

# Gränser och gränssnitt mellan människa och maskin

Staffan Larsson, fil.dr i allmän språkvetenskap

Hur ska vi människor relatera till våra maskiner? Är människan en sorts maskin? Kan maskiner vara mänskliga? Dylika frågor framstår för vissa av oss som fascinerande och intressanta, medan andra ställer sig mer oförstående till dem. Varför är dessa frågor intressanta (för vissa)? Varifrån kommer människors fascination för maskiner och speciellt då ”mänskliga” maskiner? För många är maskiner, och då speciellt datorer, det längsta man kan komma från mänsklighet; de är hårda och kalla, vi är mjuka och varma; de är rationella, vi är känslomässiga. Givet dessa skillnader mellan människor och datorer, varför vill vissa av oss spegla sig i datorn medan andra vänder sig bort med en rysning (eller en gäspning)?

Man kan kanske spekulera i att människor känner sig besläktade med maskiner i olika grad. Eller: att människan-som-maskin erbjuder en förklaringsmodell som avlastar människan en del av den (för vissa mer än andra) tunga bördan av att vara särskild – att ha en fri vilja, att vara medveten, att vara etiskt ansvarig. En delvis motsatt tänkbar målsättning är att urskilja det specifikt mänskliga, det som gör oss till etiska varelser. I detta fall är syftet att visa att vi *inte* är maskiner.

Vad, om något, utmärker det unikt mänskliga? Är det tänkbart att en maskin skulle kunna ha dessa ”unikt mänskliga” egenskaper, och vad skulle i så fall krävas? Enligt Heidegger är det som utmärker en människa inte det hon normalt tror att det är; att hon är gladlynt eller butter, morgontrött eller morgonpig, känslösam eller känslökall. Alla sådana uppfattningar är bara påhitt som tjänar till att dölja för människan att hon inte har någon essens annat än denna: att hon tolkar sig själv och sin värld.

Begreppet ”gräns” kan tolkas på åtminstone två sätt i detta sammanhang. För det första, den kontaktyta mellan människa och maskin i vilken kommunikation äger rum; ”gränssnittet”. Gränssnittsfrågan är i grunden en praktisk fråga om hur vi relaterar till maskiner i vardagen. För det

andra, det som gör människan till människa och maskinen till maskin (om det nu skulle vara så enkelt). Detta är en mer teoretisk/filosofisk fråga om hur människor och maskiner är relaterade i princip. Nedan kommer jag att relatera båda dessa frågor till ett av grundantagandena inom språkteknologi och kognitionsvetenskap: att mänsklig språkförmåga och intelligens kan simuleras i datorer genom att dessa datorer programmeras med modeller av hur mänskligt tänkande fungerar.

### *Gränssnitt*

Var går gränsen mellan människa och maskin, och (hur) kan språket användas för att överbrygga denna gräns? I dagens samhälle omges vi av mer eller mindre (oftast mindre) intelligenta maskiner som vi dagligen tvingas kommunicera med. Denna kommunikation är dock mer bestämd av den teknikorienterade ingenjörskonstens begränsningar, än av människors önskningar och behov. Problemet med detta är inte enbart att det tar lång tid att lära sig navigera den nya mobilens menysystem. I förlängningen finns risken att teknik, som skulle kunna underlätta för alla, förbehålls de teknikintresserade. Vi löper risken att få ett samhälle där medborgarens möjlighet att delta i samhället begränsas av hennes tekniska kunnande. För att undvika detta måste tekniken göras tillgänglig för alla; vi behöver ett mer demokratiskt sätt att kommunicera med maskiner. Det naturliga mänskliga språket verkar vara det självklara valet.

Kan då maskiner fås att kommunicera på människors villkor? Om du provat på t.ex. SJ:s talstyrda tågtidsinformation (020 757575) har du kanske börjat ana att det inte är så enkelt. Än så länge ligger tekniken i dessa samtalande datorer ("dialogsystem") oerhört långt efter människans språkförmåga. En orsak till detta är att de nu existerande dialogsystemen fortfarande utgår från maskinernas förutsättningar. Vad vi istället behöver göra är att bygga dialogsystem baserade på den mänskliga språkanvändningsförmågan.

Om detta ska lyckas måste vi först förstå hur människor använder språk. Wittgenstein påvisade att språklig aktivitet alltid är förankrad i annan mänsklig aktivitet; detta innebär att vi för att förstå språkanvändningen i förlängningen också måste förstå mänsklig aktivitet i allmänhet. Användningen av datorer i språk- och kognitionsforskningen har två relaterade syften: att förstå människan, samt att underlätta för henne. Båda dessa syften gagnas av försök att göra maskinerna mer lika människan. För det första,

genom att modeller av mänsklig aktivitet kan programmeras in i en dator och sedan testas, utvärderas, och utvecklas vidare. För det andra, genom att datorer som beter sig mer som människor kommer att underlätta kommunikationen mellan människa och maskin, till människans fördel.

### Gränser

För att en teori om mänsklig språkanvändning ska kunna stoppas in i en dator måste den först formuleras på ett sätt som är begripligt för datorn – det måste formaliseras. En formaliserad teori kan också betraktas som en modell av sitt objekt; om denna modell körs som ett program i en dator kan vi tala om en simulation.

Om nu en maskin kan kommunicera på ett naturligt sätt med hjälp av mänskligt språk, vad säger detta om den formella modell som ligger till grund för maskinens konstruktion? Att modellen åtminstone lyckas fånga en del av det iakttagbara mänskliga språkliga beteendet. Något mer?

För att få ett slut på ändlösa och meningslösa diskussioner om medvetandets natur och dess möjliga förekomst i maskiner, föreslog den engelske matematikern Alan Turing ett test för att avgöra frågan om huruvida en maskin kan betraktas som intelligent och medveten. I en mycket förenklad version går testet ut på att en dator ska lyckas efterlikna en människas språkliga kommunikation så väl att en annan människa inte kan avslöja den som en maskin. Om datorn uppnår detta bör vi betrakta den som intelligent och medveten.

Mot detta kan man hävda att även om maskinen kan kommunicera likt en människa, så betyder detta inte att människan fungerar så som modellen beskriver. Modellen kan kanske fånga en aspekt av människans beteende, men den kan aldrig *vara* människa. Människan kanske fungerar på ett annat sätt "inuti", även om det iakttagbara beteendet är detsamma. Till exempel kanske det som är explicit formulerat eller programmerat i modellen/programmet bara finns implicit i en samlig reflexartade beteenden som utgör en del av människans beteendemönster.

Det är i sammanhanget relevant att diskutera *hur* väl maskinen kan imitera en människa. Om vi tänker oss ett dialogsystem liknande SJ:s men något mer sofistikerat så förefaller det fortfarande vara möjligt att dess beteende baseras på en felaktig bild av hur människan fungerar; likheten är alltför skissartad för att vara övertygande. Men hur skulle vi uppleva ett dialogsystem som *verkligen* betedde sig mänskligt, in i minsta detalj, i minsta

hummande, minsta stammande, minsta skiftning i röstläget? Skulle vi tillskrivna detta system ett själsliv, eller åtminstone någon form av liv, medvetande, intelligens? Skulle vi kunna låta bli?

Jag misstänker att det är svårt, men inte alltför svårt, att bygga en regeluppsättning som kan förklara hur den mänskliga språkförmågan fungerar *när den fungerar utan hinder*; däremot är det svårare att förklara hur det går till när den knakar i fogarna eller går sönder. Med andra ord, det är svårt att göra en modell som misslyckas på samma sätt som människan. Kanske detta kan vara ett sätt att göra mer realistiska modeller av mänskligt tänkande och beteende.

Hur långt kan detta gå? Finns det någon principiell gräns bortom vilken maskinerna aldrig kan nå, något som utgör ett unikt mänskligt tillstånd? Hur skulle en mänsklig maskin se ut? Som våra datorer, eller radikalt annorlunda? Hur skulle den se ut inuti? Hur skulle den vara programmerad?

Är det överhuvudtaget meningsfullt att försöka efterhärma *alla* mänskliga egenskaper? Vill vi ha datorer som gör sig skyldiga till freudianska felsägningar? Här ser vi en skillnad mellan det teoretiska anslaget (undersökandet av *gränsen*), som vill förstå människan genom byggandet av formella modeller, och det praktiska (*gränssnittet*), som vill förbättra datorernas förmåga att kommunicera. För den senare finns det gränser bortom vilka det inte är meningsfullt att röra sig; för den förra är fältet fritt att söka efterhärma de mest meningslösa (enligt Freud: meningsfulla!) av mänskliga göranden och säganden.

Heidegger hävdar att vi i vårt normala vardagliga handlande inte är medvetna. Så mycket bättre för kognitionsvetaren: vi kan hoppas modellera de centrala mekanismerna i människans beteende utan att behöva ta i den trassliga medvetandefrågan.

### *Anledning*

Utmaningen för mig som humanist och språkteknolog är att hitta ett sätt att betrakta det jag gör som meningsfullt, som möjligt, som upplyftande. Mina trötta argument från B-kursen i teoretisk filosofi om att människor i viss mening är maskiner klingar allt tommare; vad spelar det för roll att människor *i princip* och *i viss mening* är maskiner? Den intressanta frågan är: hur ser området ut mellan å ena sidan de närmast triviala imitationerna av mänskligt beteende som nuvarande dialogsystem producerar, och

å andra sidan de avgrunder av mysterier som den mänskliga existensen härbärgerar? Eller enklare: hur sjutton fungerar egentligen det mänskliga psyket? Hur hanterar jag min omvärld? Finns det något i detta som kan fångas på ett meningsfullt sätt i en formell modell implementerad i en dator? Tvärvetenskaplighet är ett måste i denna undersökning; psykologi, filosofi, etnologi, vad som helst. Men även mixtrandet med detaljer i de program som får min dator på kontoret att ha enkla små dialoger om enkla ting.

Finns det inom humaniora utrymme för en öppen och sökande diskussion om människors och maskiners gemensamma och skilda gränser som inte är vare sig teknofobisk eller blind för teknikens risker och fallgropar? För språkteknologen, och egentligen för varje modern humanist, är en sådan diskussion nödvändig.

