IN5480 lecture notes 25 August 2021

Updated 25th August 2021

1 Introduction

Welcome.

We are two lecturers in this course; Asbjørn from Sintef and Jo from IfI. In addition, Claudia is working as a teaching assistant – and Trenton from NR/Sintef will contribute.

We have divided the course into three different modules. The three modules are:

- History and Concepts
- Design of interaction with AI
- Evaluation of AI in use

These modules are of course related, but we have nevertheless tried to separate out three distinct areas, one for each module. We start with an introduction and background module, and a guiding question: “What do we talk about when we talk about interaction with AI?”. Module two is about how to design interaction and AI interfaces. The third module is focusing on how to evaluate interaction and interfaces with AI. For these three modules, we will bring in research projects from IfI and Sintef that we have worked with and are currently working with.

There is a curriculum for each module. This consist of are articles that we will use throughout the course. It will be updated as the course proceed.

We believe that nobody is an “expert” in the field of “interacting with AI” today; it is an open and indeed evolving field. Our colleague here at UiO, Anders Kluge like to say that the computer is “i støpeskjeen”; a Norwegian idiom indicating that the computer, and the use of the computer, is evolving. Computing, and the use of computers 50 years ago are different and similar to computing and the use of computers today. Even though nobody is an expert in the field, we all have different experiences, thoughts and interests in the field of “interacting with AI” in this class. We believe that You all have unique interests and experiences with various areas that are valuable for all of us. We hope that you will engage with the course, discuss, investigate, present in class, think, reflect, do and write a lot - and by this learn something valuable about “interaction with AI”.

There will be both individual work and group work in this course. At the end of module 1, you will hand in the first group assignment; a wonder document. At the end of module 2, the mid-term report is due, and the final report from the groups are due at the end of module 3. This means that at the end of the course, you will have handed in the group assignment in three iterations.

The individual assignment is structured as follows. For each module, you will get tasks to discuss and do, and write about. You will at the end of the course have handed in the individual assignment in three iterations.

For both the group assignment and the individual assignment, we will have a peer review process; i.e. that you read and comment on each other's assignments. After each delivery, you will read and comment on another delivery, in a round robin fashion based on a list. For the group delivery, another group is reviewing, and for the individual deliveries, another student is providing the feedback.

So, one assignment in the course is to provide or give review to written reports. Equally important is to make use of the review that you receive - and learn and do something about it. For delivery 2 and 3 of both the individual assignment and the group assignment, it is expected that you write in the appendix what you have done with the feedback you have received.

Note that the peer review structure is a condition for attending this course. In order to do the course, you will be given the opportunity to learn from other participants individual writings, and from the project work done in the groups.

We will use the Vortex system during the course; for messages, for the overview of the plans and deadlines - so all communication from “the course” to you will be found there. This is an open platform that you hopefully are familiar with. It is open in the sense that what is published there is available to all other students. For the deliveries, each group will get a folder - and you can post and deliver your assignments there - both the assignment themselves - and the feedback. You are welcome to use any other platform for discussion, collaboration and communication for communication within groups.

A requirement for the course is that you have done all the assignments. In the end, there is an oral presentation.

Finally, some words about workload for the course:
This course is 10 ECT.
Lectures, 2 hrs every week, including a break.

Group work, x hrs every week.
Individual work, y hrs every week.
2+x+y=10-15 hours every week

If we say that 30 ECT is 30 hour work week, then 10 ECT is 10 hours work pr. week. If we say 30 ECT is 45 hours every week, then 10 ECT is 15 hours every week.

Before a short break; we “hand out” some “Guidelines for Human-AI Interaction”. After the break, we will discuss this.

The next section is going into Module 1 of the course.

2. Interaction with AI today

What follows is a transcription of a session with an Android phone, where: J=Jo, G=Google Assistant.

J: Please set an alarm for 45 minutes.
G: Done. Your alarm is set for 7:25.
J: Thanks
G: That's what I am here for/I was literally made for this kind of thing (and a smiley)/You’re welcome.
J: Please delete the alarm.
G: You can do that in the app

You are all familiar with setting an alarm; most of us do it frequently. Traditionally, we have used buttons and dials on physical wristwatches or alarm clocks, text commands in Linux, icons and lists in a calendar program or app - or as you saw her with your voice input via a microphone and voice output from a loudspeaker.

These are different dialogue systems, were interaction take place through or via an interface of some sort. When using voice, we may have a dialogue system based on *conversational interaction*; where we are talking and listening to the computer as if it was a conversation. We will in this course both in module 1 and in module 2 get back to conversational interaction, especially in module 2 where we will work with chatbots. And the word chatbot is a composed word of chat and robot. We will get back to what we talk about when we talk about robots more in module 1. Indeed, should the name of the course have been “Interacting with Robots?” What do we mean when talking about AI, Robots, Chatbots?

Then, back to setting the alarm with a conversational interface, and with voice commands. The example showed that this works fine as of today with setting the alarm - end to end. I find it particularly interesting and fine to get feedback, both visual and auditory from the computer (telephone) not only the relative time “an alarm is set for time=now + 45 minutes”, but that calculations are done by the computer, and it is stated that the alarm is set for 07:25. In the example in this text, you can calculate and see that the time of day when the alarm was set was 06:50. And it was Wednesday morning, preparing this example. Obviously, it was done somewhere, a particular place, in a time zone. The place was here in Oslo, at a table - with 4G wireless connection and the alarm clock was powered from the sun. This was some descriptions of the context of use, or the situated activity. Another situation of use that I am very fond of, when interacting with a clock, is in the evening when going to bed. Then to set an alarm for let's say 8 hours, and get a feedback from the machine, visually or orally, that this will be at 06:28 - gives both an assurance that the alarm is set ok - and I can judge if it is too late - or an ok time to be woken up. For me, this feedback from clock machinery is sometimes valuable, but is it an “intelligent” interface because it works like this? Is it a “smart” system?

Following up on the example a bit further, when trying to delete the alarm, I did not succeed in the dialogue, or in the conversation. I was sent to an “app”, where I was asked to find the alarm manually, find the appropriate places on the screen and press this. It was not deleted directly after the command, or part of the conversation “please delete the alarm”.

This, relatively concrete, mundane and un-remarkable example of a use situation can be investigated further - analysed, talked about - and it is definitely possible to work with understanding this interaction to design new ways of doing this. If the “media / or alarm audio level" is set to 0, or the machine is on “silent”, it would be good to get feedback on this for example. Or if we by applying insights into Universal Design can make this kind of functionality accessible to more users, in more situations, for more human activities?

However, there is another issue that I want to bring up with this example - apart from the specific interaction taking place. A clock, or a timepiece, or a time-system is an example of **autonomous** technology. It does something “on its own”, its primary function is handled, or operated autonomously by the clock machinery, without human intervention. When working, it just goes on doing what it does without any input or control from humans, or users. Even under water, in the sky - or here among us.

This clock technology can be understood and explained by dividing it into some modules that work together so that it is a clock. The modules are:

- An oscillator, something with a regular frequency.

- A counting mechanism that counts the oscillations.
- A display/interaction module that enables presenting the time or setting the time.

In his seminal book Computer Power and Human Reason, from Judgement to Calculation, Joseph Weizenbaum uses the clock as an introduction and explanation of autonomous technology. The first chapter of the book is called “on tools”, where he describes the evolution of important tools that we humans have made - both the tools that make us do different physical activities like cutting down trees and chop wood with axes, shoot pray over distance with guns, and sensory instruments like the microscope and telescope to let us discover more things in the world we live in. Among these tools, the clock is presented as having had a specific role, as an important **autonomous technology**. It is perhaps hard to see this now, we are familiar with the clock in our daily lives - but it is possible to imagine life before the clock - and see clock machinery as an autonomous machine – or as an Artificial Intelligence, and AI.

This is an example of a wall clock. It can be used to tell the time. This particular wall clock is set automatically from a very precise master clock in Germany, over the air. Much like the clock integrated into the smartphones we use today. You do not have to do anything to make it work, it works autonomously. And it brings its own logic with it. It does something very much more precise than any human can do, indicating “clock” time accurately. And it does it “on its own”, even “adjusting and correcting itself” from a master clock machine somewhere in Germany.

At 14:15 we started today. We are all users of this autonomous technology, and it enabled us to meet “here”. When it indicates 16:00, we will end the lecture. It structures the social setting here, and we are of course all very familiar with this way of meeting - it is a mundane, everyday technology that is one of the conditions for the possibility of living in the world we live in today. The point I want to make here is that it is an autonomous technology, and that autonomy is at the very core of many technology developments. Autonomous weapon system, autonomous cars and vehicles, autonomous conversational agents are examples of areas where AI is at work.

We are humans, and autonomy is used to describe one of the conditions of being human. Integrity is another word, often used to describe some of the same. What does happen when we mix these two types of autonomy? What happens when machine-autonomy and human-autonomy meet?

We are humans, not machines. We can experience thoughts and emotions like “today I am like a machine”, or being on auto-pilot, and even as a slave to a real machine; but we “are not” machines.

Of course, there are many similarities between humans and machines. Analytically, we can select some areas where there are similarities between humans and machines.

“Physical”; human bodies and "hardware" for computers.
"Sensors", vision with eyes for humans and camera or sensors for the computer.
"Effectors", the voice for humans and loudspeakers for computers.
"Reasoning", thinking for humans and calculations for computers.
“Emotions”, feelings for humans and simulations for computers.

Both the established research fields *evolutionary computing,* and *bio-inspired computing* deals with learning from humans, and other species that have evolved, making models of what is done by humans and other species - and then engineering and programming it.

There are some differences between machines and humans. Weizenbaum, Dreyfus, Suchman, Winograd, Agre, Searl and others have gone to great length to discuss important differences. One topic that often is described is about “making sense”, or “finding meaning” as a basic difference - where we humans seek to make sense and experience the world we are in - whereas machines do not. We experience being-in-the-world, being present here, now - whereas machines do not. It makes sense for me to be here now, for example, to do my job as employed by the university. For computers, it does not (as far as I know..). You have knowledge about calculations and nutrition for example, whereas machines do not possess this kind of “knowledge”. You learn new things every day. Machines do not learn, at least not in the way we have historically thought about learning. This lecture may make sense (or not) for you. This lecture does not make sense or mean anything for the computer. You may understand the equation 1+1=2. The computer does not understand 1+1=2.

And here, we are at an important issue with ways of talking and thinking about machines and of us humans. We use many words that historically meant something about being human, and apply it to systems and machines. Knowledge, understanding, communication, interaction, perception, intelligence, learning, sensing, agency, intention, dexterity, curiosity, affections, autonomy - these are all words that we have used to talk about human conditions, being a human.

Within AI, many of these concepts and words are used merely as metaphors, but in other instances, they are applied more literally. What is important for this module in the course is to discuss what we mean when using these words, and be open to discuss it. The concept of anthropomorphism and zoomorphism are ways of talking about this.

A pragmatic view is that we apply and use metaphors in computing - and when talking about using computers. The "desktop" on "Windows" is not a window, and a “web page” is not a

page in the historical sense. Navigating the web is different from navigation and wayfinding in the mountains or at sea. And machine learning is different to human learning. In what ways are humans and machines different? Do machines learn, reason, feel and dream?

Interaction with AI - a pragmatic discipline - a pragmatic view

AI is an established research topic in computer science and in industry. "Interaction with AI" is not an established research field; neither in computer science nor in industry. In the following, we will say something about the AI field as a pragmatic field.

In 2016 during the annual Google I/O conference, at the keynote opening session by the chief executive officer Sundar Pichai, it was announced that the old slogan "mobile first" is replaced by "ai first". This was repeated, and elaborated upon in later conferences. Google is communicating to us that AI will affect the products and services they work with in many different ways and many different areas, such as in the health domain. This is important messages since it is from one of the largest ICT enterprises of our time, and since services and products from this company are used by so many different people, in so many different situations, in so many different activities. Other examples are Siri from Apple, IBM Watson, Alexa from Amazon and Cortana from Microsoft. These are all said to be "intelligent" in some way, and powered by AI.

Microsoft, on their official website, writes a slogan for AI: "Amplifying human ingenuity; we' re excited about the opportunities that AI bring to organisations and people."

One of the many companies making platforms for AI, viv.ai has this slogan:
"Radically simplifying the world by providing an intelligent interface to everything." - and "intelligence as a utility". The way the company present itself on the web is, on the "front page" (it is very hard not to use the metaphor "page" here, of course, it is not a "page" literally (unless you have printed it onto a page), but text and images on a display in front of you):
"Viv is an artificial intelligence platform that enables developers to distribute their products through an intelligent, conversational interface. It's the simplest way for the world to interact with devices, services and things everywhere. Viv is taught by the world, knows more than it is taught, and learns every day."

What does this mean? It sounds incredible, that something is distributed through an intelligent, conversational interface, and that "the platform..knows more than it is taught, and (that) it learns every day". This can be interpreted in a pragmatic way that the company is

providing a platform that applies machine learning, conversational interfaces and knowledge systems or big-data systems. It is not “really” about learning, knowledge or conversation - as we have used these words and concepts traditionally.

Mass media is discussing AI. It is possible to find statistics on the frequency on articles, documentaries on this topic over the years since the term “AI” was coined. Indeed, it could be interesting to see the number of mass media articles over the years, we know that during the AI winters with little funding, AI got minimal attention from mass media. Today, it seems like there is much discussion of AI in mass media. Many people are engaged with different emotions, such as great hopes and great fears, and news about the hopes we have for what AI can do for us - and fears we have for unintended or unseen consequences are there.

And lastly, AI has been, is - and probably will continue to be explored in films, books and games. To name a few local and global: the film 2001 A space Odyssey from 1968 directed by Stanely Kubrick and inspired by a short story by Arthur C. Clarke. The TV series Blindpassasjeren by Jon Bing and Tor Åge Bringsværd from 1976 is another example.

So, we can say that AI is on the agenda among the technology companies, in scientific research, in media worlds and portrayed in films, books and games. On a macro level, on the political and social agenda, on a meso level, organisations for example and finally on a micro level - you interacting with a system.

What do we talk about when we talk about interaction with Artificial Intelligence? This is the guiding question for the first module.

Next week; we continue with this question, and will go into AI and Universal Desing. We will also continue with forming groups.

See you next week (here or there;)