

Den kunstige intelligens er eksponent for

Rationalitetens ufornuft

Forskningsmiljøet omkring kunstig intelligens undertrykker enhver problematisering af området, men kendsgerningen er, at man slet ikke har opnået de resultater, man forudsagde for 10 eller 20 år siden

Af BENT FLYVBJERG

EN SOMMERDAG i 1964 ankom Hubert Dreyfus, en ung filosofilærer fra MIT, Massachusetts Institute of Technology i Cambridge, til RAND Corporation i Californien. RAND er et stort rådgivende konsulentfirma inden for forskning og udvikling, særlig kendt for to ting: Sit arbejde for det amerikanske militær. Og at en meget stor del af USAs topviden-

KRONIKKEN

skabsfolk på et eller andet tidspunkt har arbejdet for RAND.

Nobelpristageren Herbert Simon er en af dem. Sidst i 50'erne og først i 60'erne udførte han verdenskendt pionerarbejde med udvikling af kunstig intelligens hos RAND. Hubert Dreyfus' bror, Stuart, var ansat hos RAND som datalog. I en privat korrespondance havde Hubert og Stuart diskuteret RANDs arbejde med kunstig intelligens. Stuart havde fortalt Hubert om den euforiske stemning over de første resultater. Hubert havde svaret, at kunstig intelligens-folkene burde læse den tyske filosof Martin Heidegger, inden de gik videre i deres arbejde. Filosofien lå inde med viden, som kunne bruges til at forstå muligheder og begrænsninger ved kunstig intelligens. Og ifølge Hubert var der god grund til at slå koldt vand i blodet.

STUART SKAFFEDE kort efter Hubert et midlertidigt job hos RAND med at evaluere forskningen i kunstig intelligens. Det var derfor, Hubert tog til Californien i 64. Efterhånden som han gennemgik firmaets arbejde punkt for punkt, blev hans forudnelser bekræftet: Optimismen for kunstig intelligens var overdrevet. Den skyldtes, at det var lykkedes forskerne at få en computer til at manipulere symboler

og ikke bare tal - dengang et vigtigt resultat i sig selv - og herigennem løse nogle simple logiske opgaver. Men Hubert Dreyfus fandt intet grundlag for Herbert Simons og hans forskerholds krav på at have kastet lys over de processer, som indgår i forståelse, intuition og indlæring. Og hvad mere er, han mente at kunne sige, at disse processer ikke kunne imiteres med eksisterende computerteknologi. I sin rapport til RAND, provokerende kaldet *Alkymi og kunstig intelligens*, konkluderede Hubert Dreyfus, at *»en signifikant udvikling af kunstig intelligens må afvente computere af en helt anden slags (end digital computeren), hvoraf den eneste eksisterende prototype er den lidet udforskede menneskelige hjerne«*.

I dag, mere end 20 år senere, holder denne konklusion stadigvæk. Herbert Simon forudsagde i 1958, at i løbet af ti år ville verdensmesterskabet i skak holdes af en computer. I 1984 slog en skotsk skakspiller, David Levy, som er placeret omkring nr. 1000 på verdensranglisten, verdensmesteren i computerskak, programmet Cray-Blitz, fire spil mod nul. Cray-Blitz kører på verdens hurtigste computer (en Cray X-MP) og kan vurdere cirka ti millioner slutpositioner i valget af hvert træk.

I 1965 var Simons profetier blevet mere vidtløftige:

»I løbet af tyve år vil maskiner være i stand til at udføre ethvert arbejde et menneske kan udføre«. Andre kunstig intelligens-entusiaster er kommet med tilsvarende forudsigelser. Fælles er, at de alle er gået lige så lidt i opfyldelse som skakforudsigelsen. Computere kan stadig ikke imitere menneskelige intelligens, selv om de er blevet mange gange bedre til at imitere den del, som består i løsning af simple veldefinerede problemer.

SITUATIONEN ER også den samme på en anden måde. Forskningsmiljøet omkring kunstig intelligens undertrykker typisk enhver problematisering af området, et udtryk for dets uvidenskabelige karakter. I 1964 satte Herbert Simon sig kraftigt imod, at Hubert Dreyfus' rapport blev offentliggjort. Det tog et års lobby-virksomhed for Stuart internt i RAND, efter Hubert var taget tilbage til MIT, at få rapporten publiceret. Resultatet var et ødelagt forhold til Simon. Rapporten blev læst over hele verden og er i dag en klassiker på området.

På MIT blev Hubert Dreyfus frosset ud af de lokale kunstig intelligens-forskere. Joseph Weizenbaum, som i dag oversættes til dansk, var den eneste datalogiprofessor med tidlige betænkeligheder over for området. Når han og Dreyfus mødtes, skete det uden for universitetet, så Weizenbaum ikke ville få problemer med kollegerne.

Hubert Dreyfus er i mellemtiden blevet professor i filosofi ved University of California, men løber stadig ind i tilsvarende vanskeligheder. Da han for nylig var inviteret til at optræde i et tv-program om kunstig intelligens, som skulle sendes over hele USA, stillede de deltagende kunstig intelligens-tilhængere, blandt andet Stanford Universitetets John McCarthy, et ultimatum til produceren af udsendelsen. Det skete samme aften, som optagelsen skulle finde sted: Enten blev Dreyfus taget af programmet, eller også deltog tilhængerne ikke. Kunstig intelligens skulle ikke problematiseres nationalt for åben skærm. Og blev det heller ikke. Dreyfus blev sendt hjem.

Men selv om Dreyfus ikke fik lejlighed til at stille dem, er de ubehagelige spørgsmål for kunstig intelligens stadig: Hvad er de uløste problemer? Har der været fremskridt i forsøgene på at løse dem? Hvad er grundlaget for at tro, de kan løses?

HVORFOR ER debatten om kunstig intelligens blusset kraftigt op de seneste år? Hvorfor undertrykker tilhængerne kritik?

For di man kapruster. Og fordi der er penge i oprustning. Krig og krigsforberedelse tåler ikke kritik. Heller ikke når der er tale om en fiktiv computerkrig. *Kunstig intelligens er computerindustriens Star Wars program*, og det promoveres i øjeblikket lige så hårdt. Fjenden er ikke Sovjetunionen, men det teknologisk, økonomisk og konkurrence-mæssigt mere betydningsfulde Japan. Ligesom der snakkes om en *missilkøft* til retfærdiggørelse af den militære oprustning, snakkes der om en *kunstig intelligens-køft*. Og det sidste er en lige så tænkt og interessebestemt størrelse som det første. Dets eksistens legitimerer en forsknings- og udviklingsindsats i milliard (\$)-klassen. »Kløften« er lige så god som penge i banken. Og sådan ønsker kunstig intelligens-folkene at bevare den. Derfor er der stærke interesser forbundet med at kunne præsentere et indtryk udadtil af, at området er videnskabeligt og uproblematisk.

DETTE ER MÅSKE i sig selv nok til at gøre kunstig intelligens interessant, men ikke efter min mening. Der har været så mange fantasiprojekter fra teknologioptimister og computer-freaks i tidens løb. Og i en vis forstand er kunstig intelligens bare endnu et af disse, selv om den uægtelig er andet og meget mere end det.

For mig at se har kunstig intelligens sin virkelige betydning i det faktum, at den kan ses som den ultimative prøvesten og eksponent for den type af rationalitet, som Max Weber kaldte *occidental*: Vestens instrumentelle, analytiske rationalitet. I dag er det næppe rimeligt at betegne denne rationalitet som et rent vestligt fænomen, medmindre det er for at beskrive dens historiske rødder. I dag er rationaliteten global i dens udbredelse og ikke mindst, i dens konsekvenser.

Kunstig intelligens er kunstig i begge betydninger af ordet, når den ses i dette lys. *Jäk-le konstig*. Den er eksponent for en rationalitet, der, samtidig som den er klart nødvendig og uomgængelig for opretholdelsen af det dagligliv, vi kender, også er den største trussel mod dette livs fortsatte eksistens.

Lyder det overdrevet? Lad mig blot nævne rationaliteten i kaprustning, i teknologisk udvikling (f.eks. inden for mikrobiologi), i udnyttelsen af naturressourcer og i markedsstyringen af produktion og forbrug. Noget så grundlæggende for mennesket og for andet liv som vand, energi, dyrkbar jord og skove er truet, for ikke at tale om faren for A-krig. Det handler om modernitetens potentielt fatale rationalitetsbegreb. Men lad det ligge til senere, og lad mig vende tilbage til kunstig intelligens i sig selv.

Hvorfor er man ikke nået længere end tilfældet er med kunstig intelligens? Hvorfor har de optimistiske forudsigelser ikke holdt stik?

For at besvare disse spørgsmål er det nødvendigt at se på den måde, mennesker tænker og lærer på, og sammenligne den med computeres funktionsmåde.

DETALJEREDE STUDIER af menneskers læreproces tyder på, at folk går igennem mindst fem forskellige trin ved indlæring af færdigheder.

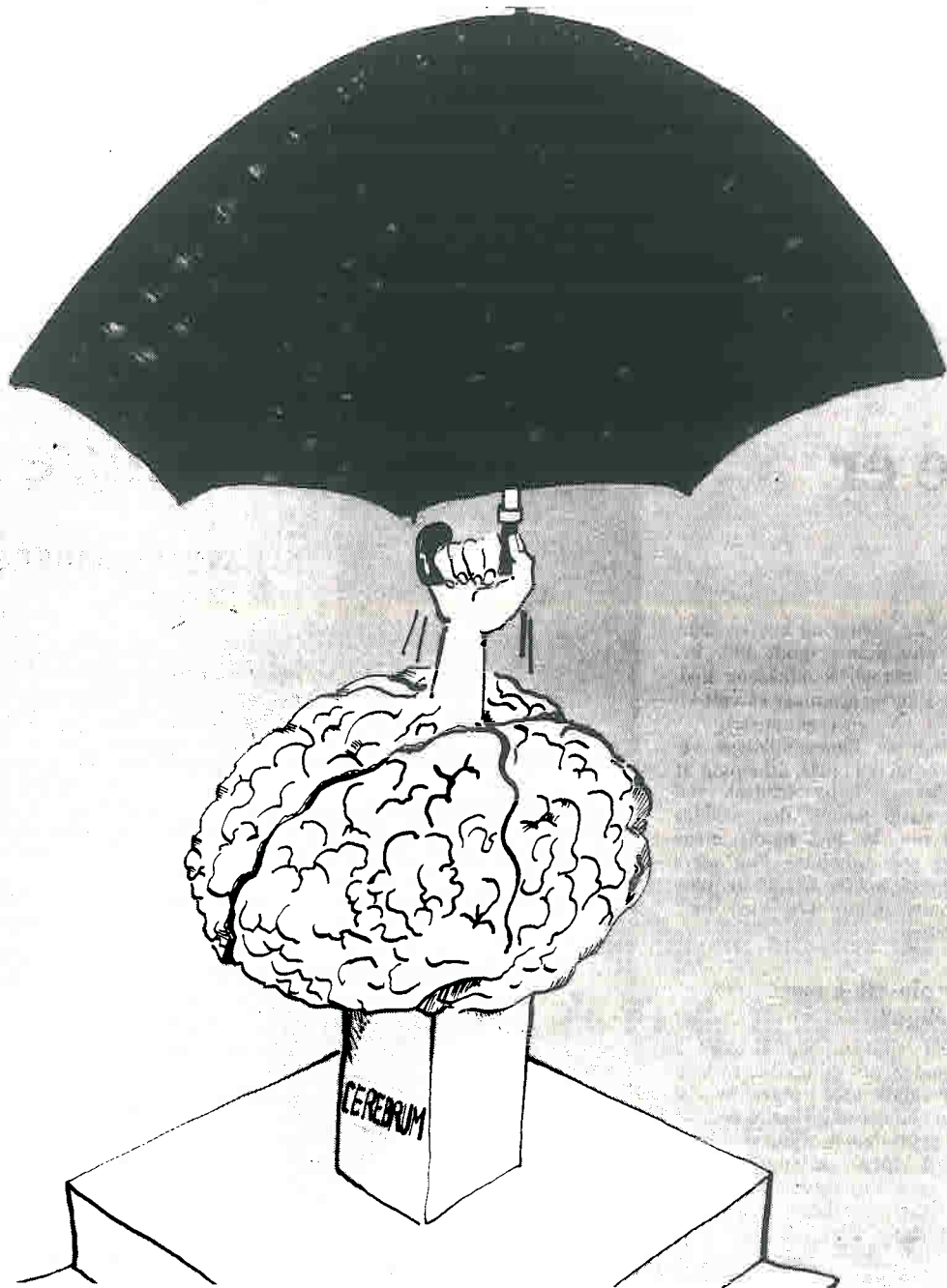
1. *Nybegynder*. Nybegynder er dem, som første gang står over for et givet problem, f.eks. at køre bil. Ved instruktion læres forskellige begreber, objektive fakta og regler, som er nødvendige for handling, f.eks. begreberne hastighed og omdrejningstal og regler for, ved hvilket omdrejningstal der skiftes gear. Definitionerne er så klare og objektive, at de kan forstås uden henvisning til den konkrete situation. De er *situations-uafhængige*, og nybegynders handlinger er *regelbaserede*. Efter nybegyndere har lært en håndfuld regler eller flere for en eller anden færdighed, bliver udøvelsen af dem så kompleks og kræver så meget koncentration, at funktionsdueligheden nedsættes. De første regler er nødvendige for at opnå de første erfaringer. Men regler bliver hurtigt en begrænsning i læreprocessen, og deres betydning må mindskes for at nybegyndere kan komme videre. Det sidste sker ved at gøre konkrete *erfaringer* på egen krop.

2. *Avanceret begynder*. Den avancerede begynder genkender i en konkret situation, gennem sine erfaringer, bestanddele, som er relevante for at udøve den aktuelle færdighed. Bestanddelene genkendes direkte fra situationen uden anvendelse af generelle fakta og regler. De er *situationsafhængige*, og den avancerede

begynder er kendetegnet ved at operere både situationsafhængigt og -uafhængigt. En bilist, som er avanceret begynder, kan skifte gear pr. rygmærk ud fra vibrationer og motorlyde uden bevidst refleksion. Og der er ingen, som bevidst kombinerer fakta og regler for at identificere lugten af frisklavet kaffe. Situationsafhængig erfaring er i dette trin af læreprocessen vigtigere end nogen form for situations-uafhængig information.

3. *Kompetent udøver*. Igen er det nødvendigheden af at håndtere en kompleks virkelighed, som er drivkraften frem til næste trin. Med mere erfaring bliver antallet af relevante genkendelige bestanddele i en bestemt situation overvældende, og folk begynder en løbende *prioritering* og *planlægning*. Den kompetente sygeplejerske går i modsætning til begynderen ikke automatisk rundt fra patient til patient i en forudbestemt rækkefølge, men vurderer løbende behovet for opmærksomhed og pleje og planlægger sit arbejde tilsvarende. Kognitivister og andre, der tænker på tænkning som problemløsning, fokuserer på de tankeprocesser som kendetegner dette trin.

4. *Kyndig udøver*. Den kyndige udøver er et stort skridt hinsides analytisk rationalitet og den langsomme ræsonneren, som kendetegner de første tre trin af læreprocessen. Kyndige fodboldspillere opfatter øjeblikket for en dribling ved, at *hele* den visuelle situation foran dem og fornemmelserne i deres egen krop udløser erindringer om tidligere tilsvarende situationer, hvor driblinger er lykkedes. Der er ikke noget, der tyder på, at fodboldspillere ved hjælp af generelle regler kombinerer forskellige fakta om egen og modstanderes position, bevægelse, hastighed o.l., og derefter træffer et valg. Den kyndige udøver opfatter tilsyneladende situationer *intuitivt* og *holistisk*. Intuitionen er ikke gætteværk, irrationalitet eller overnaturlig inspiration, sådan som visse kognitivister ynder at beskrive den (for derefter at tage afstand fra den). Den er en egen-skab, ethvert menneske konstant bruger i dagligdagen. Kyndige udøvere bruger dog stadig også elementer af analytisk beslutningstagning, idet strategier, der intuitivt fremtræder som fornuftige i en givet situation, af og til vurderes i deres delelementer og efter generelle regler. Der er tale om en vekselvirkning, hvor ud-



gangspunktet er situation og intuition.

5. *Virtuos*. Med tilstrækkelig erfaring når folk frem til et niveau, hvor det ikke kun er situationer, som genkendes intuitivt og i deres helhed, men også - og samtidigt - de relevante beslutninger, strategier og handlinger. Problemer ses ikke som ét og løsninger som noget andet. Der planlægges ikke. Vi er nået til niveauet for den flydende, utvungne præstation, som ikke afbrydes af analytiske overvejelser. Det er virtuosens niveau, niveauet for *intuitiv, holistisk* og *synkronisk* ekspertise. Det er det niveau, de fleste af os befinder sig på med hensyn til at gå, stå, tale, læse og mange andre dagligdags aktiviteter. Vi er virtuose

til dagligdags forståelse og dagligdags handling, til at være dagligdags mennesker. I kontrast er der andre områder, hvor kun en brøkdelen af nybegyndere når frem til at være virtuoser, f.eks. skak, violinspil og øjenkirurgi. I skak udgøres virtuosniveauet af de såkaldte stormestre og internationale mestre, hvoraf der eksempelvis kun er cirka 70 i USA. Virtuoseres færdigheder er blevet så meget en del af dem, at de ikke er mere opmærksomme på færdighederne end på deres egen krop. Flypiloter fortæller, at de som begyndere oplevede, at *de fløj deres maskiner*, mens de som virtuoser simpelt hen oplever, at *de flyver*. Adskillelsen mellem person og maskine, subjekt og objekt, er forsvundet. Der er kvalitet i oplevelsen, det er kvalitet, der præsteres.

GENNEMGANGEN AF de fem trin er meget summarisk og er klart åben for kritik, f.eks. for overforenkling og individfokusering. Alligevel er der så meget om de fem trin, at de er værd at reflektere nærmere over i forhold til kunstig intelligens og i forhold til diskussionen om rationalitet.

Moralen er i den forbindelse, at der er mere mellem himmel og jord og mellem menneskets ører end analytisk rationalitet. Det er der imidlertid ikke mellem pladerne på de computere, vi kender i dag. Kunstig intelligens kan med den eksisterende computerteknologi ikke nå længere end til trin 3, kompetence, og befinder sig i dens kendte udgaver oftere på trin 1 og 2, nybegynder og avanceret begynder. Der er et kvalitativt spring fra de tre første til de to sidste trin. Springet består i,

at regeltænkning opgives, og menneskets intuition sættes i arbejde. Springet er en forudsætning for menneskets bedste præstationer. Det kan ikke tages af eksisterende computere. Og derfor kan de ikke simulere menneskelig intelligens, selv om de kan slå mennesket mange gange i det, maskinen er bedst til, nemlig løsning af veldefinerede og velafgrænsede opgaver. Kunstig intelligens er halv intelligens. Tilsvarende er analytisk rationalitet halv rationalitet, hvilket burde være tankevækkende, da denne rationalitetsform er blevet prist som den ypperste i vores samfund fra Sokrates til i dag.

Takket være forskningen i kunstig intelligens har Platons og Kants teorier om, at menneskets hjerne arbejder på grundlag af regler, endelig fundet deres empiriske test. Forsøget på at bringe logiske maskiner til at producere menneskelignende forståelse, viser, at den to tusind år gamle konventionelle og stadig dominerende opfattelse er utilstrækkelig.

TRODS EN SPIRENDE udvikling i kunstig intelligens forskningen i mere holistisk retning kan man med god sandsynlighed for at få ret påstå, at kunstig intelligens også fremover vil slå fejl inden for områder, som kræver kyndighed og virtuositet i udøvelsen af færdigheder. Det gælder uanset, om der er tale om professionel ekspertise på højt niveau eller om dagligdags aktiviteter. Og det gælder uanset kunstig intelligens-lobbyens påstande om, det modsatte og uanset de mange milliarder, som pumpes ind i området for tiden. Kunstig intelligens vil fortsat fejle med hensyn til

sprogforståelse, indlæring, genkendelse af tale, forståelse af fortællinger og andre områder, som reflekterer strukturen i vores dagligdags fysiske og sociale verden, og hvis beherskelse er nødvendig for virtuos adfærd.

Der, hvor kunstig intelligens vil kunne få en nyttig betydning, og allerede har det, er på områder, som kan afgrænses snævert fra dagligdags forståelse og fra forandringer. Områder som har veldefinerede problemer og klare regler for, hvordan problemerne løses. Dvs., områder præget af en stor portion rutine. Her er kunstig intelligens i rivende udvikling i de såkaldte ekspert-systemer. Det gælder inden for medicin, kemi, geologi og datalogi, hvor de mest udviklede systemer findes. Men mange andre områder er også godt i gang. Der er imidlertid ikke tale om kunstig intelligens i begrebets egentlige betydning, der dækker imitation af menneskelig intelligens som sådan. I ekspert-systemerne imiteres kun et mindre, velafgrænset hjørne af den menneskelige intelligens.

Alligevel er der også for ekspert-systemer grund til at slå koldt vand i blodet. Deres historie er kendetegnet ved en række overdrivelser af, hvad systemerne kan. Som sagt er der penge i den slags overdrivelser. Men det gør dem selvfølgelig ikke mere rigtige. Selv for ekspert-systemer kan det forudsiges, at inden for områder, hvor folk udviser holistisk og intuitiv forståelse, vil systemer baseret på heuristiske regler (de tommelfingerregler systemerne arbejder efter) aldrig bliver så gode som virtuose eksperter, også selv om det er de selv samme eksperter, der har leveret reglerne til syste-

merne. Tests af eksisterende systemer underbygger dette til dato.

Hvorfor? Fordi virtuose eksperter simpelt hen ikke bruger regler. De genkender tusindvis af specielle tilfælde direkte, holistisk og intuitivt på baggrund af deres erfaring. Det er ægte ekspertise. Reglerne til ekspert-systemerne formuleres kun, fordi systemerne kræver det. De er karakteristiske for systemerne, ikke for de virkelige eksperter. Det er grunden til, at konstruktørerne af ekspert-systemer ofte har det allerstørste besvær med at få eksperter til at formulere de regler, som er nødvendige for systemerne. Og det er grunden til, at heuristiske ekspert-systemer ikke når længere end til trin 3 i læreprocessen, kompetence.

DET BETYDER IKKE, at ekspert-systemer ikke kan være nyttige. Det kan de, punktvis. Kritikken af kunstig intelligens og ekspert-systemer giver ikke grobund for en generel maskinstormerholdning. Ekspert-systemer finder som sagt allerede i dag nyttig anvendelse på en række områder. Og selv om et system måske aldrig kan blive så godt som de eksperter, hvis kunnen det søger at simulere, kan det godt være bedre end gennemsnitspræstationen på et område. Lægen på sygehuset i X-købing, som står over for en lungesygdom, kan måske forbedre sin diagnose og behandling ved adgang til et ekspert-system, som anvender den viden, lungeeksperterne på Rigshospitalet har, eller den akkumulerede viden om lungesygdomme for den sags skyld. Det kunne være en ny form for kollektivitet: Den samlede læ-

gestands viden involveret i den enkelte diagnose og behandling.

Men der er en fare her. Så længe ekspert-systemer bruges som grundlag for og hjælpemiddel til at træffe beslutninger, men det stadig er mennesker der træffer beslutningerne, er der ikke det store problem. Hvis systemerne derimod bliver et substitut for menneskelig beslutningstagning er den gal. Og det er præcis her, skillelinjen mellem fornuftig og ufornuftig anvendelse af computere går.

Beslutninger taget af ekspert-systemer vil som sagt aldrig blive mere end kompetente, hvilket ikke altid er nok, og i nogle tilfælde kan være direkte farligt. Og hvad vigtigere er, hvis folk generelt begynder at forlade sig på ekspert-systemer, computere og andre analytiske værktøjer til at træffe afgørelserne for sig, så udvikles evnen til virtuos adfærd ikke. Menneskets største potentiale, som gør, at det kan klare sig i utroligt komplekse situationer og under vidt forskellige forhold, kort sagt dets intuition, undermineres og underudvikles, simpelt hen fordi den ikke bruges. Denne tendens til dequalificering gør sig allerede gældende en del steder i vores samfund. Med en almindelig udbredelse af computere kan tendensen i sin konsekvens medføre, at dette teknologiske fremskridt får en endnu mere fordummende virkning end fjernsynet.

Star Wars-projektet illustrerer problemerne i ekspert-systemer uhyggeligt klart. Projektet omfatter et system for computerstyret missilaffyring. Militære eksperter afgør på forhånd under hvilke omstændigheder systemet skal

fyre. Alle krisesituationer skal forudses og regler for reaktioner programmeres, så en computer derefter automatisk kan styre eventuelle reaktioner fra systemet. Fordelen ved automatikken er et hurtigt svar i tilfælde af angreb. Problemet er, at systemet, som andre ekspert-systemer, kun er kompetent. Og i en krisesituation er kompetence ganske enkelt ikke godt nok. Især ikke når det drejer sig om atomkrig. Virtuosekspertise er nødvendig. Det er dødsensfarligt at forbigå det menneskelige element, rent bortset fra, at det er dødsensfarligt at følge kaprustningens logik.

Star Wars-projektet udtrykker i sin oprindelse og i sin funktionsmåde rationalitetens ufornuft i en nøddeskal, den rationalitet, som gør, at vi i dag lever i en historisk unik periode, i en verden, som aldrig har eksisteret før, den fatalt-risikable verden. Mere om det i en følgende kronik.

To aktuelle bøger om computere og kunstig intelligens, som artiklen delvist bygger på:

Hubert & Stuart Dreyfus: »Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer«, The Free Press.

Theodore Roszak: »The Cult of Information: The Folklore of Computers and the True Art of Thinking«, Pantheon.

Bent Flyvbjerg er gæsteforsker ved University of California, Los Angeles.

Bent Flyvbjergs næste kronik, Den fatalt-risikable verden, bringes i morgen.