

## Under strecket

## Sedan 1918

# Är vi tillräckligt smarta för artificiell intelligens?

Spekulationer om hur artificiell intelligens kommer att utvecklas antar ofta närmast teologiska dimensioner. I en ny bok föreslår den svensk-amerikanske kosmologen Max Tegmark i stället att AI bör betraktas som ett nytt stadium av mänskligt liv, och därför bäst kan förstås ur ett biologiskt perspektiv.

För ett tiotal år sedan sökte jag upp John McCarthy, en av portalgestalterna i den tidiga forskningen om artificiell intelligens. Tillsammans med Marvin Minsky, Herbert Simon, Allan Newell och en handfull andra forskare hade John McCarthy i mitten av 1950-talet lagt grunden för vad som länge såg ut att vara ett projekt med mer fantastik än egentlig forskning: att utveckla intelligenta datorer. Nu var läget emellertid ganska dystert. Det som forskarna själva kallade "AI-vintern", det vill säga en period präglad av krympande forskningsanslag och minskat intresse för frågor om artificiell intelligens, var fortfarande kännbar.

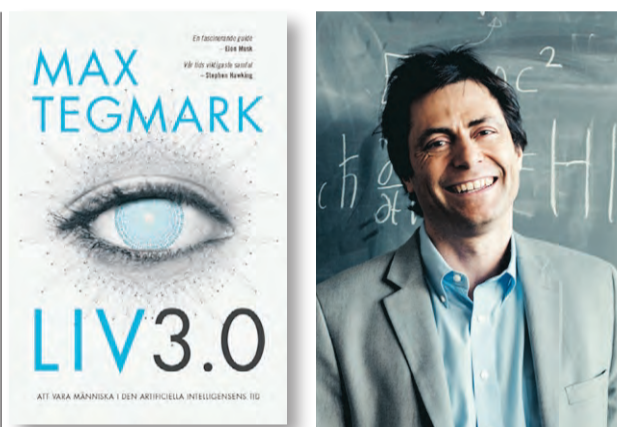
När jag sökte upp McCarthy var det för att diskutera detta misslyckande. Min första fråga till honom var om han inte var besviken. McCarthys buskiga ögonbryn flög upp i förvåning. Absolut inte, förklarade han, och tillade att det ju var en särdeles dum fråga. Visste jag inte att 50 år efter att Gregor Mendel upptäckt ärftligheten så hade genetiken nästan glömts bort, men att de därpå följande 50 åren gett oss sekvenseringen av det mänskliga genomet? Utvecklingen, förklarade han, kommer igång först ett halvt sekel efter upptäckten. "Så är det alltid med stora genombrott."

McCarthy gick bort 2011, men han hann ändå märka hur det fält han hade varit med om att grunda snabbt gick en ny vår till mötes. Den del av forskningen som möjliggjorde detta var ironiskt nog en som han själv avfärdade som ett sidospår, men fältet exploderade och är idag ett av de hetaste forskningsområdena i världen. Framstegen kommer slag i slag och vi står nu inför nästa skede i den tekniska utvecklingen. Där tekniken tidigare var ett verktyg som förstärkte vår fysiska förmåga, och som kunde automatisera arbetet i industrialismens tidevarv, har tekniken idag kommit till en punkt där den inte bara kan automatisera utan också reproducera vår kognitiva förmåga: lösa problem, utforska nya områden och lära sig nya uppgifter.

Kanske är det vi ser också början på en ny livsform? Det är åtminstone grundtesen i den svensk-amerikanske kosmologen Max Tegmarks bok "Liv 3.0 – Att vara mänskliga i den artificiella intelligensens tid" (övers: Helena Sjöstrand Svinn och Gösta Svinn; Volante). Tegmark tecknar ett schema där livet börjar som enkelt biologiskt liv med förmåga till reproduktion, utvecklas så att det kan förändra sin egen mjukvara genom kultur och sedan kan förändra också sin hårdvara genom teknologiska landvinningar. Detta tredje stadium är nytt, och vi måste – menar Tegmark – börja förhålla oss till det på ett mer medvetet, reflekterat sätt.

Tegmarks bok tjänar detta syfte väl. Det är en öppen och inkluderande bok, som även om den inte skyr tekniska detaljer noga förklarar och utvecklar dem så att även en läsare utan bakgrund i datavetenskap eller fysik kan följa med. Det är på sätt och vis en perfekt bok för den som vill förstå de samhälleliga utmaningar som möter oss i en värld där vi lyckats skapa generell artificiell intelligens.

Just ordet "generell" är viktigt här. Vi har idag lyckats ta fram mjukvara som kan utföra en mängd olika uppgifter bättre än människor: schack, vissa medicinska bedömningar, bilkörning – allt detta kan idag utföras med hjälp av specialiserade tillämpningar av artificiell intelligens. Samtidigt är denna intelligens låst till sina uppgifter. En mjukvara som kör bil kan inte automatiskt spela schack, och vice versa. Jakten på en generell artificiell intelligens är jakten på en mer mänsklig intelligens som



Max Tegmark är professor i fysik vid MIT. I kväll håller han en öppen föreläsning på Kulturhuset Stadsteatern i Stockholm. Foto: Anders Lindén

också förmår skifta mellan olika uppgiftsdomäner. Det är egentligen ett artificiellt, hypersnabbt lärande – en intelligens i rörelse.

Anta nu att vi lyckas med detta – vad skulle det betyda, och vilka frågor måste vi då besvara? Tegmark menar att denna diskussion inte bara är intressant utan helt enkelt nödvändig för oss att ta tag i. Den förändring som en sådan teknisk landvinning skulle innebära är så omvälvande att den kräver att vi tänker om på alla möjliga olika områden: ekonomi, politik, juridik, forskning, medicin – de flesta av samhällets beståndsdelar förändras i grunden när, eller om, detta sker.

Innan vi kan påbörja den diskussionen måste vi dock fråga oss hur en sådan omvälvning kommer att se ut. Tegmark påpekar att en överväldigande andel av de olika framtidsscenarioer som brukar spelas upp tycks förutsätta två grundläggande antaganden. Det ena är att utvecklingen mot en sådan "superintelligens" accelererar exponentiellt, och att det bara kommer att vara en fråga om dagar från det att den mänskliga intelligensen blir förbisprungen till att den blir fullständigt underlägsen. Det andra regelmässiga antagandet är att det handlar om en enskild intelligens som systematiskt och globalt tar över alla uppgifter.

Ingen av dessa premisser är absoluta, påpekar Tegmark. Det är fullt möjligt att tänka sig utvecklingen av flera olika artificiella intelligenser och en mer evolutionär takt i utvecklingen. I konsekvens med att vi talar om artificiell intelligens som en livsform borde vi anlägga ett mer biologiskt än datavetenskapligt perspektiv, menar han.

Det är ett viktigt påpekande. Alltför ofta förvandlas bilden av artificiell intelligens till bilden av ett platonskt väsende som omfattar allt, en sorts helt obunden existens som överstiger vårt förnuft, misstänkt likt en gud. Det finns också något trösterikt i tanken att vi skulle kunna överlåta alla våra problem – klimatförändringar, fattigdom, pandemier – åt en varelse som är överjordiskt intelligent. Diskussionen om artificiell intelligens kan alltså ibland få en bismak av teologi, inte minst i diskussionen om singulariteten – den punkt då den artificiella intelligensen accelererar så snabbt att den ögonblickligen blir universell.

Ett biologiskt perspektiv fordrar i stället att vi funderar på sådant som den artificiella intelligensens metabolism och livslängd, eller dess konstitution och storlek. Idag kräver en mänsklig hjärna mycket mindre energi än en dator för att spela brädspelet go, men schack kan datorn

spela med mindre energiåtgång än oss. Detta spelar stor roll – artificiell intelligens som förutsätter en energiåtgång som är betydligt större än om en människa utför samma uppgift vore inte särskilt värdefull.

Med ett biologiskt perspektiv får vi också anledning att fråga oss om även artificiella intelligenser kommer att utveckla ekosystem, leva i symbios med artificiella bakterier och kanske drabbas av högst verkliga om än artificiella virus. Livet är trots allt ganska väl organiserat, och föreställningen om den artificiella intelligensen som något starkare, renare och friare kanske är mer ett uttryck för vår egen flyktinstinkt från köttets snöda verklighet än en god intuition?

Här är, i de grundläggande frågorna kring vad som skulle kunna hända, finns mycket att hämta ur Tegmarks tankereda och fantasi. Han skriver medryckande och med inslag av fiktion på ett sätt som stimulerar läsarens egna tankar.

Samtidigt finns också högst verkliga och mer aktuella frågor som vi måste ställa oss. Bland de mer akuta finns frågan huruvida det är rimligt att militarisera den allt kraftfullare tekniken. Vill vi se autonoma drönare som själva identifierar och väljer sina mål? Instinktivt ryggar nog alla tillbaka inför den frågan, men tänk om det kan visas att vi därigenom kan minimera förlusten av civila vid olika tillslag? Inom medicinens område stöter vi på en annan akut fråga: vill vi överlåta diagnoser och kirurgi till robotar helt och hållet? Och så vidare.

Den franska filosofen Simone Weil utvecklade i sin bok "Att slå rot" (1949) tre kriterier för utformningen av maskiner som ekar i de frågor som Tegmark ställer i "Liv 3.0". Hon menade att en bra maskin ska vara säker, den ska kunna användas för en mångfald syften och den ska alltid opereras av en människa. Det sista kravet, menade Weil, hade att göra med att arbetets mening annars försvinner – och kanske också dess mänsklighet. Men hur ställer vi sådana krav mot effektivitet och precision?

Vi ställs redan nu inför frågor om hur vi hanterar system som nått en sådan grad av komplexitet att det inte längre är möjligt för oss att förklara hur de fungerar i detalj. Dessa system är inte svarta lådor i den bemärkelsen att de hålls stängda för oss, men de är hermetiskt till slutna av den komplexitet som finns inbyggda i dem. Samtidigt fungerar de allt bättre. Tegmark pekar på den forskning som nu krävs om säkerhet och tillförlitlighet, och menar att denna kommer att bli avgörande för om vi accepterar en växande roll för artificiell intelligens på områden där vi hittills krävt en mänsklig hand.

Perspektiven blir emellertid verkligen svindlande när Tegmark kommer in på gränssytorna mellan kosmologi och artificiell intelligens. Om artificiell intelligens är nästa steg i livet betyder det att vi borde förvänta oss att den expanderar på samma sätt som livet självt gör det. Tegmark noterar förstulett att artificiell intelligens skulle kunna öka biosfärens storlek med en 32 storleksordning, och om vi letar efter utomjordiskt liv så är det nog bäst att leta efter oorganiskt liv, eftersom det mycket lättare kan sprida sig i rymden. Här finns en intressant evolutionstanke: att organiskt liv endast är ett genomgångsleed (och ett ganska kort sådant) i evolutionen.

Max Tegmark har inte bara skrivit en bok om dessa frågor, utan även grundat ett institut som ska syssla med dem – Future of Life Institute. Han har samlat in en försvärlig summa pengar, men framför allt en stor skara namnkunniga personer som gemensamt förbundit sig att driva dessa frågor vidare.

Sanningen är förstås att vi inte vet om vi någonsin kommer att lyckas skapa en superintelligens – och hur den i så fall kommer att gestalta sig. Samtidigt vore det oansvarigt att inte överväga vad det skulle innebära om vi lyckades. Och hur detta "vi" skulle förändras i grunden när en ny intelligens träder in i det.

Nicklas Berild Lundblad

Adjungerad professor i innovation vid KTH och samhälls-politisk chef för Europa på Google  
understrecket@svd.se