

Human Computer Interaction (In English)

Prof. Rajiv Ratn Shah

**Department of Computer Science and Engineering
Institute IIT Madras**

Interface

Hello everyone. Welcome to the class of Human Computer Interaction. Let's get started. So this week we will be talking about interface. So this is the outline of previous lecture that we have covered. We discussed about user perspective.

We have discussed the importance of user, how we can learn their requirements, about their preferences and so on. We discussed about persona. We discussed about data synthesis techniques. We discussed about, through some real world examples, about how we keep user in the center while solving any real world problems.

This is the outline for this week. We'll be talking about interface, interaction, and their types. We'll be talking about conceptual modeling and cognitive aspect of the interfaces. We'll be also talking about emerging trends in the interfaces, some cognitive aspects and so on. We'll also be having two hands-on session this week.

So one with Adobe Express and another with Behance. Okay, so let's get started. Let's start with what is interface, interaction, and their types. So interface are the medium through which users and system communicate. In human-computer interaction, they're designed to make interactions effective, efficient, and satisfying, which will eventually contribute toward the user experience.

the key element of an interface is input which a user give to the system output which user will receive from the system and some kind of feedback again some kind of output will be getting from system in response to probably what user actions were done do not normal said the best interface is no interface. It means your interface should be in such a way that it shouldn't look like a barrier. It should be intuitive and simple enough to interact with it. So when the user doesn't notice the system, but simply achieve their goal. So importance of the interface is visible by these points, it enhances the usability.

Good interfaces make system accessible and intuitive. For example, ATM interfaces with clear and step-by-step instruction for withdrawing the money, checking the balance, or whatever. It should facilitate accessibility. So inclusive design ensure usability for people with disabilities. We have discussed several accessibility guidelines, which need to be

followed.

For example, screen reader for visually impaired user. So design is not just about what it looks like and feels like. Design is how it works, T.E.

Jobs said. Additionally, it should drive the innovation. Interfaces open new possibilities for interaction. So it is not just limited with the traditional way of interacting with the system, but to gesture control in Tesla vehicle, which has basically make it quite popular and also improve the better user experience with the Tesla users. So design must reflect the practice and aesthetic in business. But above all, good design must primarily serve people, said by Bell.

Similarly, users spend most of their time on other sites. Make your work the same way, Jacob said. So in human machine interaction, as you can see that there are primarily three components. One is the human, which with diverse needs, skills, goals, limitation, capabilities, emotion, cognitive and physical capability, and these cognitive and physical abilities influence the way user interact with the system that is a machine here. So machine is hardware and software based system.

For example, it could be computer, vehicle, robots, mobile phone, smartwatch, and so on. And it could include different sensors, actuators, and algorithms. An interface, that is something you can see, it is a medium that enable the interaction between human and the machine. For example, it could be touch screen, It could be swiping, it could be button clicking, it could be voice assistance, it could be some gesture. Interaction refers to the communication and the action between a user and a system facilitated by interface.

It involves Input from the user, as discussed, it could be anything like clicking, typing, gestures, speaking, and the system response corresponding to that, it could be probably displaying some result, giving some feedback, probably generating some sound, or some haptic input, and so on. Key elements of interaction, user, system, context, and the goals. So definitely user is one of the key element where an individual engage with the system. It could be a group of individual as well. System, the software or hardware enabling the task which the user want to perform.

Context, the environment and or the situation where interaction occurs. For example, for a Tesla car, it is driving on the road or probably in the parking lot you are trying to park or maybe in the garage. And all the interaction will be accordingly. Goals, the objective of the user aims to achieve. For example, on the road, you want to drive the car safely and probably reach to your destination.

So interaction is the essence of human-computer interaction. It transforms static system into a dynamic one and user-centric experience. So importance of interaction in HCI, I believe it's quite clear. It enhances usability and efficiency. It also improves satisfaction and the engagement of the user with the system.

And it bridges the gap between the user goals and the expectation that they have with the system functionality which is completed. Real-world examples, I believe there are so many examples you can see around. One example is mentioned here is airline check-in kiosk. So touchscreen for selecting seats, printing boarding pass and so on, that is something you perform. Augmented reality, user interacting with virtual objects in AR games like Pokemon Go is one of the most popular AR games has been built so far.

Educational apps where interactive exercises, quizzes, progress tracking and so on can be visualized. So there are several interaction types in HCI, instructing, you simply issue some command and selecting some options to perform. You can even converse with the system. Conversation can happen in many ways, it could be through speech, it could be through text and so on. So interacting with the system as if having a conversation, manipulating So interacting with an object in a virtual or physical space by manipulating them.

For example, zooming and tilting. Exploring, so moving through a virtual environment or a physical space. AR, VR is something that we do. Responding, so the system initiate the interaction and the user chooses whether to respond. For example, often you interact with some website and some alert comes.

So you need to respond to that. For example, moving to a new website, you're leaving the current website and moving to some other new website. So let's go about all of them one by one. So let's talk about first instructing. So where user instruct a system and tell what to do and often these are some well-defined tasks.

For example, tell the time, print a file, save a file, delete a file, move a file, stop the car and so on. A very common conceptual model underlying a diversity of devices and systems, for instance, word processor, VCR, and the vending machines. Simple instruction you give, okay, so dispose this item. Or probably, I want to buy this item, so you just get the cold drink or whatever. So the main benefit is that instructing supports quick and efficient interaction.

So it is good for repetitive kind of actions. So that's a point you must keep in mind. So often instruction is very, very useful when the task is quite repeated in nature and performed on multiple objects. One thing probably you can also note that the chances of

error in instructing is quite less. In fact, if the instruction has been defined well, and the corresponding action has been implemented, there is no chance of doing error.

So let's see two examples. You have two vending machine, A and B. Which one is easiest and why? What do you think? In this case, you have different items here. And in order to get the item, you'll have to press whatever the code is mentioned here, and then probably insert money and get it done. On the other hand, on the right side, you can see there are different drinks are there. And in order to get the item, you simply have to press these buttons.

What do you think now? Which one is easier? I believe all of you think and agree with me that the B is the easiest option because the chances of error in this case is too low or in fact negligible as compared to the first one because here you may enter wrong code. probably the code which is written, maybe the sticker is old. It is somehow erased and probably even you can't read it and so on. But in this case, it's quite easy. Just insert the money and press the button.

You'll get the corresponding item. So again, so this is kind of example of instruction and you'll have, so you also have to design and define instruction in such a way that the chances of getting error is negligible. So next one is conversing. So underlying model of having conversation with another human or another machine. So it ranges from simple voice recognition menu driven system to more complex natural language dialogues. So especially in the era of chat GPT, now I believe conversing become quite easy, quite accurate.

Because earlier the conversation used to happen probably in some fixed set of options that you have. If you speak in natural language, if you type in natural language, probably the machine or the system may not able to understand you, but now the things are changed and in good direction. So example include timetable, search engines, advice, giving system, help system and so on. So there are also virtual agents, chatbot, toys, paid robot designed to converse with you. Often at many galleries or probably museums, you might see their robots are there, which are there to help you, especially with the young kids to tell more about what all the arts are there, what all the unique items are, antique items are there and so on.

So there are good and bad things for sure. So it allows users, especially novice one to interact with the system in a way that is familiar to them. It's quite easy for them to probably follow. So it can make them feel comfortable at ease and less scared. On the other hand, It could also be that misunderstanding can arise when the system doesn't know how to parse what the users say, especially when the users start speaking in natural

language and probably in language other than English or probably in some other dialects. I think it's very hard for the system to understand.

In chatbot as well, for example, people start using Hinglish rather than probably English then again the system is not able to understand you so in those case it fails so like in this case example voice assertion can misunderstand what children or person with speech disorder say another way of another way is manipulating So here basically it involves dragging, selecting, opening, closing, and zooming actions on the virtual objects. It exploits user knowledge of how they move or manipulate physical world. So for example, in real life, when you have to basically throw the garbage, you always, for example, throw in the dustbin. So in this case, can we do something similar? So often you might have seen that on Windows desktop, there is a recycle bin box. You can drag and drop or kind of trash any of your file by just dragging the file to and putting into the recycle bin.

What happens? You are doing some kind of manipulation. You are deleting the file, throwing the file by putting that in the recycle bin. That is the garbage box here. so it can involve action using the physical control so like nintendo or air gesture such as microsoft connect is one of the popular one to control the movement of on screen of that it can also tag physical objects that are manipulated in physical world result in a physical or digital event such as animation so direct manipulation this term has been coined by One of the renowned scientists, Ben.

So it has three components. Continuous representation of object and action of interest. Physical action and button pressing instead of issuing commands with the complex syntax. Rapid reversible action with immediate feedback on object of interest. So one example that I've given, direct manipulation you've seen, just drag in the file and probably putting the recycle bin, you see immediately, the file has been removed from the desktop.

Now it is in the recycle bin and so on. So benefit of that manipulation, it's quite easy for novices to probably use it. So novice can learn the basic functionality quickly and adopt the system. Experienced user can work extremely rapid to carry out a wide range of tasks, even finding new functions. Intermittent user can retain operational concepts over time. Error messages rarely needed because you can see often with eyes itself what is happening.

So for example, deleting the file, you can see that now it has removed from the desktop. So users can immediately see if their actions are furthering towards their goal or not. If not, then do something else. So again, take the example of file. If you are trying to

remove the file, delete the file, and even after, even after just dragging it to the recycle bin, you still see the file is still there.

It means it has not, immediately you can see there is, file has been deleted. So user experience less anxiety because they are satisfied, they are able to see what is happening. And if you keep your user informed, I think it's quite happy for them and at least they can understand they are inching towards their goal. so user gain confidence and mastery and feel in control because they know what is happening and they are able to perform everything but at the same time there are some disadvantages as well of the direct manipulation some people take the metaphor of direct manipulation too literally similarly not all tasks can be described by object and not all actions can be done directly Some tasks are better achieved through delegating. For example, spell checking can become screen space gobblers.

So basically in that case, often you see your screen is full of all those direct manipulation. So moving a cursor using a mouse or touchpad can be slower than pressing function key to do the same, right? So often, for example, if you have to select all lines in a file, for example, there are 100 pages. Direct manipulation, just selecting all the lines, all the pages, you know how time taking and cumbersome process is this, but simple instruction you can do to select everything. Control A that you can do, or just go to the file and do select all.

So the next way of interaction is exploring. So where it involves moving through virtual or physical environment. So in that case, user can explore aspects of virtual 3D environment. It is more immersive in nature. So physical environment can also be embedded with sensors that when detect the presence of someone will trigger digital or physical event to happen. So many example of virtual environment, including cities, park, building, rooms, data sets, enable user to fly over them and zoom in and zoom out in different parts so that is exploration you can do in the virtual world through augmented reality in the virtual reality mixed reality and so on so for example in this case you can see that you can visualize a probably an insect in a cave In this case, sometimes things larger than life in virtual reality.

So for example, here the insect probably looks same size as human, which is not true. And similar thing you can see here in the example of space when the person is visualizing the space here. Another way of interacting is responding. System takes the initiative to alert user to do something that is think is of interest. Often you see that system does this by detecting the location and or presence of someone in the vicinity and notify them on their phone or watch.

You might have seen that sometime ago Facebook had a feature Is there any friend nearby? In that case, I mean, when you pass by the vicinity of some of your friend, your friend get an alert. Okay. So probably you're passing by. So you can, again, you can respond to that or no, but again, in a way it may be somehow you breach of privacy as well. Sometimes it would, for example, you are a tourist you are in some new location some new city and there basically you get a kind of notification okay so these are places of interest then probably you will like it so alert the user nearby coffee bar with some friends are meeting similarly user fitness tracker notify them of a milestone reach so often many of you might have might be using smart watches these days, where it continuously counts how many steps you have walked today.

And you can set up a milestone, for example, every day I want to walk 10,000 steps. And probably what all progress you are making in the day, it tries to notify you, okay, so I think you're doing well. So far, you have walked 5,000 steps, half of the day is left, probably you can do it, or probably more encouraging and so on. So Automatic System Alert without any request made by the user. So choosing an interaction type, which one to choose? Because as we discussed, there are many of them.

So each interface type offers unique advantages and drawbacks. Selecting depends on many things. So firstly, purpose, suitability for the task and user, Second is the environment where the context of interaction is very important. Thirdly, the resource availability of the technology and the budget as well because you can't apply AR VR everywhere because it's costly. Direct manipulation is good for doing things.

For example, designing, drawing, flying, driving and so on. Issuing instruction is good for repetitive tasks. for example, spell checking, file management, and so on. Having a conversation is good for certain services, for instance, finding information or requesting music, where basically you want to probably looking for information on Google search, for example, but probably you may not recall exact keyword to find that. Probably whatever you're typing, you can get suggestion and based on that, probably you can further inch towards the things you're looking for. Hybrid conceptual models are good for supporting multiple ways of carrying out the same actions.

So this is the general diagram of multimodal human-computer interaction. Multimodal because You can interact with the system in more than one way. Multiple modalities are there to interact with you.

And you have multimedia systems. Again, it could be any machines. It could be digital systems. And there are interfaces. And again, you can interact with the system in several ways.

So, for example, you can also speak. You can also probably type. You can also give gesture. You can also probably, based on probably your physiological signal it can do something so there is probably another advanced area like brain computer interaction so for example it can sense the your brain activities and probably understand what you want to say so that is also one of the you can say the futuristic and probably doable now direction people are working on. Just like as in human, when there is something wrong, I mean, we apologize. And the other person usually say, it's okay. But do we have such thing in computers? Do we have such thing in system initiate notifications? So does it allow users to correct the advice or information? So it's more like, If the system is at fault, I mean, does it listen to you? Does it recover from it? That is something we have to be careful about.

So there are several types of interfaces. GUI, which has probably changed the fortune of Apple. So graphical user interface. So it uses visuals like icons, menus, windows, and so on. So for example, Microsoft Windows, Mac OS interface, mobile apps, and so on.

Pros, intuitive and user-friendly because you can visualize, you can see. Support multitasking effectively. So you can perform multiple things and get you done. Reduce learning curve for new user because they can see what is happening. And the cons is, resource intensive because often these GUIs requires I mean front-end programming where basically for example you are in different javascripts and so on so that requires a lot of memory so requires advanced hardware often for example if you are more interacting with in immersive environment it requires advanced hardware So limited flexibility for power users because usually, often you've seen that many hackers or probably system engineers, they prefer to do command line based interaction rather than GUI because that is faster, that is safer.

So complex design may overwhelm user in the advanced applications. Design is complex. Again, too many functionalities, too many functions are running, and it may simply overwhelm your system as well. And often that's why you've seen that many web pages are unresponsive after some time. So command line interface, so that is text-based.

And as I mentioned in previous slide, it's primarily for the experts. So like Linux terminal, Python interpreter, and since it's quite lightweight and operate quickly and ideal for repetitive and complex tasks, flexible and powerful for the expert users. But at the same time, it's too hard for the people novice user to learn. So a steep learning curve for the beginners. not visually appealing because of course it doesn't have more attractive interface that you can see and and often you may not also get a quite good sense of what is happening so not quite engaging error prone as it require exact sentence So for

example, if you're working on Linux system, in order to probably list all the files, you want to write LS. In that case, I mean, yeah, it has to be exactly LS and so on, and the corresponding what all the syntax are there.

If you don't write that, it say, I mean, I don't find that command. Another types of interfaces is natural user interface. So it's more natural human being so for example simple gesture voice or touch and so on that in natural way as a human we do for example Apple Siri you can speak to it Kinect gaming system you can play with it so the pros are intuitive due to reliance on the natural human actions accessible with hands-free option for example you can do with voice command, engaging and highly interactive for users. But one of the big cause that you can see that Melek precision for detailed task. Challenging in noisy or crowded environment.

For example, you're speaking with Alexa or Apple Siri. If it is noisy environment, you are in traffic, it may not able to understand you. So it requires advanced technology, increasing cost and so on. Another type of interface is tangible user interface, where basically you interact with physical objects. So for example, in this case, you can see the Sparrow robots in STEM education that is being used. So pros are encourage collaboration in group settings, combine physical and digital interaction effectively.

It enhances learning experiences, especially for the children. And the cons is it can be costly to produce and maintain, requires physical space for installation and use. And of course, scalability is often limited. Mobility is often limited. So next type of interface is virtual reality. So virtual reality is more immersive 3D environment where probably you feel you are in a virtual world.

and you are interacting with the probably virtual objects and so on, and you are part of that virtual world itself. So for example, Oculus Rift for gaming and training simulations. So pros are, provides full immersive and engaging experience, everything in 3D, ideal for simulation and training scenarios. For example, we might have seen that, for example, you are in the space itself and probably, fighting war with star wars facilitate experiential learning safe environment often there is a many difficult task or complex task or you can say dangerous task which often you don't want to go in while you are learning for example welding if you're doing welding again if you directly go into the learning to welding, it may basically you may hurt yourself. So what if you try in the virtual world itself? There is a lot of research is happening on how you can upskill the people especially for these complex tasks, dangerous tasks first through the virtual world and then go in the real world how to do so.

So the cons is high hardware cost, setup requirements and may cause motion sickness.

and fatigue. So limited application in everyday tasks. So next is augmented reality. So just like virtual reality, as a user, you will be fully in a virtual world. So in augmented reality is basically what happens, it overlays virtual element in the real world.

For example, if you're taking a picture of a real world, then probably you can insert some object there. So that is something Pokemon Go game is based on. And you'll be surprised to know that despite its very simple applications, but it has broken all the record when this game has been launched. And of course, the owner has made huge amount of money from this.

Similarly, Ikea, place app and so on. So Froze S combines real and virtual world seamlessly, versatile across fields like retail, healthcare, education and enhances real world interaction with overlays. Cons, again, heavily dependent on the hardware, and it will again cost your battery life. There is no free lunch. And extensive to the implemented scale may distract user from their surrounding sector.

And this is the way the evaluation, you can see that evolution of interfaces in SCI happens. So interfaces have evolved from simple text-based command to now immersive and the context-free systems now. So these are things you can see like started with like simple MS-DOS, after that Windows, MacOS. Now we have the web-based interfaces, Google Chrome, browsers and so on. We have natural user interfaces like Alexa, Google Assistant and so on, where you can interact with voice and so.

Immersive interfaces like VRAR and so on. For example, Oculus Rift is one of the popular VR headset and so that you can see. So real world examples are for GUI, simple Google Doc, you can perform different operations. So voice interface, Google Assistant, Amazon Alexa, Apple Siri, and all those things you can do. So for VR, Google Earth VR, Oculus, and all those we have discussed. AR, Snapchat filters, often you might have seen that many of you Basically, Instagram lovers, Snapchat lovers and so on.

So when you take any pictures, you can add the filters to it. For example, probably on your face, you make yourself cat, dog, princess, king, whatever. So touch interface, for example, iPhone's multi-touch gesture, pinch to zoom is very interesting thing which happened. So Alan Cooper said, no matter how beautiful, your interface should always serve a purpose. I mean, you can compromise a little with aesthetics, but at least it should basically serve the purpose it's supposed to do, helping the user accomplishing their goals.

So interaction type and interface styles. So if you see the difference, Interaction type is about a description of what user is doing when interacting with the system. For example,

instructing, talking, browsing, responding, and so on. At the same time, interface style is a kind of interface used to support these interactions.

For instance, command, menu base, gesture, query, pen, form filling, voice, graphic, etc., and so on. So this is the end of part 1. Thank you so much. See you in the next lecture.