AI – LECTURE 2

Social impact of artificial intelligence

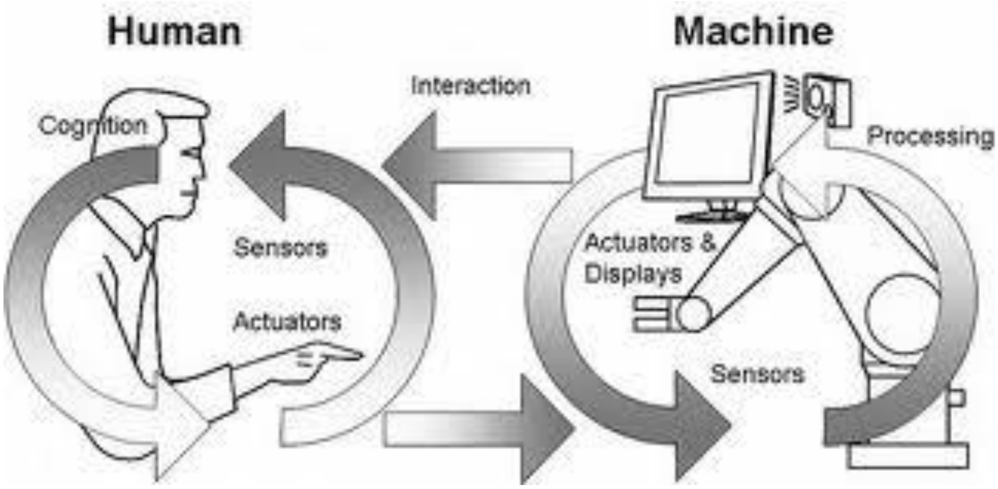
- When we are able to create machines smarter than humans, then those machines could do likewise, but much faster.
- The problem of the threat of unfriendly artificial intelligence and losing control of the machines we have built remains unsolved.
- Artificial autonomous agents are, in one sense the next stage in the development of technology. *Autonomous agents perceive, decide, and act on their own.*
- A machine (such as your brain) cannot predict (and thus cannot control) a machine of greater algorithmic complexity, which is a bound on a formal measure of intelligence. As a consequence, AI is an evolutionary process: each generation experimentally creates modified versions of itself without knowing which versions will be smarter.
- Moral destruction: could we become less human-indeed, less than human-as a result of advances in AI? This might happen if people were to come to believe that purpose, choice, hope, and responsibility are all sentimental illusions. Those who believe that they have no choice, no autonomy, are unlikely to try to exercise it. But this need not happen, for our goals and beliefs-in a word, our subjectivity-are not threatened by AI. The philosophical implications of AI are the reverse of what they are commonly assumed to be: properly understood, AI is not dehumanizing.
- Political destruction could result from the exploitation of AI (and highly centralized telecommunications) by a totalitarian state. If AI research had developed programs with a capacity for understanding text, understanding speech, interpreting images, and updating memory, the amount of information about individuals that was potentially available to government would be enormous.
- Many people fear that in developing AI, we may be sowing the seeds of our own destruction, our own physical, political, economical, and moral destruction. Computer technology (and AI in particular) cannot in principle achieve the reliability required for a use where even one failure could be disastrous.
- Do you really want a future where you can have everything you want, including the ability to change what you want?

AI is now mature in applications approach, both as a science and, in its technologies and applications, as an engineering discipline. Many opportunities exist for AI to have a positive impact: -

- AI researchers and development engineers have a unique perspective and the skills required to contribute practically to addressing concerns of global warming, poverty, food production, arms control, health, education, the aging population, and demographic issues.
- Assistive technology for disabled and aging populations is being forged by many researchers.
- Assisted cognition is one application but also assisted perception and assisted action in the form of, for example, smart wheelchairs and companions for older people and nurses' assistants in long-term care facilities.
- AI could potentially automate all human labor.
- There are issues of robot liability and insurance being fronted by ethical researchers. There will have to be legislation that targets robot issues. There will have to be professional codes of ethics for robot designers and engineers just as there are for engineers in all other disciplines.
    - What should we humans do ethically in designing, building, and deploying robots?
    - How should robots ethically decide, as they develop autonomy and free will, what to do?
    - What ethical issues arise for us as we interact with robots?

Man-Machine Interface (MMI):

The discipline of Man/Machine Interface is the study of communication between machines and humans.



Human-machine interaction is goal-oriented. The overall goals of human-machine systems are mainly
(1) Productivity goals,
(2) Safety goals,
(3) Humanization goals, and
(4) Environmental compatibility goals.

The **productivity goals** include economic as well as product and production quality goals.
The **safety goal** is strongly influenced by the application domain. This goal class dominates all others in many large-scale systems and, particularly, in risky systems.
The **humanization** goals comprise team and work organization, job satisfaction, ergonomic compatibility, and cognitive compatibility.
The **environmental compatibility** goals refer to the consumption of energy and material resources as well as to impacts on soil, water, and air.
Human-machine interaction and human-machine systems research require multi-disciplinary or interdisciplinary views and approaches.

The following three domains contribute to human-machine interaction and systems research:
(1) **Cognitive science** and ergonomics (as the human sciences),
(2) **Automation and systems engineering** (as the systems sciences), and
(3) **Information and communication** engineering (as the computer sciences).

General Principles

- Feedback to user - The ideal form of feedback is allowing the user to see things happen.
- The user should be in control - the computer should not be issuing commands and expecting the user to obey. The user prefers that he can make commands and the computer obeys. The computer should be a helpful servant, not a cruel master.
- Predictability - users should be able to guess what will happen if a key is pressed, caps lock is used, text is highlighted etc
- Transparency - means that the user can easily find out what is going on behind the surface.
- Never interrupt the user - Time and attention are precious resources to many software users. Stealing those resources is a sign of disrespect to the user's work situation.

- Can I guess what the user wants? - Modern software has many of these "smart" features. But they violate the principle of predictability and the principle that the user should be in control. The software should never do anything on its own initiative, not even correct errors.
- Error tolerance - Human language is generally quite error tolerant. If one word of a sentence is wrong or missing then you can still guess what the meaning should be.
- WYSIWYG - **what you see is what you get**. It is a common buzzword in the design of user interfaces. It means that when you write or draw something in a computer program, then the screen should look exactly as the final document will look when you print it out.
- Speak the user's language - It is nice that the program tells the user what it is doing while the user is waiting for the document, but this message must be understandable.
- Avoid anthropomorphic interfaces - Some designers are tempted to make user interfaces look like cartoon characters or communicate with the user in a human dialog style. Experiences with anthropomorphic user interfaces are generally not very good. For example, Microsoft's animated paper clip called "Office Assistant" was so unpopular that they had to remove it. Many people found it annoying, and some even considered it an insult to their intelligence.
- Design should reflect the user's logic, not the constructor's logic - operations that represent different working situations to the user should be kept in different parts of the user interface.
- The design of a button should reflect its importance - The buttons on a control panel or keyboard should be arranged so that the buttons that are used most often are big and have the most prominent positions, while buttons that are seldom used may be smaller and put away in a corner. Buttons that are used for basic setup and configuration may be put away under a lid, so that people know that these are not for everyday use.
- Provide alternative ways out of a situation – e.g Proceed to next step, Return to previous step, Undo last operation, Undo more than one operation, Redo undone operation, Cancel operation, Save data and interrupt operation so that I can resume later, Help about this step, Help about program in general, Leave program
- Accessibility to handicapped users - Some of the handicaps you may consider are: missing or paralyzed limbs, impaired motor skills, shaking or imprecise movements, impaired hearing, reduced vision, blindness or color blindness, reduced cognitive skills
- Novices versus experienced users - New users that haven't seen your product before need guidance on how to use it. The best guidance may be a self-explaining design that shows what the apparatus can do. If it is a machine then there should be a button or handle for every function. The heavy user who has to do the same tasks hundreds of times a day will find it annoying if he has to click his way through a lot of submenus every time he does a simple task. Not only is this time-consuming, heavy use of the mouse also causes strain injuries which in the worst case can permanently disable the user. Therefore, there should be shortcut keys for all common commands.
- Standardization - The lack of a universal standard for numeric keypads is a problem. The user will type faster and make fewer errors if the keyboard layout is always the same. And blind people would certainly prefer that all telephones have the same key placements. If you design a control panel or other hardware user interface, be sure to check if there is a standard or common practice for the colors, shapes, and placements of buttons etc
- Open standards - All technical products that have connections with other products need a standardized interface.

Interaction with smart environments

Smart environment as "a small world where different kinds of smart device are continuously working to make inhabitants' lives more comfortable." Smart environments aim to satisfy the experience of individuals from every environment, by replacing the hazardous work, physical labor, and repetitive tasks with automated agents. There are three different kinds of smart environments for *systems*, *services* and *devices*:

1. virtual (or distributed) computing environments,

2. physical environments and human environments,
3. hybrid combination of these:

- *Virtual computing environments* enable smart devices to access pertinent services anywhere and anytime.
- *Physical environments* may be embedded with a variety of smart devices of different types including tags, sensors and controllers and have different form factors ranging from nano to micro to macro sized.
- *Humans environments*: humans, either individually or collectively, inherently form a smart environment for devices. However, humans may themselves be accompanied by smart devices such as mobile phones, use surface-mounted devices (wearable computing) and contain embedded devices (e.g., pacemakers to maintain a healthy heart operation or Augmented reality **(**AR) contact lenses).

Smart environments are broadly classified to have the following features

1. *Remote control of devices*, like power line communication systems to control devices.
2. *Device Communication*, using middleware, and Wireless communication to form a picture of connected environments.
3. *Information Acquisition/Dissemination from sensor networks*
4. *Enhanced Services by Intelligent Devices*
5. *Predictive and Decision-Making capabilities*

To build a smart environment, involves technologies of

1. Wireless communication
2. Algorithm design, Signal Prediction & Classification, Information theory,
3. Multilayered Software Architecture, Corba, middleware
4. Speech recognition
5. Image processing, Image recognition
6. Sensors design, calibration, Motion detection, temperature, pressure sensors, accelerometers
7. Adaptive control
8. Computer Networking
9. Parallel processing
10. Operating Systems

Smart Device

Is an electronic device that is cordless mobile, always connected (via WiFi, 3G, 4G etc.) and is capable of voice and video communication, internet browsing, geolocation and that can operate to some extent autonomously.

Smart Device Characteristics

- *Smart Devices* in general can take a much wider range of form-factors than appliances.
- *Smart Devices* support the ubiquitous computing properties.
- information appliances focus on remote interaction with computing environments that tend to be personalized
- Usually personalised devices, specified owner.
- Locus of control and user interface resides in the smart device.
- Dynamic component-oriented resource extensions & plug-ins of some hardware resources
- Access to specific external environments: human interaction, physical world interaction and distributed ICT / virtual computing interaction.