

Skal man tage hensyn til robotter?

Spørgsmålet om etiske hensyn til robotter rejser en lang række problemer, for vi er nødt til at vide noget om deres indre følelsesliv, før vi ved, hvilke hensyn vi skal tage.

Alle er lige, undtagen dyr

De væsener, vi mennesker tager mest hensyn til, er andre mennesker. Ikke at vi altid gør det, men vi gør det da i det mindste somme tider. Og at vi anser os selv for at have en helt speciel etisk status, fremgår med al ønskelig tydelighed af de meget krævende regelsæt, vi har udformet om behandlingen af mennesker. Tag bare menneskerettighedserklæringens artikel 1: "Alle mennesker er født frie og lige i værdighed og rettigheder. De er udstyret med fornuft og samvittighed, og de bør handle mod hverandre i en broderskabets ånd". Sådan er det ikke med dyr. Dyr holder vi i fangenskab, ofte på meget lidt plads. Det gør vi som regel for at fodre dem op, slagte dem og spise dem. I Danmark har en burhøne nogenlunde et A4-arks plads at boltre sig på, og i Europa produceres der årligt mere end 4,5 milliarder kyllinger. Der er altså langt flere kyllinger end mennesker i EU, men de optager en meget lille procentdel af den samlede plads. Det har ikke meget med broderskabets ånd at gøre.

Hvis man tilhører arten homo sapiens, er man altså automatisk del af et helt særligt fællesskab, som giver én rettigheder og bemyndigelse til at herske over andre væsener. Men hvor rummeligt bør dette fællesskab egentlig være? Skal der være plads til andre væsener med menneskelige træk, som blot har det lille handicap, at de ikke helt igennem er mennesker? Det kunne for eksempel være chimpanser eller fremtidens robotter. Eller skal det være en eksklusiv klub, hvor kun væsener, der er fuldgældige medlemmer af homo sapiens, har adgang?

Forkæmperne for dyrs rettigheder

I løbet af 1980'erne tog diskussionen om de forskellige væseners etiske status en helt ny drejning. Fra flere sider blev det hævdet, at medlemskab af arten homo sapiens ikke i sig selv berettiger én til en særlig status i forhold til andre væsener. At hævde noget sådant er blot en måde at undertrykke væsener, der ikke tilhører gruppen. I den forstand er der ikke nævneværdig forskel på medlemskabet af homo sapiens og de andre gruppetilhørsforhold, der i tidens løb er blevet brugt til at retfærdiggøre særlige privilegier for netop den gruppe. At lægge vægt på, om man tilhører den ene eller anden art, skaber således en ulighed, der svarer til racisme: "Når der er interesse modsætninger mellem racer, tilsidesætter racister lighedsprincippet ved at lægge mest vægt på de interesser, medlemmerne af deres egen gruppe har. På samme måde lægger de personer, jeg vil kalde speciesister, mest vægt på de interesser, medlemmerne af deres egen art har, når der er interesse modsætninger mellem arter".

Men hvis man opgiver speciesismen, hvad skal man så lægge vægt på, når man har med andre væsener at gøre, hvis man altså ønsker at handle på en etisk acceptabel måde? Ifølge forkæmperne for dyrs rettigheder er svaret på dette spørgsmål i princippet meget enkelt. Man skal se på, hvilke interesser væsener, man står over for, har, hvad enten det er et dyr, et menneske eller noget helt tredje. Og dyr har tydeligvis interesser. Dyr virker til at have en slags bevidsthedsmæssige oplevelser på samme måde som mennesker. De kan for eksempel føle smerte, sult og andre former for ubehag. Og det er naturligvis i deres interesse at slippe for disse oplevelser og så i stedet føle tilfredshed og behag. Og burhønen er næppe i en tilstand af behag, når den sidder fastlåst på sin A4-arks-plads. Det er netop det, der gør det etisk betænkeligt overhovedet at holde burhøns.

Men hvordan finder man ud af, om forskellige væsener har bevidsthedsmæssige oplevelser og dermed interesser? Kan man overhovedet være sikker på, at dyr har mentale oplevelser? Og hvad med robotter?

Hvad ved vi om andre væsener?

Den berømte filosof Rene Descartes (1596 - 1650) gav i sine værker udtryk for den opfattelse, at dyr må betragtes som en slags automater eller robotter, som ikke har mentale eller bevidsthedsmæssige oplevelser. Det var en opfattelse, som var vel hørt i samtiden, hvor den blev brugt til at retfærdiggøre, at man kunne behandle dyr, som man havde lyst til. For hvis dyr hverken kan føle smerte eller behag, har de vel ingen interesser, og så er der ikke nogen etisk begrundelse for at tage hensyn til dem overhovedet.

Umiddelbart virker det som taget ud af den blå luft, at dyr ikke skulle kunne føle for eksempel smerte eller ubehag. De udviser jo i mange sammenhænge en adfærd, der tyder på, at de har den slags følelser, blandt andet udstøder de lyde, ser paniske ud eller forsøger at stikke af. Men for Descartes og hans samtid var der en enkel forklaring på det. Man anså dyrenes reaktioner for at være ren adfærd, dvs. at der ikke er knyttet en indre bevidstheds- eller oplevelsesmæssig side til adfærden. Der er altså alene tale om ydre bevægelser, som ikke opleves indefra af et væsen, det kan gøre ondt på.

Descartes og hans samtids forståelse af dyrene rejser et dybt filosofisk problem, som aldrig siden har forladt filosofien: Hvordan kan vi egentlig vide, om andre væsener har mentale eller bevidsthedsmæssige følelser og oplevelser - og i givet fald hvordan disse oplevelser er? Spørgsmålet virker nok forståeligt, når man snakker om hamstere, regnorme og firben. Det er nemlig helt umuligt at sætte sig ind i, hvordan verden ser ud for disse dyr. Men spørgsmålet har en dybere pointe end det. For sagen er, at vi - også i vores omgang med andre mennesker - aldrig har direkte adgang til andre individers bevidsthedsmæssige oplevelser. Det eneste, vi kan erkende, er jo deres adfærd - at de bevæger sig, siger noget osv. Og derfra slutter vi, at de har visse "indre" oplevelser. Men strengt taget ved vi ikke noget om det, for de oplevelser, vi går ud fra, at andre har, er under alle omstændigheder strengt private. Derfor kan man principielt set aldrig kontrollere andres påstande om egne indre oplevelser. Sætningen "det gør meget ondt" forekommer naturligvis meget troværdig, hvis man lige har fået sin arm savet af. Men nogle mennesker hævder, at de permanent føler sig så udmattede og energiforladte, at de ikke er i stand til at leve en normal tilværelse. Men hvis man ikke kan se hvorfor, de skulle have det sådan, hvordan skal man så forholde sig til det? Og giver det overhovedet mening at sammenligne indre oplevelser? Spørgsmål som "Er du lige så træt som mig?" "Elsker du mig højere, end jeg elsker dig" eller "Gør din forbrænding mere ondt end min?" er vel ret beset så umulige at have med at gøre, at de i nogle sammenhænge nærmer sig det meningsløse. Dette illustrerer netop, at man i bund og grund intet ved om, hvad der rent bevidsthedsmæssigt foregår "inde" i andre, hvis der da overhovedet foregår noget.

Har robotter et indre liv?

Mange nutidige filosoffer tager det ovennævnte problem meget alvorligt. De hævder, at vi strengt taget ikke ved noget om, hvorvidt andre personer eller væsener har et indre liv. Det er en antagelse, som kan være mere eller mindre velbegundet, men sikker viden har vi ikke.

Men hvad taler henholdsvis for og imod den antagelse, at robotter har et indre liv, som giver dem etisk status? Svaret på dette spørgsmål er helt afgørende for, om man skal tage hensyn til robotter. Hvis robotterne har en eller anden form for indre liv, forekommer det sandsynligt, at de også har interesser. For den måde, man behandler robotterne på, må så påvirke robotternes indre liv. Hvis man sparker en hest, gør det ondt på den. Derfor er det uetisk at gøre det uden en meget god grund. Og hvis man også kan påvirke

robotters oplevelser i positiv eller negativ retning, så er det heller ikke irrelevant fra et etisk synspunkt, hvordan man omgås dem.

Men er der grund til at tro, at der på et tidspunkt udvikles robotter, som har et indre liv? Det mener visse forskere, og for nogle af dem er der en meget enkel grund. Rodney A. Brooks hævder, at mennesket historisk set altid har været tilbøjelig til at se sig selv som et ganske særligt væsen, i nogle perioder med den begrundelse, at mennesket som det eneste væsen er skabt i Guds billede. Og hvis mennesket allerede i udgangspunktet og næsten pr. definition anses for at være noget særligt, kan det naturligvis bruges til at begrunde, at vi har en række egenskaber, som andre væsener ikke har.

Er mennesket i grunden også en maskine?

Forestillingen om menneskets sjæl er et eksempel på, hvad der kan betegnes som menneskets adelsmærke. Det har været tanken, at mennesket har en sjæl, som bibringer mennesket udødelighed og en række andre ophøjede egenskaber. Men hvis man opgiver ideen om menneskets særlige status, åbner man ifølge Rodney Brooks samtidig for, at også andre væsener end mennesket kan have indre oplevelser på stort set samme måde, som mennesket har. Brooks mener, at man kan nå frem til denne konklusion gennem en erkendelse af, at mennesket i bund og grund ikke selv er andet end en avanceret maskine: "Every person I meet is also a machine - a big bag of skin full of biomolecules interacting according to describable and knowable rules. When I look at my children, I can, when I force myself, understand them in this way. I can see that they are machines interacting with the world."

Hvis mennesker er maskiner, så kan det naturligvis ikke udelukkes, at andre maskiner, nemlig robotterne, en gang i fremtiden udvikler egenskaber, der kan sammenlignes med det menneskelige bevidsthedsliv. For spørgsmålet er ikke så meget, om det er muligt at skabe andre væsener med et indre liv - det må det i udgangspunktet anses for at være, for det handler blot om at udforme det materielle fundament på den helt rigtige måde. Så spørgsmålet er snarere, om vi bliver dygtige nok til at overskride de nuværende begrænsninger, så vi en gang i fremtiden faktisk kommer til at stå over for et sådant væsen.

Hvis vi kommer til at stå over for et avanceret robotvæsen med indre liv, må vi ifølge Rodney Brooks endnu en gang udvide kredsen af individer, som er omfattet af etikken: "When our robots improve enough, beyond their current limitation, and when we humans look at them with the same lack of prejudice that we credit humans, then too we will break our mental barrier, our need, our desire to retain tribal specialness, differentiating ourselves from them. Such leaps of faith have been necessary to overcome racism and gender discrimination. The same sort of leap will be necessary to overcome distrust of robots".

Hvilken karakter har robotters indre liv?

Forestil dig, at der en dag går en robot hen til dig og siger: "Jeg er et bevidst væsen!". Hvordan ville du forholde dig til det? Ja, det ville jo i første omgang afhænge af situationen. Du ville ikke tage det alvorligt, hvis der var tale om en robot, der gentog alt, hvad du selv sagde - og du lige havde sagt: "Jeg er et bevidst væsen". Du ville heller ikke tage dig særligt af det, hvis du havde med en robot at gøre, der var programmeret til helt mekanisk at sige bestemte ting i bestemte situationer. Nej, ytringen ville først for alvor være interessant og måske ligefrem uhyggelig, hvis du havde grund til at tro, at taleren selv forstod sætningen og havde en hensigt med at udtale den.

Men hvad skulle der til, før du faktisk ville tro, at en talende robot havde bevidsthed? Spørgsmålet er vanskeligt at svare på, eftersom de kriterier, vi mennesker i al almindelighed bruger til at tilskrive andre væsener bevidsthed ud fra, er meget komplicerede og upræcise. Det har blandt andet noget at gøre med, om det pågældende væsen har en tilstrækkeligt kompleks hjerne og et centralnervesystem. Og det

afhænger også af, hvordan væsenet opfører sig, om det for eksempel udviser smerteadfærd eller undgår bestemte situationer.

Men man kan jo ikke uden videre antage, at de kriterier, vi bruger til at tilskrive andre væsener bevidsthed ud fra, ville være de samme som dem, det ville give mening at bruge i forhold til robotter. Alle andre væsener er ifølge den gængse darwinistiske forståelse et produkt af den naturlige selektion, som har fundet sted over millioner af år. Og mange af os har endda fælles forfædre. Vi har derfor helt automatisk visse fællestræk. Men robotter ville jo være blevet til på en helt anden måde. Så formodentlig ville det være nødvendigt at forholde sig til dem ud fra en helt anden målestok. Det må tiden vise.

Men en gang i fremtiden bliver vi måske faktisk enige om at tilskrive robotter bevidsthed. Forestil dig nu videre, at robotten efter at have sagt "jeg er et bevidst væsen" med rusten stemme fortsatte med at sige: "og jeg synes, min tilværelse er meget meningsforladt". Hvad ville du nu stille op? Tja, hvis det var et helt almindeligt menneske, der på samme måde talte om meningsforladthed, ville du måske ikke være så meget i tvivl om, hvad du skulle stille op. Du ville spørge til, hvorfor tilværelsen virkede meningsforladt. Og ikke mindst ville du givetvis have nogle ideer om, hvordan personen skulle hjælpes, for eksempel: "Hvad med at få et arbejde, så du får noget fornuftigt at give dig til og kommer lidt mere ud blandt andre mennesker?". Som udgangspunkt ville du gå ud fra, at det andet menneske og du havde så mange fælles værdier, forventninger og erfaringer, at du ville kunne sætte dig ind i vedkommendes situation. Netop derfor ville du være i stand til at gå ind i en dialog med ham eller hende og eventuelt endda støtte personen i at få et indhold i livet. I nogle situationer ville du måske endda føle dig etisk forpligtet til at gøre dette, for eksempel hvis personen var en af dine venner eller noget af din familie.

Kan vi forstå robotter?

Den berømte østrigske filosof, Ludwig Wittgenstein (1889-1951), skrev i sit værk *Filosofiske Undersøgelser*, som udkom to år efter hans død, en sætning, der siden hen er blevet kendt: "Hvis en løve kunne tale, kunne vi ikke forstå den". I sig selv er sætningen ikke noget særligt. Det særlige var de begrundelser, Wittgenstein gav for den. Ifølge Wittgenstein er det nemlig meget forenklet sagt sådan, at vi kun til fulde kan forstå det andre siger, hvis vi til en vis grad indgår i den samme livsform som dem. Der bliver altså nødt til at være nogle fælles erfaringer, værdier, gøremål, handlemønstre osv., for at man kan have en meningsfuld dialog. Hvis man tror, at løver ville bruge menneskesprog, men blot tale om noget andet end os mennesker, tager man altså fejl. Løverne ville bruge et sprog, der passede bedre til deres livsform, end menneskesproget gør. Netop derfor ville det være svært at forstå dem, også selv om man gjorde sig store anstrengelser for at oversætte det, de sagde. Det ville nemlig ikke uden videre kunne oversættes.

Hvis man havde med robotter at gøre, ville situationen så være anderledes end med løver? Umiddelbart forestiller man sig jo nok, at "den bevidste robot" taler flydende menneskesprog og kan fortælle om sig selv og om sine oplevelser og interesser. Men hvis Wittgenstein har ret, er det slet ikke så enkelt. For man kan naturligvis ikke uden videre gå ud fra, at bevidste robotter har den samme type af indre oplevelser eller den samme type af interesser, som vi andre har. Måske er menneskesproget ikke egnet til at beskrive deres erfaringer. Så det kan være, at vi slet ikke er i stand til at sætte os ind i deres situation. Men hvis vi skal tage hensyn til dem, kræver det utvivlsomt et vist minimum af forståelse for, hvad deres eksistens overhovedet går ud på. Det er nærmest en umulig opgave at varetage deres interesser, hvis man i bund og grund ikke kan sætte sig ind i, hvad deres perspektiv på verden er.

Et tilsvarende problem kunne i øvrigt opstå, hvis man - i stedet for at udvikle robotter helt fra bunden af - videreudviklede allerede eksisterende mennesker ved at bygge dem sammen med forskellige former for teknologi. I nogle tilfælde ville det formentlig ændre det mentale liv hos de cyborgs, man fik ud af det. For

eksempel er det svært at forestille sig, at det ikke ville forandre et menneskes oplevelser fundamentalt, hvis hjernekapaciteten blev mangedoblet, og han eller hun samtidig fik mulighed for at koble sig trådløst op på nettet alene ved tankens kraft. Måske ville personen få en helt anden og meget mere decentreret identitetsfølelse. Igen ville det måske være svært for andre mennesker at forstå dette individs specielle perspektiv på verden. Og dette kunne gøre det til et problem, hvordan man rent etisk skulle behandle personen.

Robotter: Hvad er det etisk rigtige?

Det kan ikke udelukkes, at der en gang udvikles robotter med mentale eller bevidsthedsmæssige oplevelser og et "indre liv". Udvikles der sådanne væsener, skal de naturligvis tildeles en etisk status, så de er omfattet af etikken og bliver taget hensyn til for deres egen skyld. Det lyder jo meget enkelt, men der er intet enkelt ved at udvikle en samlet robotetik.

Det første svære spørgsmål er, hvordan vi skal afgøre, om robotter har et indre liv? Hvis vi har programmeret dem til at tale og opføre sig som menneskelige væsener, hvordan skal vi så kunne skelne imellem det, der kommer fra os selv, og det der kommer fra robotten selv?

Det andet svære spørgsmål er, hvordan vi skal behandle en robot, hvis vi er blevet enige om at tilskrive den en etisk status, så den er omfattet af etikken og skal tages hensyn til på samme måde som mennesker og andre sansende væsener? Svaret afhænger selvfølgelig af, hvordan dens indre liv fungerer, og hvordan dens perspektiv på tilværelsen i øvrigt tager sig ud. Vi bør jo heller ikke behandle dyr og mennesker fuldstændig ens, når de nu har forskellige behov og forskelligt perspektiv i forhold til verden. Men er der overhovedet grund til at tro, at vi kan sætte os ind i, hvordan det er at være robot? Det er svært at svare på. Under alle omstændigheder er der naturligvis et indbygget problem i, at den bevidste robot formodentlig vil være afhængig af de mennesker, der helt fra grunden har skabt den til brug for egne formål. Medmindre en robots bevidsthed er så upåvirkelig af det, der sker i den ydre verden, at den er helt ligeglad med, hvordan man behandler den.

Det tredje svære spørgsmål er, hvordan vi skal forholde os til, at vi ikke kan besvare selv de mest grundlæggende spørgsmål om robotetik. Måske er det etisk rigtige i sidste ende at lade være med at udvikle robotter, hvis vi blot har den mindste smule mistanke om, at de udvikler bevidsthed.