

Om maskiner kunde tala

Staffan Larsson, forskarassistent i lingvistik

Ludwig Wittgenstein skriver i *Filosofiska Undersökningar* (1953): ”Om ett lejon kunde tala, så skulle vi inte förstå det.” Varför inte? Enligt en tolkning vill Wittgenstein med detta göra oss uppmärksamma på att språket först och främst är ett socialt fenomen som uppkommer i samspelet mellan människor. Då det mänskliga språket uppkommit i interaktionen mellan just människor, snarare än mellan människor och lejon, så är det också tillgängligt bara för människor. Ett språk som uppkommit mellan lejon (om de nu kunde tala) skulle på samma sätt vara otillgängligt för oss människor. Dessutom finns förstås avgörande skillnader mellan människors och lejons fysiologi, kognition och levnadsmiljöer som ger olika grundförutsättningar för de sociala interaktioner som människor respektive lejon kan delta i. Sammantaget resulterar dessa skillnader i att människors och lejons livsvärldar, och därmed deras språk, skulle vara mycket olika.

Inom språkteknologin utvecklas metoder som möjliggör kommunikation mellan människor och maskiner med hjälp av naturliga (mänskliga) språk. Vid en första anblick verkar detta motsäga Wittgensteins påpekande; det mänskliga språket har inte uppkommit i interaktion mellan människor och maskiner och är således inte tillgängligt för maskiner – de kan inte förstå vad vi talar om. I denna artikel vill jag undersöka en samling tankegångar och idéer från olika håll i ett försök att kasta ljus på frågan om maskiners (o)möjlighet att förstå och använda mänskliga språk.

Hjärnans maskinkod?

I Neal Stephensons science fiction-roman *Snow Crash* (1992) visar det sig att ett gammalt sumeriskt språk fungerar som ett ”hjärnans maskinkod” som gör det möjligt att infektera människor med ett ”mentalt virus” och därmed fullständigt kontrollera deras beteende och tankar. Analogin med datorprogrammering och datorvirus är tydlig. Givet antagandet att hjärnan

i princip fungerar på samma sätt som en dator följer att det också måste finnas en hjärnans motsvarighet till datorernas maskinkod, som möjliggör total kontroll över maskinens tillstånd. Men är en hjärnans maskinkod verkligen möjlig?

Den så kallade *simulationsemantiken* är en kognitivt orienterad teori om språklig betydelse som bygger på idén att vi, när vi förstår språk, exekverar en *mental simulation*; vi skapar oss (enkelt uttryckt) en mental bild av det språkliga innehållet. I en simulation aktiveras samma hjärnstrukturer som skulle ha använts om vi faktiskt befunnit oss i den verkliga situation som motsvarar den (fiktiva) simulationen. Varje ord är kopplat till en simulator (ordets ”betydelse”) som aktiveras då vi känner igen ordet, och språket gör det således möjligt att kontrollera simulationen. En roman kan ses som ett program skrivet i naturligt språk som, när det körs, ger upphov till en simulation i läsarens medvetande. Enligt detta synsätt kan mänskliga språk bäst förstås som ett hjärnans programmeringsspråk – skillnaden mot datorernas programmeringsspråk är helt enkelt att mänskliga språk exekveras på människohjärnor istället för på datorer. Finns det då någon avgörande skillnad mellan en människa och en dator vad gäller förmågan till språkförståelse?

Komplicerat eller komplext?

Inom det tvärvetenskapliga studiet av komplexa system görs en distinktion mellan *komplicerade* och *komplexa* system. Ett komplicerat system, till exempel ett flygplan eller en dator, kan bestå av ett stort antal komponenter, men kan alltid ges en uttömmande beskrivning i termer av de enskilda komponenterna. En nyckelegenskap hos komplicerade (men inte komplexa) system är att deras beteende är begripligt, beskrivbart och förutsägbart. Ett komplext system kan däremot inte förstås enbart i termer av sina komponenter, då det inte är komponenterna själva (de kan vara mycket enkla) utan den rika interaktionen mellan dem som är det primära. I komplicerade system finns ofta en komponent som kontrollerar de övriga komponenterna (till exempel centralprocessorn i en dator), men i komplexa system har ingen komponent kontroll över de övriga komponenterna. Exempel på komplexa system är klimatsystem, ekologiska system, den mänskliga hjärnan, (de flesta) sociala system samt mänskliga språk.

En nyckelegenskap hos komplexa system är möjligheten till *självorganisering*, det vill säga förmågan att utveckla och anpassa sin interna struktur för att manipulera och hantera förändringar i sin omgivning. Ett självorganiserande systems struktur är inte designat och planerat utan uppstår i interaktionen mellan dess komponenter. Att nya fenomen kan uppstå spontant gör också att komplexa system inte är förutsägbara, i motsats till de enbart komplicerade systemen. En annan konsekvens av förmågan till självorganisering är att komplexa system är föränderliga och anpassningsbara och kan hantera förändringar i omgivningen. Komplicerade system, däremot, är ofta statiska på så sätt att det är svårt (dock inte omöjligt) att förändra dem, och sådana förändringar kräver ofta medveten och noggrann planering av en designer eller konstruktör (eller programmerare).

Språk och kod

I *My Mother was a Computer* (2006) ställer N. Katherine Hayles frågan vad den grundläggande skillnaden är mellan språk (mänskliga språk) och kod (i betydelsen programmeringsspråk). Ett svar är att språk är avsedda att förstås enbart av människor, medan programmeringsspråk måste förstås av både datorer och människor (för att det ska vara möjligt för en människa att programmera en dator). Programspråkskonstruktörer lägger ner mycket möda på att göra programmeringsspråken mer begripliga för människor; ett exempel på en framgångsrik innovation på detta område är de så kallade "objektorienterade" programmeringsspråken såsom Java.

Hayles utvecklar utifrån Saussure och Derrida en teoretisk ansats till att förstå betydelse hos programmeringsspråk i kontrast till betydelse hos mänskliga språk. Jag kommer här inte att närmare gå in på denna ansats utan istället fokusera på ett annat sätt att förstå betydelsen hos datorkod. Enligt denna alternativa ansats, som antyds men inte utvecklas av Hayles, kan en maskins beteende beskrivas på flera olika nivåer: som elektriska spänningsförändringar i datorns kretsar; som ett flöde av ettor och nollor; som exekverandet av ett programmeringsspråk som till exempel Java; eller utifrån ett grafiskt gränssnitt på datorns skärm. De olika beskrivningsnivåerna erbjuder olika sätt att tala om, tänka på, och kontrollera datorns beteende. Eftersom datorn inte är komplex, utan bara komplicerad, är det i princip möjligt att ge en komplett förklaring av datorns beteende enbart utifrån fysikens lagar. En sådan förklaring skulle dock vara ohanterlig för

en människa och därför använder vi oss hellre av de ”högre” beskrivningsnivåerna när vi pratar och tänker om datorns beteende.

Hayles påpekar att programmeringsspråkens kraftfullhet ligger i just att de både är sätt att tala om och tänka på datorns beteende (dvs. begripligt för människor) och att koden har direkta fysikaliska och kausala effekter på datorns beteende som gör det möjligt för oss att kontrollera detta beteende in i minsta detalj. Vi kan förtydliga genom att tillägga att detta är möjligt just tack vare att datorn är just ett komplicerat system snarare än ett komplext sådant. Ett komplext system tillåter helt enkelt inte den typ av strikt kontroll som är grunden för traditionell datorprogrammering.

Både mänskliga språk och programmeringsspråk kan ses som sätt att påverka beteendet hos andra (människor eller maskiner) genom användandet av språkliga yttranden (talade, skrivna eller programmerade). Enligt Hayles är yttranden i mänskliga språk knutna till det yttre beteendet genom komplexa kedjor av orsak och effekt i den mänskliga organismen. Utifrån kontrasten mellan komplicerade och komplexa system kan vi kontrastera detta mot de ”bara” komplicerade orsakssammanhangen mellan programkod och iakttagbart beteende i en dator.

Människospråk och datorkod skiljer sig även med avseende på de sociala strukturer som reglerar deras utveckling. Både naturliga språk och programmeringsspråk förändras över tid, men på radikalt olika sätt. Programmeringsspråk kontrolleras normalt av en liten grupp utvecklare som har mer eller mindre fullständig makt över programmeringsspråket. Mänskliga språk, däremot, utvecklas genom människors dagliga interaktioner där vissa kanske har mer makt än andra men där ingen kan förhindra att språket förändras.

Om människor börjar använda ett ord på ett nytt sätt (till exempel ”surfa” för att beteckna Internetanvändning) så kommer ordets betydelse att förändras. Om däremot en eller flera programmerare försöker använda ett ord i ett programmeringsspråk på ett annat sätt än vad som avsetts av programspråkets utvecklare, så blir resultatet inte att språket förändras utan att programmen inte fungerar. Även detta är en konsekvens av skillnaden mellan komplexa och komplicerade system. Människors hjärnor är komplexa, självorganiserande och därmed flexibla system som kan förstå och anpassa sig till språkliga innovationer. Datorer däremot är komplicerade och därmed rigida strukturer som inte kan anpassa sig på detta sätt.

Komplexa maskiner, komplicerade människor

Vi kan nu ge ett svar på frågan om huruvida det kan finnas en "hjärnans maskinkod" som skulle göra det möjligt att kontrollera en människas beteende i detalj. Svaret är helt enkelt: Nej. Maskinkodens förmåga att direkt kontrollera beteende kommer just av att den fungerar i ett komplicerat men förutsägbart system. Hjärnan är ett komplext system där sådan kontroll inte är möjlig, och därför är en hjärnans maskinkod inte möjlig.

Samtidigt är denna gränsdragning inte absolut. På senare år har forskare konstruerat datorprogram som kan lära sig rudimentära språk på ett sätt som påminner om mänsklig språkinläring, och genom att låta flera sådana program interagera i ett socialt nätverk kan nya språkliga konventioner uppstå genom självorganisering. Dessa experiment ligger oerhört långt från den mänskliga interaktionens rikedom, och datorernas radikalt människofrämmande fysiologiska, kognitiva och ekologiska förutsättningar gör att det verkar osannolikt att datorer (av den typ vi har idag) någonsin skulle kunna närma sig människans förmåga att förstå och använda mänskliga språk.

Lika väl som det finns komplexa maskiner, så kan man också hitta exempel på försök att reducera bort människors komplexitet och göra om dem till förvisso komplicerade men ändå förutsägbara komponenter i rigida och centraliserade sociala system. Byråkratiska och kanske framför allt militära organisationer är exempel på detta.

Det är nu dags att återvända till vår ursprungliga fråga. Hur kan språkteknologin vara möjlig, givet det som sagts ovan? Hur kan datorer, som inte är komplexa och dynamiska och som därmed inte har förmågan att ingå i komplexa sociala interaktioner och system, ändå fås att interagera med människor med hjälp av mänskliga språk? Ett kanske nedslående svar är att det bara kan ske om (de flexibla) människorna anpassar sig till (de oflexibla) maskinerna. Det "naturliga språk" som datorerna använder är i själva verket bara datorkod utklätt till mänskligt språk. Även språkteknologin kan således anklagas för att tvinga människor att reducera bort det komplexa till förmån för det enbart komplicerade.

Å andra sidan finns det en stor mängd vardagliga mänskliga aktiviteter (till exempel biljettbokning) som mer eller mindre stelnat i sin form och som därmed blivit tillräckligt förutsägbara för att datorer ska kunna delta

i dem. I dessa aktiviteter förefaller det möjligt att introducera språkteknologiska hjälpmedel utan negativa effekter. Detta gäller även interaktioner mellan människa och maskin som redan nu sker helt på maskinens villkor, till exempel handhavandet av mobiltelefoner. Om det synsätt som utvecklats här är giltigt, så har det konsekvenser för när vi kan förvänta oss positiva effekter av att introducera språkteknologiska hjälpmedel.

Cybernetiska språk?

Slutligen: i den mån datorer kan tilldelas den komplexitet, flexibilitet, inlärning och självorganisering som krävs för att kunna delta i komplexa sociala interaktioner så kan man spekulera om möjligheten för människor och maskiner att ingå i nya sociala nätverk där nya språk kan uppstå, utifrån och trots våra skilda förutsättningar. Lejonen kan vi lämna därhän; de skulle knappast vara särskilt intresserade av att tala med oss människor även om de kunde. Men för maskinerna, som vi byggt för att tjäna våra egna syften, är situationen en annan. Resultatet skulle inte bli att datorerna förstår och använder *mänskliga* språk, utan att det i interaktionen mellan människor och maskiner uppstår nya människa-maskin-språk (vi kanske kan kalla dem *cybernetiska språk*) som kan förstås och användas av både oss och dem. Det är en öppen fråga om detta är möjligt, kanske likaså om det är önskvärt.