots (Vinanzi et al. 2019).

A Bayesian network is a probabilistic graphical model which we met in Lecture AIML01-02-03.

Slide 13 Autonomy – whether or not a robot is self-governing and determines its own goals – is not black and white issue.

There are degrees of autonomy

While there are two extremes,

A fully autonomous car without a driver, for example,
versus
a robot mobile robot in a nuclear site tele-operated by a human

Slide 14 Most of tasks in which intelligent robots will interact with people involve a variable degree of collaboration

requiring a robot with a variable degree of autonomy to adapt to the user needs and to the environmental circumstances.

This is sometimes referred to as
Adjustable autonomy
Shared autonomy
Sliding autonomy
Subservient autonomy

Typically, the system controls its own behavior to a greater or lesser extent,
But the goals are determined by the human with which it is interacting

Slide 15 This could be for example the case of intelligent assisted driving and the six levels of driving automation identified by the car industry.

Slide 16 Future intelligent vehicles will dynamically switch from situations in which

the car performs some lane-following, steering and acceleration tasks autonomously

when it can easily recognise the road and traffic condition
(level 3: conditional driving automation)

to situations in which the driver must take control when the car is unable to perform the driving task (level 0: no automation).

Slide 17 Another important future research direction in AI for collaborative robotics concerns the quality of the interaction,

That is, the design of long-term and trustworthy interaction and wellbeing in human-robot collaboration.

Long-term interaction requires the robot to be able to engage in continuous, meaningful, and contextualized interaction

over a series of interactions lasting for days, weeks, or longer.

This will require robots to have an ability to recognize a person and their personality and preferences, to remember recent interactions, and to engage in empathic behavior that takes the person's needs into account.

Slide 18 This is a significant challenge in robotics, as the great majority of current intelligent robots are only capable of short-term interaction (typically one-session only).

Trustworthy interaction, a growing field of research, requires people's acceptance and trust of the robot's behavior and decision-making process.

This is also linked to ethical issue explainable AI (which we will cover in Lecture AIML01-04-04) and to the achievement of peoples' and robot's reciprocal theory of mind
In other words, robots will infer people's intentions and goals and people will infer robots' intentions and goals.

Slide 19 Which brings us back to this sequence of pictures depicts a situation in which the iCub humanoid robot (www.icub.org) is interacting with a woman. Now, after reading her intention to get her phone from her bag, the iCub alerts her to the fact that it is on the desk, hidden from her by the laptop, and the woman reads the iCub's intention, and sees the phone.

Again, as we noted, these abilities these pictures have not yet been implemented: they represent the future AI as it enables effective collaboration with robots.

To summarize:

The ability for two agents, such as a person and a robot, to collaborate requires poses many challenges.

1. The need for robust perception and action.
2. The ability to read intentions.
3. The ability to establish a theory of mind.
4. The ability to exhibit different degrees of autonomy, and
5. The ability to exhibit trustworthy behavior.

Here is some recommended reading.

Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijsers, M., & Šabanović, S. (2020). Human-robot interaction: An Introduction. Cambridge University Press.
Chapter 2 – What is Human-Robot Interaction? pp. 6-17.
https://www.human-robot-interaction.org

Sandini, G., A. Sciutti, and D. Vernon, Cognitive Robotics. In M. Ang, O. Khatib, and B. Siciliano (Eds.), Encyclopedia of Robotics. Springer, 2021.
http://vernon.eu/publications/2021_Sandini_et_al.pdf

Here are the references cited to support the main points in what we covered today.

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: a survey.
https://doi.org/10.1098/rstb.2018.0032

Mou W., Ruocco M., Zanatto D., Cangelosi A. (2020). When would you trust a robot? A study on trust and theory of mind in human-robot interactions. Proceedings of RO-MAN 2020, 29th IEEE International Conference on Robot and Human Interactive Communication, Naples, August 2020.
https://arxiv.org/abs/2101.10819

Vinanzi S., Patacchiola M., Chella A., Cangelosi A. (2019). Would a robot trust you? Developmental robotics model of trust and theory of mind. Philosophical Transactions of the Royal Society B., 374.
https://doi.org/10.1098/rstb.2018.0032

Vinanzi S., Cangelosi A., Goerick C. (2021). The collaborative mind: Intention reading and trust in human-robot interaction. iScience, 24(2).
https://www.cell.com/iscience/fulltext/S2589-0042(21)00098-5