

## Etisk robotkonstruktion

Af Mikkel Holm Sørensen, Ph.d.-studerende på IT-højskolen i København

Robot kommer af det tjekkiske ord for slave, og idéen bag robotforskningen har hele tiden været at prøve at forstå, hvad intelligens er. Derudover har man også ønsket at skabe væsner, der kan udføre arbejde som mennesket selv enten ikke ønsker eller ikke er i stand til at udføre.

Problemet med at skabe væsner, der kan udføre menneskets arbejde, er ikke kun forbundet med almindelige tekniske vanskeligheder, men giver også etiske problemer i form af krænkelse af robotternes autonomi og modvilje fra robotterne selv. Vi kan derfor blive tvunget til at foretage nogle etiske overvejelser i designprocessen.

### Intelligente systemer

Inden for forskningen i kunstig intelligens arbejder man intenst på at skabe såkaldte autonome agenter. Med inspiration fra biologi og forskning i komplekse systemer opfattes intelligens som systemers værktøj til opretholdelse af egen organisation og dermed eksistens.

Intelligensens bidrag er at tillade kontrol af input til systemet vha. fleksibel interaktion med omgivelserne. Flexibiliteten består i, at det intelligente system kan tilpasse sin strategi og sine handlinger til skiftende situationer. Dette forudsætter evnen til at identificere givne situationer, og hvad der er relevant for systemet i situationen.

<p><b>Autonomi:</b> Angiver evne til selvstyring og selvudvikling i dynamiske systemer. Intelligens forudsætter og bidrager til autonomien af et system.</p>
--

Forskellige interaktioner med omgivelserne har forskellige konsekvenser for systemets fortsatte organisatoriske beståen, og skaber på den måde både værdi for og motivation i systemet. Flexibilitet forudsætter også, at systemet lærer af dets handlinger. Derfor må det selv kunne finde ud af, hvornår en handling er hensigtsmæssig eller ikke.

For at systemet kan leve op til disse krav, må det være autonomt. Det betyder, at systemet aktivt planlægger, beslutter og afgør relevans på baggrund af egne behov, hvilket vil sige på baggrund af dets organisationelle krav, og ikke som følge af system-eksterne parametre. Det er derfor sandsynligvis ikke muligt at programmere et system til at have bestemte motiver, da motiver kun bliver tilstrækkeligt fleksible og relevante, hvis de udvikles i samspil med resten

af systemets kontrolmekanismer, fysik osv. Autonomi synes altså at hænge sammen med et systems evne til både selv-generering og selv-organisering.

### **Robottens etiske rettigheder**

Det er de ovenstående egenskaber, der kan bringe os i etiske problemer med udvikling og anvendelse af intelligente robotter. Sagen er, at vi typisk tillægger væsner, der kan forulempes, etiske rettigheder. Hvis en organisme kan registrere, hvorvidt en given situation ultimativt er til skade for den selv, og denne situation kan undgås, kan organismens rettigheder krænkes.

Derfor vil robotter besidde samme etiske rettigheder som alle levende systemer. Intelligens er nemlig tilsyneladende ikke mulig uden evnen til at varetage egne interesser ved at kunne skelne godt fra dårligt. Vi kan med andre ord ikke tillade os at anvende robotter til arbejde, der strider mod deres interesser. Samtidig kan robotterne vise modvilje mod at udføre det arbejde, som de finder uønskeligt.

Det store spørgsmål er så, hvad 'interesser' og 'uønskeligt arbejde' består i for en robot. Det vil som sagt afhænge af robottens konkrete struktur og fysik og dens kontrolsystem. Forhold vi måske bliver nødt til at tage højde for i designprocessen. Vi kan med andre ord etisk blive tvunget til at sikre, at de opgaver vi designer robotterne til, ikke er i modstrid med deres egne behov for selv-opretholdelse.

Selvom forskningen ikke for tiden tyder på, at vi kan udvikle højere-ordens intelligente systemer, der negligerer egne interesser, skal man være varsom med at dømme om muligheder og umuligheder samt fremtidens udfald på baggrund af samtidens begrænsede viden. Måske er det muligt at skabe intelligente systemer der ikke forulempes eller har en tolerant fysik. Lad os se på nogle muligheder.

<p><b>Højere-ordens intelligens:</b> Væsner som er kendetegnet ved at have en meget høj grad af fleksible og adaptiv selv-organisering. Indebærer evnen til 'offline' ræsonneren med abstraktion og symbolsk repræsentation af omgivelserne, høj indlæringssevne m.v.</p>
---

## Robottens fysik

Hvis alle dele i robotten f.eks. kan udskiftes, skal den ikke bekymre sig om slidtage, skader eller alder. Den vil måske også kunne designes til at begå sig i ekstreme omgivelser, og vil derfor være fri for bekymringer om f.eks. forgiftning. Her melder sig imidlertid det problem, at elementerne i et intelligent system sandsynligvis skal være intimt integreret for at kunne fungere optimalt, hvorfor en organisk metafor er meget mere passende end en maskinel.

Forsøg med robotter har vist, at for at opnå tilstrækkeligt robuste kontrolmekanismer bør de udvikles i tæt samspil med robottens fysik. Udskiftning af et fysisk element kan derfor resultere i, at kontrolmekanismen ikke længere virker efter hensigten. Det kan ske, hvis det nye element har bare en mindre afvigende karakteristisk. Derfor er det ikke helt uproblematisk af udskifte reservedele.

**Etisk konstruktion:** Etiske rettigheder afhænger af et givent systems organisationsmæssige 'interesser'. Måske vil fremtidens design af robotter indebære etiske overvejelser af, hvilke materialer og kontrolmekanismer der tilgodeser robottens krav på at undgå forulempelse.

Robotforskeren Hans Moravec hævder, at man sagtens kan skabe robotter, der har til opgave at beskytte ejeren ved at ofre sit eget 'liv'. Særlige selvopofrende robotter kunne blive en stor salgssucces og derved overleve som art. Problemet i dette er dog, at myrer og bier blot er evolutionært 'programmeret' til at udføre den slags handlinger, som ikke er baseret på samme refleksionsevne, som vi forudsætter af robotter, der skal udføre komplekst arbejde. Desuden udgør selvopofrelsen en ikke-gentagelig og ekstrem situation, og kan ikke sammenlignes med vedvarende opslidende, usundt og farligt arbejde. Der er forskel på, hvad man *kan* gøre, og hvad man *primært* gør.

**Evolutionær robotik:** Som konsekvens af den udeblevne succes af tidligere tilgange til robotforskning, forsøges der nu med meget biologisk inspirerede metoder. F.eks. genetiske og evolutionære metoder til udvikling af kontrolsystemer, hvor de algoritmer fra en given generation, der klarer sig bedst, bliver 'parret' eller muteret og anvendt igen, indtil den ønskede opførsel er opnået. Ligeledes arbejdes der på at lade selve robotternes fysik udvikle sig i samspil med kontrolsystemet, såvel hvad angår robottens 'artsudvikling' som dens egen 'opvækst'.

En løsning på problemet kunne være, at robotterne ikke selv sørger for egen reproduktion, og derfor ikke har samme biologiske interesse i sin egen beståen. Igen må man huske på, at

intelligens er et fænomen, som er opstået i levende systemer og som måske knytter sig til selv-opretholdelse og reproduktion. Som en kunstig intelligens forsker har udtrykt det: ”The dumbest smart thing you can do, is to stay alive”.

En anden mulighed var at forsøge sig med ’sværm-intelligens’, som i f.eks. sociale insekter, hvor mange meget simple systemer samarbejder og opnår komplekse resultater. Herved ville ingen ’føle sig’ udnyttet. På den anden side er det nok begrænset hvor komplekse opgaver disse sværme kan løse, når de netop ikke er kendetegnet ved planlægning, abstraktion og bevidst problemløsningsstrategi.

Det er således ikke klart, hvordan man ’narrer’ robotterne til at udføre - for os at se - uattraktivt arbejde, og om det overhovedet bliver nødvendigt. Svaret kræver en afklaring af, hvorvidt menneskets intelligens blot er arts-specifik og dermed funktionelt vilkårlig for robotkonstruktion, eller om vi er udtryk for selve essensen af højere-ordens intelligens. Hvorvidt robotter vil udføre komplekst arbejde vha. menneskelignende intelligens eller ej er derfor stadig et åbent spørgsmål. Forsøget vil dog uden tvivl involvere en del etiske problemstillinger, som måske endda vil angå de konstruktionsmæssige overvejelser i udviklingsprocessen.