

Human Computer Interaction (Hindi mein)

Professor Rajiv Ratn Shah

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

Conversational AI: Human-Centric Interaction through HCI and NLP

Lec44

[Sangeet] Namaskar main Aman main Dr. Rajiv ke saath kaam karta hoon unke kaafi saare products projects ke andar kai saare projects jaise ki hamara audio ka ek bahut bada platform hai jiske andar hum audio ke data sets aur in saari cheezon ke pe kaam karte hain. So aaj main aapko bataunga HCI aur uske conversation systems aur applications ke baare mein. Toh chaliye shuru karte hain. Toh is saari cheezon ke andar hum kya cheezein cover karenge? What is conversation system? Kya hote hain conversation system? Kaise use hote hain? Kya-kya cheezein hum kar sakte hain? Kaun se kaun se bots hain? Kya-kya prakaar ke hote hain woh? Aur ek uske baad ek demo bhi hum run karenge. Chaliye shuru karte hain. Ab definition ke baare mein. Ab iske andar aap dekhenge toh likha hua hai a computer system that can communicate and interact with human using natural language. Iska matlab kya hota hai ki ek computer system jo ki aapki natural language ke andar aapse baat kar sakta hai. Iske andar do cheezein jo hai humko especially note down karni hai. Ek computer system aur ek hamara natural language. Computer system kya hota hai? Jaise ki aap jaante hain ki computer ki ek apni khud ki bhasha hoti hai jo ki numbers ke andar hoti hai. Binary format ke andar one zero ke andar aur natural language kya hoti hai? Natural language hoti hai ki jo hum insaan aapas mein baat karte hain. Jaise ki main aapko abhi bata raha hoon Hindi ke andar English ke andar Tamil, Telugu kisi bhi conversation ke andar jo do insaan do human being ke beech mein hoti hai usko hum natural language bolte hain. Toh uske using that natural language hum kaise computer ko apni cheezein samjha sakte hain ki hey aap yeh karke dikhao mere ko. Toh iske kuch-kuch examples hote hain jaise ki chat bots, virtual assistants like Siri, Alexa, Google ya fir Google Home jo hum gharon ke andar use karte hain. Aajkal smart phone hai jaise ki hum Google Home ko kuch command dete hain ki hey Google can you please turn on the light. Ab Google jo hai woh aapki light jo hai woh switch on kar dega ya switch off kar dega. Voice assistance. Voice assistant hote hain ki kuch-kuch bots hote hain jo ki aapke ek specific cheezon ke upar trained hote hain. Agar aap usko bologe ki please aap mere liye calendar book kar do ya Uber book kar do ya isko call kar do toh woh aapko us directory mein jaakar us person ko call kar denge aapke behalf ke upar. Aapko bas apni natural language ke andar usko samjhana hoga. So ab iske andar aata hai ki what is essence of HCI aur uske key principles kya hote hain? Ab HCI ki essence jaise ki ismein likha hua hai HCI ensures that conversation system are user friendly. User friendly ka matlab kya hota hai ki ab main aapse Hindi mein baat kar raha hoon. Abhi agar mujhko mere jo bots hain ya fir

conversation systems hain woh agar Hindi samajh hi na paaye toh woh aadhi se zyada janta ke liye bekaar ho jayenge. Unko samajh hi nahi payenge ya unke unko help hi nahi kar payenge. Toh us cheez ka koi point nahi bachega. Toh isliye hum usko user friendly banate hain jiske andar multi language hoti hai. Multi language ka mera matlab hai ki Hindi ho sakti hai, English ho sakti hai, Tamil, Telugu and etc. So and so forth. Iske andar kuch bhi aa sakta hai. Koi bhi language jo ek user ko aur user friendly ka doosra matlab yeh hota hai ki aapke hisaab se cheezein zyada complex na banayi jaaye. Ek human ke hisaab se. Koi bhi ek normal janta normal banda usko use kar sake. Jaise ki maine aapko doosra bataya iske andar key principles kya hote hain? Iska hota hai user centered design. Jaisa ki maine aapko bataya tha ki user ke hisaab se particular thing design honi chahiye taaki ek normal janta usko use kar paaye. Doosra hota hai accessibility and inclusibility. Ab iske andar kya hota hai? Accessibility ka matlab kya hota hai? Ki aap jo hai bahut saari language bahut saari cheezein support kar pa rahe ho in normal day to day activity natural language ke andar aap kisi chatbot ko kisi conversation AI ko aap bata pa rahe ho ki mere ko yeh yeh cheezein karni hai aur woh aapke behalf pe woh cheezein kar pa raha hai. Ab iske andar kya hota hai ki hum ek chat bot ko agar kuch bole karne ko aur agar woh task perform kar pa raha hai woh different accent samajh pa raha hai. Accent se mera matlab hai ki ek Haryanvi bande ka accent, ek Delhi ke bande ka accent aur ek UP ke bande ka accent. Although teeno bolte Hindi hain. Lekin teeno ki jo accent hogi bolne ka Hindi ka jo accent hoga woh different-different hoga. Ab agar ek AI ka bot in teeno ki Hindi ko agar nahi samajh pa raha hai toh iska matlab ki woh kisi na kisi accessibility mein jo hai achha perform nahi kar pa raha hai. Toh humko yeh make sure karna padta hai ki hum ek aisa tool banayein jo ki robust ho. Ab iske andar key components kya-kya aate hain. Main aapko batata hoon ki key components kya hote hain natural language ke andar. Key components ke andar sabse pehle aata hai conversation voice to text. Voice to text ka mera matlab hai jo main abhi bol raha hoon jo mere munh se jo sentence aa rahe hain usko ek microphone ek jo bhi hamara device AI bot hota hai woh jahan se bhi voice capture kar raha hai wahan se woh voice capture kare aur uske baad us data ko ek textual format mein convert kare. Textual format ka mera matlab hai ki jo hum Word document ke andar PDF ke andar ya fir kisi PPT ke andar jo hota hai jaise ki yeh jo humne likha hua hai this is a word format jo bhi humne text iske andar likha hua hai yeh saara word format ke andar toh hamare bole hue words ko jab ek conversational bot usko text mein convert karta hai tab woh process karna kaafi easy ho jaata hai. Ab natural language ka mera matlab iske andar yeh hota hai ki natural language ke andar usko samajhna chahiye. Hindi, English, Tamil, Telugu, Kannada, Spanish koi bhi XYZ koi bhi language ho sakti hai uske andar. Agar woh conversation bot usko samajh pa raha hai toh usko text form mein convert kar payenge. Ab dialogue management teesra cheez hota hai. Dialogue management ka mera matlab hai ki agar kisi bot ko ya kisi person ko bhi agar hum ek specific format yaani ki word by word ya fir ek speaker one speaker two ke hisaab se agar hum kisi insaan ko bhi agar samjha nahi samjha payenge toh insaan ka jo hai perspective badal jaata hai. Toh isiliye humko LLM ko bhi AI models ko bhi aise hi samjhana padta hai ki uske andar human jo hai human one yeh bol raha hai. Human two yeh bol raha hai aur uske baad human two ka sentence itne hain aur human one ke sentences itne

hain. Toh humko yeh bot ko samjhana padta hai. Ek dialogue ke format mein dena padta hai. Textual format ke andar jab hum de rahe hain. Natural language generation. Ab natural language generation ka matlab yeh hota hai ki jaise ki maine abhi kuch command di apne AI tool ko. Usne kya kiya ki mere words ko pehle textual format mein convert kiya aur jo maine usko command di thi us command ke upar usne work kiya. Uske baad jo work karne ke baad uska output aaya woh jo hai humko usi language ke andar humko wapas return kar raha hai woh ki jaise ki main aapko abhi aane waale example mein dikhaunga main. Google ko bolunga ki Google can you please make a call to is banda ek x banda. Toh Google jo hai us bande ko call kar lega. Ab woh command kaise samajh pa raha hai? Mere jo words hain usko pehle woh textual format mein convert kar raha hai. Phir apni processing kar raha hai peeche ki user ne jo hai woh kya poocha hai? Mere ko kya jawaab dena hai ya fir kya command follow karni hai. Us hisaab se fir woh us cheez ko process kar raha hai. Output de raha hai humko. Ab chatbot ke jo use hote hain woh kaafi zyada jagah par use hote hain. Jaise ki aap dekh paate ho customer service chatbots providing 24/7 assistance. Aajkal aapne dekha hoga jab bhi aapke paas koi bank ki query aati hai toh bank ke andar kya hota hai ki aapko ek IVR system hota hai wahan par aapko kuch words ke andar aapko unko apni problem batani padti hai taaki woh IVR chatbot aapko specific us department ko call forward kar sake ya fir solution provide kar sake toh yeh jo hota hai woh kaise samajh paata hai? Aapko kuch batana padta hai ki mera credit card ke andar problem aa rahi hai. Mere bank account ke andar problem aa rahi hai. Mere ko FD mature karwani hai ya fir FD close karwani hai, open karwani hai. Toh woh aapko us concern department ko forward karta hai. Woh kaise samajh paata hai? Humne bola ki FD close. Ab woh isko as a textual lega. Usko apne LLM ke andar process karega aur ek output dega. Iske andar output kya hai ki jo FD ka department hai usko call transfer kar dega woh. Yeh saari cheezein humne khaali bol ke kari. Humne koi usko command nahi diya. Na mouse se na select karke na keyboard se na kisi aur cheez se. Toh yeh saari cheezein hamari voice se ho rahi hai. Ab virtual assistant hote hain. Virtual assistant jaise ki maine aapko bataya ki agar hum aap koi koi bhi aap mein se Google ka koi phone use kar raha hai. Google Pixel ya fir aajkal koi bhi Android koi Apple ka phone use kar raha hai toh uske andar inbuilt Siri ya fir Google Assistant aate hain. Agar main unko bolunga ki aap please mere liye ek appointment schedule kar do 4:00 baje ki suppose 4:00 pm ki. Toh woh jhat usi time calendar mein jaakar mera 4:00 baje ka ek schedule daal denge. Aur main unko bolunga ki 4:00 baje mere ko dentist ke saath jaana hai. Toh woh 4:00 baje dentist ko appointment ek schedule kar denge. Toh woh kaise hua? Woh aise hi hua ki humne usko command di aur usne process kiya. Then ek output return aaya. Output kya aata hai ki 4:00 baje ka ek dentist ka appointment book karna. Ab kuch-kuch hote hain educational tools. Educational tools ke andar aajkal jo hai ChatGPT ka ek bahut hi naya version aaya hai jiska hum jisko hum voice se bhi command de sakte hain. Ab woh voice se command kaise hota hai? Hum usko bole hain ki GPT can you please tell me about HCI. Humans inter design kya hota hai ya fir humans interaction kya hota hai? Toh woh aapko batayega ki us education ke beech mein ki kya hota hai asal mein woh. Aap usko kuch aur command doge ki GPT ya fir Google ya koi bhi chat bot ki yeh pen kaise work karta hai? Yeh monitor kaise work kar raha hai? Is monitor ki dimensions

kya hai? Toh calculate karke aapko ek output generate karke deta hai. Toh yeh hamara educational chat box ke andar aa jaata hai. Mental health chat box again yeh aapko ek emotion provide karta hai. Emotion ka matlab hota hai ki jaise ki hum apne aapas mein doston se baat karte hain. Main apne dost ko batata hoon ki hey Raj yeh problem mere ko ho gayi. Ab jo mera friend Raj hai woh bolega ki Aman don't worry about it. Zyada mat socho ya fir koi unko solution dega. Toh yeh jo uska jo human ka jo behavior hai yeh chatbot usko generate karne ka try karta hai. Exactly same. Ab kya hota hai ki woh hum chatbot se maine bola ki hey Google mere ko suppose mere ghar ke andar XYZ mere mere brother ko fever ho gaya hai. Toh chatbot jo hai woh bolega ki don't worry Aman he will be cool. Thode dino mein theek ho jayega woh ya aap usko aisi-aisi medication do. Toh mere ko consolidate karne ka try karta hai. Mere ko sadbhavna dene ka try karta hai ki aap please tension mat lo. Toh yahan par hum mental health ke andar yeh chatbot jo hai aajkal use kar rahe hain. Ab conversation system ke andar bahut saare conversation system aate hain. Jaise ki maine aapko pehle bataya tha ki ek yeh Google ka Google Home hai. Yeh Google Android phones ke andar jo aata hai Google Assistant hai ya fir hamara Echo Dot hai jo ki Alexa Amazon provide karta hai. Ya fir jaisa ki maine aapko bataya tha ChatGPT ka ek aajkal jo version aa raha hai jo aapko batata hai ki aapki conversation ke andar normally day to day jo aap natural language ke andar baat kar rahe ho uske andar aapki query leke aapko ek answer provide karta hai just like a human being but uske paas infinite knowledge hai. Ab main chahunga ki aapko aapki further jitne bhi yeh readings hain aap iske baad ek baari padho jaake yeh saari ki saari aapko important hai aur aapko ek overview de dengi ki cheezein kaise chalti hain. Is conversation AI ke andar jo maine aapko bataya main aapko summary de deta hoon ki iske andar hum kya-kya cheezein process karte hain ya fir ek computer kaise process karta hai. Ek human jo hai pehle usko ek ek natural language ke andar ek command dega. Command kuch bhi ho sakti hai. Suppose maine usko bola ki ek meeting schedule kar do. Ab woh meeting schedule karega. Meeting kab ki karni hai? 4:00 baje ki karni hai. Ab kiske saath karni hai? Suppose dentist ke saath hi karni hai. Purana example hai but ya. Ab yeh jo command hai yeh command AI chat bot hamara text format mein convert karega. Textual format mein convert ho gaya. Jaise hi hamara textual format mein convert hone ke baad isko process karega. Isko process karne ka jo hota hai woh machine language ke andar process hota hai. Machine language ke andar woh process karega toh woh binary format mein hota hai. One zero jitne bhi codes hote hain uske andar isko banayega aur usko process karega. Then usse kya hoga ki machine samajh payegi ki ek user ne ek human ne usko bola kya hai karne. Jaise hi woh isko process karega ki meeting setup karni hai 4:00 baje ki dentist ke saath woh ek aapko output generate karke de dega ki hey Hi Aman maine aapki 4:00 baje ki meeting setup kar di hai with dentist. Yeh uska output rahega. Yeh poora ka poora process jo hai woh ek audio chat bot ka kaam aata hai. Iske andar aur bhi kai saari cheezein hum add kar sakte hain. Kai saari cheezon ka mera matlab hai ki language detection maybe jaise ki abhi main aapse baat kar raha hoon. Yeh poori shuddh Hindi nahi hai. Yeh English bhi hai beech mein. Toh hum isko Hinglish mein bolte hain. Ek normal conversation ke andar isko Hinglish bolte hain. Toh ab Hinglish ko process kar raha hai woh language ko. Maybe Spanish ho, maybe Tamil ho, maybe Telugu ho, kuch bhi

language ho sakti hai. Toh language specify bhi kar sakta hai. Yeh wala chat bot accent specify bhi kar sakta hai. Maybe tone bhi specify kar sakta hai. Emotion bhi specify kar sakta hai kabhi-kabhi. So ab main aapko ek demo dikhaunga Alexa ke saath. Unfortunately mere paas Alexa toh nahi hai magar main aapko apne phone ke andar Google Assistant ka prayog karke aapko dikha sakta hoon ki cheezein kaise chalti hai. Toh aaiye ek meeting setup karte hain Google ke through ki Google kaise meri natural language ke andar main bataunga usko aur woh ek meeting setup kar sakta hai. Toh chaliye shuru karte hain. Okay Google kya aap mere liye Okay Google kya aap mere liye 4 baje ki meeting setup kar sakte hain dentist ke saath Aapke liye shaam 4:00 baje ki dentist ke saath appointment schedule kar di gayi hai. Toh jaisa ki aapne dekha maine Google ko bataya ki 4:00 baje mere ko dentist ke paas jaana hai. Aap mera ek baari ek meeting setup kar do. Toh unhone 4:00 baje ki exactly mere liye ek calendar Google calendar ke upar meeting setup kar di hai. Toh aise hi ek conversation bot jo hai mere ko kai saari cheezon ka answer de sakte hain. Kai saari cheezein bata sakte hain. Toh ek aur cheez hum kar sakte hain. Maybe agar main kahin baitha hoon aur main kahin bore ho bore ho raha hoon bahut zyada. Mere ko cheezein seekhni hai ya fir mere ko ek joke sunna hai toh kya karunga main? Main Google ko bolunga ki Google mere ko ek joke sunao please. Toh Google dekhte hain kya reply karta hai. Okay Google Okay Google Kya aap mujhe ek joke suna sakte hain? Main joke nahi suna sakta lekin main aapko kisi aur tarah se hasne ki koshish kar sakta hoon. Okay Google mujhe aaj ki news batao. Aaj ki kuch mukhya khabarein yahan di gayi hain. Kartavya Bhavan ka udghatan Pradhan Mantri Narendra Modi ne Delhi mein navnirmit Kartavya Bhavan 3 ka udghatan kiya jo kai sarkari mantralayon aur vibhagon ka naya thikana hoga. Rahul Gandhi ko jamanat. Jharkhand ki Chaibasa court ne Amit Shah par tippani ke maamle mein Rahul Gandhi ko jamanat de di hai. Chennai Grand Masters tournament sthagit. Chennai Grand Masters tournament ko ek din ke liye sthagit kar diya gaya hai. Toh jaisa ki aapne dekha maine Google se poocha ki mere ko ek news batao. Aaj ke headlines batao. Joke sunane ka maine try kiya. Toh Google ne kuch respond kiya. Ab woh respond kaise kiya? Usne uske apne intelligence ke hisaab se process karke cheezein humko ek output di. News ke andar jo bhi latest abhi news chal rahi hai usne woh bataya. Jokes ke andar uske paas kuch nahi hoga toh maybe usne nahi bataya. Maine calendar invite karne ke liye ek meeting setup karne ke liye usko bola toh usne calendar ka ek invite daal diya 4:00 baje ka. Agar main usko bolun ki kisi ko phone kare toh Google phone bhi karega. Toh aise hi jo hai woh ek poora human center ke around ek audio ka ek chatbot aise hi kaam karta hai ki aapki natural language ko apne computer ke andar pehle text ke andar format convert karta hai. Phir computer ke language ke andar usko convert karta hai aur phir usko process karke natural language ke andar hi aapko output deta hai. Aasha hai ki aapko saari cheezein samajh aayi hongi jo maine aaj aapko samjhane samjhaya. Agar aapko kuch aur kisi type ki problem hai toh maine jo aapko links pehle ki isse purani wali slide ke andar diye the aap usko refer kar sakte hain. Cheezein aise hi work karti hain ki aap apni machine ko ek natural language ke andar batate hain. Woh aapka text mein convert karta hai. Phir machine language ke andar convert karta hai. Phir apne intelligence ke hisaab se process karke aapko natural language

ke andar hi output deta hai. Toh yahi hum iske andar samjhane ka prayas kar rahe the aapko. Toh aasha hai ki aapko sab kuch samajh aaya hoga. Dhanyavaad. [Sangeet]