

HANDELSHÖGSKOLAN
vid Göteborgs Universitet
Institutionen för Informatik

Intelligenta System

- *är de att lita på?*



Examensarbete i informatik, C-nivå
Höstterminen 2001

Handledare: Göran Walske

Författare: Tony Oddhammar

Förord

För att kunna sammanställa min uppsats har jag fått hjälp av ett antal personer. Jag vill framföra ett stort tack till alla intervjuade som trots att de alla är upptagna med sina respektive arbeten tog sig tid att besvara mina frågor. Dessutom vill jag tacka min handledare Göran Walske som har agerat bollplank och trott på möjligheten att genomföra projektet.

Göteborg den 5 januari 2002

Tony Oddhammar

Abstract

Idag har det blivit allt vanligare att våra vägar korsas av s.k. intelligenta system. Tanken med dessa system är att de skall hjälpa oss med våra vardagliga sysslor på jobbet, i hemmet m.m. Idag när vi alltmer bygger in intelligenta funktioner i maskiner, robotar och i datasystem anser jag att man måste utreda frågorna kring konsekvenserna kring bristfälligt hanterande och för stor tilltro till sådana här system.

- Litar vi för mycket på systemen?
- Bör vi lita på dem?
- Och vem skall ställas till svars för misstagen systemet gör!

Syftet med denna uppsats är att undersöka vad forskare inom området tycker om vi som människor blir alltmer beroende av Intelligenta System och om det finns någon fara i att förlita sig på dessa system för mycket?

Koncentrationen av utredningsarbetet kring artificiell intelligens och speciellt intelligenta system, där misstagen kan få ödesdigra konsekvenser för tredje part i sammanhanget. Jag begränsar också omfånget av uppsatsen genom att utelämna att beskriva hårdvaran till intelligenta system. Intressenterna för uppsatsen är alla de som intresserar sig för området AI och den effekt den kommer att få på oss människor, lekman som expert.

I mitt fall har jag valt att göra en kvalitativ undersökning med djupintervjuer av personer som arbetar med intelligenta system. ett antal frågor med öppna svarsalternativ som jag valde att e-posta till ett antal människor. De intervjuade personerna valdes endast ut på kriterierna att de skulle jobba inom området AI.

Att inte kontrollera systemen och helt förlita sig på dem även om de har sk. intelligens anser jag vara riskfyllt, för även om intelligensen blir mer avancerad är den ändå likgiltig!

Naturligtvis behöver man inte övervaka ett system som Deep Blue som om det gör ett misstag som värsta konsekvens får att den förlorar sin schackmatch utan system som ex. genomför en diagnos av en patient och därmed tredje part kan drabbas av misstagen bör övervakas av en mänsklig expert som bedömer rimligheten i diagnosen, i fall att! Min uppsats kanske ger oss en tankeställare om hur vi vill att våra intelligenta system skall se ut i framtiden och hur vi skall hantera användandet av dem.

Jag skulle tycka att det skulle vara intressant om någon skulle ta reda på och skriva om ett ”real case scenario” där vi som människor förlitat oss för mycket på ett intelligent system. Exempelvis den tragiska bombningen av Kinesiska ambassaden i Jugoslavien.

Innehållsförteckning

Förord	2
Abstract	3
Innehållsförteckning	4
1. Inledning	6
1.1 Bakgrund	6
1.2 Problem	7
1.3 Syfte	7
1.4 Avgränsningar	8
2. Metod	8
2.1 Metoddiskussion	8
2.2 Tillvägagångssätt	9
2.2.1 Teoretisk referensram	9
2.2.2 Intervjuer	10
2.2.3 Källkritik	10
3. Vad är Artificiell Intelligens?	12
3.1 Vad menar vi med intelligens?	12
3.2 Artificiell intelligens	13
3.2.1 Turing's idéer	13
3.2.2 Moderna synen på AI	14
4. Intelligenta System	15
4.1 Kunskapsbas	16
4.2 Intelligenta tekniker	17
4.2.1 Neurala nätverk	17
4.2.2 Genetiska algoritmer	19
4.2.3 Fuzzy Logik	20
4.2.4 Expert system	21
4.2.5 Intelligenta hybrid system	22
4.3 Användningsområden för intelligenta system	23
4.3.1 Intelligenta System och Finansmarknaden	23
4.3.2 Intelligenta System och Sjukvård	23
4.3.3 Intelligenta och Industrin	24
5. Resultat	24
5.2 Resultat av intervjuerna	24
6. Analys, Diskussion och slutsats	29
6.1 Analys och diskussion	29

6.2 Slutsats	32
6.3 Förslag till vidare studier	33
8. Referenslista	34

Bilagor

Appendix I
Appendix II

1. Inledning

1.1 Bakgrund

Idag har det blivit allt vanligare att våra vägar korsas av s.k. intelligenta system. Tanken med dessa system är att de skall hjälpa oss med våra vardagliga sysslor på jobbet, i hemmet m.m. Dessa intelligent system bygger på idén att de själva skall kunna tänka och resonera likt en människa dvs. de skall ha en inbyggd konstgjord intelligens s.k. artificiell intelligens.

Artificiell intelligens har varit ett erkänt forskningsområde sedan i mitten på 1950-talet då det hölls en konferens kallad Dartmouth konferensen. På denna konferens användes termen artificiell intelligens för första gången. Innan dess var forskningsområdet endast drivet av entusiaster.

Forskningen inom området ökade likt snöbollseffekten efter Dartmouth konferensen och det dröjde inte länge förrän man hade ett flertal olika tekniker att framställa artificiell intelligens.

Under 1970- och 1980-talen så implementerades en hel del intelligenta robotar inom industrin bl.a. inom övervakning av löpande band samt till att sköta lager sysslor.

Och en dag så hände det: den 4 juli 1981 skulle Kenji Urata på Kawasaki Heavy Industries reparera en av maskinerna på växellådsavdelningen. Han beträdde den s.k. förbjudna zonen för att göra ett försök att reparera maskinen utan att behöva stänga ner produktionen. När han stod med ryggen till under reparationsarbetet märkte han inte att en robot närmade sig bakom hans rygg. Roboten slog honom upprepade gånger innan den krossade honom till döds¹.

Detta fall blev känt över hela världen eftersom människor då fick för sig att robotar var intelligenta onda ting som var ute efter att krossa människans samhälle men idag vet vi bättre. Olyckan inträffade mer på grund av maskinens brist på intelligens än tvärtom. Dagens forskare anser att om maskinen istället varit intelligent hade den märkt att något varit fel och slutat med sina desperata försök att nå fram till produktionslinan.

Även om vi har långt kvar till situationen att ett av våra mest användbara redskap kan bli en jämbördig part och kanske just därför, måste vi vara försiktiga i användandet av intelligenta system. Människor får inte överge eller glömma sin roll som slutgiltiga bedömare eller övervakare av systemen och deras likgiltiga intelligens.

¹ Rothfeder F, s.155

1.2 Problem

Mitt första möte med begreppet ”Artificiell Intelligens” var när jag i ung ålder såg filmen ”2001 – Ett rymdäventyr”² av Stanley Kubrick. Filmen handlar om en grupp människor som är på en rymdfarkost/-station. Stationen har ett inbyggt intelligent system som kallar sig HAL. Det visar sig att HAL inte bara är ett system som kan hjälpa till med problemlösning mm. utan även har känslor och medvetande. I filmen börjar besättningen på rymdfärjan känna sig motarbetade och hotade av HAL. De får för sig att stänga av systemet och då försvarar sig HAL genom att börja döda besättningsmedlemmarna.

Detta är naturligtvis bara Science Fiction, det är långt kvar till vi kan skapa en dator med ett medvetande om någonsin. De typer av intelligenta system som vi lyckats skapa och används idag består mest av relativt enkla analys- och problemlösningsapplikationer och kan inte på något sätt jämföras med HAL.

- Vem skall stå till svars när HAL dödar?

Idag när vi alltmer bygger in intelligenta funktioner i maskiner, robotar och i datasystem anser jag att man måste utreda frågorna kring konsekvenserna kring bristfälligt hanterande och för stor tilltro till sådana här system.

I våran verklighet är frågan vem skall stå till svars när systemen gör fel som får allvarliga konsekvenser?

- Hur mycket bör exempelvis de personerna som sysslar med fondförvaltning till yrket och ansvarar för mycket pengar lita på den tekniska analysen som datasystemet gör? Eller läkaren som kör sina provsvar genom någon typ av beslutstödsystem kan han lita på resultatet?

Jag har av moraliska och etiska skäl svårt att acceptera att vi skall överlåta viktiga beslut helt och hållet åt så kallade intelligenta system. Så frågorna som dyker upp är:

- Litar vi för mycket på systemen?
- Bör vi lita på dem?
- Och vem skall ställas till svars för misstagen systemet gör!

1.3 Syfte

Syftet med denna uppsats är att skaffa förståelse för området Artificiell intelligens samt att undersöka vad forskare inom området tycker om vi som människor blir alltmer beroende av Intelligenta System och om det finns någon fara i att förlita sig på dessa system för mycket?

² Org.titel ”2001: A Space Odyssey”, bygger på boken med samma namn av Arthur C Clarke

1.4 Avgränsningar

Jag kommer att lägga koncentrationen av mitt utredningsarbete kring artificiell intelligens och speciellt intelligenta system, där misstagen kan få ödesdiga konsekvenser för tredje part i sammanhanget.

Avgränsning kommer också att ske genom att avstå från att intervjua ”vanliga människor” d.v.s. personer som inte kan någonting om artificiell intelligens och intelligenta system eftersom de naturligtvis är antingen odelat fascinerade eller skrämde av fenomenet. Genom att välja personer som vet vad begreppen innebär har de antagligen både negativ som positiv bild av sådana här system.

Jag begränsar också omfånget av uppsatsen genom att utelämna att beskriva hårdvaran till intelligenta system. Jag tänker heller inte bli för teknisk och specifik i min beskrivning av de intelligenta teknikerna utan tänker hålla det på en nivå så att också en person som ej är insatt i området skall kunna få ut någonting av att läsa den. Det betyder att jag anser att intressenterna för uppsatsen är alla de som intresserar sig för området AI och den effekt den kommer att få på oss människor, lekman som expert.

2. Metod

2.1 Metoddiskussion

Valet av metod bör alltid utgå från problemet man definierar. Den metod man väljer kan vanligen knytas till vilken syn man har på vetenskap och därmed även till inställningen man har till sina medmänniskor och det samhälle de lever i.

Det finns två huvudinriktningar dessa är kvalitativ- och kvantitativmodell. Kvantitativa metoden är präglad av kontroll från forskarens sida dvs. undersökningen är mycket formulerad och strukturerad. Kvantitativa metoden präglas också av att den anger vilka svar som är tänkbara och är speciellt användbar vid undersökningar av statistisk karaktär.

I mitt fall har jag valt att göra en kvalitativ undersökning med djupintervjuer av personer som arbetar med intelligenta system. Varför jag använder mig av just denna metod är att den kvalitativa metoden karaktäriseras av att man *försöker nå förståelse hos olika individer*³. Målen för den kvalitativa forskningen innebär att söka, förstå och beskriva hur olika sammanhang hänger ihop. Metoden används främst för forskning inom human- och samhällsvetenskap.

³ Hartman J. s.239

Den kvalitativa metoden utgår från ett hermeneutiskt synsätt som just uttrycker förståelse för individer eller en grupp individer genom att beskriva de föreställningar de har om sig själva och den situation de lever i⁴. Den hermeneutiska cirkeln eller spiralen innebär att man genom att först delarna också förstår helheten.

En kvalitativ undersökning är till skillnad från en kvantitativ subjektiv genom att det är forskarens tolkningar utav svaren som utgör resultatet i forskningen. I och med att denna metod är subjektiv kan den inte observeras och kvantifieras, utan den måste förstås genom att forskaren sätter sig in i de olika människornas situation.

Valet av den kvalitativa metoden gjordes på grund av att jag anser att problemet är utformat på det sättet att endast en kvalitativ undersökning är intressant också för att lämna utrymme för de intervjuade att förklara sina synpunkter på mina frågeställningar och att ge dem möjlighet att ge mig ökad förståelse inom forskningsområdet.

Mitt angreppssätt måste anses som analytiskt med en induktiv ansats vilket betyder att jag under min insamling av data inte använder mig av några teorier. Detta görs för att undvika att påverka de intervjuade i någon riktning.

Det är viktigt att den problemställning jag har för avsikt att undersöka uppvisar validitet, det innebär att ställa rätt frågor som överensstämmer med syftet för min undersökning, görs inte det så fås inte korrekta svar och undersökningen blir därmed missvisande.

Förutom validitet skall en undersökning också uppvisa reliabilitet som innebär pålitlighet. Finns reliabilitet så skall en undersökning kunna uppvisa samma resultat vid upprepade undersökningar av samma slag. I denna uppsats kan kanske reliabiliteten ifrågasättas på grund av mitt begränsade antal lyckade intervjuer.

2.2 Tillvägagångssätt

2.2.1 Teoretisk referensram

Datainsamling kan ske på flera olika sätt bl.a. litteratursökning på bibliotek, sökning av Internet, enkäter intervjuer samt andra observationer. *De data man samlar in till sin undersökning kan vara av två slag, data som redan finns (sekundärdata) och data som måste samlas in (primärdata)*⁵. Ofta är det lämpligt att man börjar undersökningen med data som redan insamlats av andra.

Jag började min undersökning genom att leta igenom litteratur och tidskrifter på biblioteket samt Internet för att få en överblick av problemområdet och kunna bygga mig

⁴ Hartman J. s.240

⁵ Wiedersheim-Paul F, s.76

en teoretisk referensram samt skaffa mig en grund till intervjuerna. Jag fick även tips av en del intervjuade av vad jag kunde hitta mer information om ämnet någonstans (oftast egen författade artiklar).

Referensramen delades upp i två delar dels Artificiell Intelligens där jag beskriver bakgrunden och utvecklingen bakom fenomenet AI och dels Intelligenta System där jag talar om det specifika området IS vilket större delen av utvecklingen av AI ligger i idag. Jag beskriver bl.a. vilka de vanligaste teknikerna är samt användningsområdena för IS.

2.2.2 Intervjuer

Ofta när man genomför en sådan här uppsats är man tvungen att genomföra enkäter, intervjuer eller båda delarna för att få svar på sin problemställning. Valet av metod beror på problemets karaktär. Det finns ett antal intervju- och enkätmetoder bl.a. besöksintervjuer, telefonintervjuer, postenkäter och e-post intervjuer.

När kände mig tillräckligt teoretiskt förberedd så författade jag ner ett frågeformulär, ett antal frågor med öppna svarsalternativ som jag valde att e-posta till ett antal människor vars adresser jag fick tag i via en s.k. resurs sida på nätet kallad ”AI on the Web⁶”. Där kunde jag hitta e-post adresser till i stort sett varje känd forskare inom området. De intervjuade personerna valdes endast ut på kriterierna att de skulle jobba inom området AI.

E-post intervjuer har vissa fördelar och några av dessa är möjlighet till geografisk bredd, låg kostnad, ingen intervjuareffekt (dvs. intervjuaren styr svaren med sin närvaro) osv. Nackdelar är naturligtvis att det kan vara svårt att följa upp frågor och att det finns risk för låg svarsfrekvens.

Eftersom de intervjuade har en möjlighet att besvara frågorna när de vill och om de vill med denna metod tror jag att de känner sig mer öppna till att faktiskt svara än om man tvingar sig till en besöks- eller telefonintervju som kan vara tidskonsumerande och måste planeras in.

2.2.3 Källkritik

Det finns ett antal källkritiska kriterier. Några av dessa är⁷:

- Samtidskrav, dvs. att om någon återger någon fakta ska den vara färsk och inte kunnat förvrängas med tiden.
- Tendenskritik, dvs. vilka egna intressen har uppgiftslämnaren.
- Beroendekritik, dvs. källorna skall vara oberoende av varandra

⁶ <http://http.cs.berkeley.edu/~russell/ai.html>

⁷ Wiedersheim-Paul F, s.83

Mycket av min undersökning bygger på sekundärdata och jag måste säga att litteraturen varierar mycket i kvalitet. Eftersom det finns en begränsad mängd böcker inom området har det varit svårt att få tag i referenslitteratur av genomgående hög kvalitet och därför är referensramen också uppbyggd på ett fåtal böcker. Jag måste dock säga att tillgången på information över nätet har varit mer än tillfredsställande, men som ofta är fallet med Internet är att materialet är svårarbetat och har därmed mest använts till kortare inlägg.

Intervjuerna gjordes med forskare som har aktuell kunskap i området och måste därför anses klara samtidskravet, dock kan tendenskritik ges eftersom problemområdet är forskarnas arbetsområde kan de ha en tendens till att skydda det eller har en övertro på det och därmed får vi kanske omedvetet en snedvridning i intervjuvaren. Och därmed kan reliabiliteten i intervjuerna ifrågasättas.

3. Vad är Artificiell Intelligens?

Termen Artificiell Intelligens (AI) dök upp första gången under Dartmouth konferensen⁸ 1956. Idéerna kring området dök dock upp mycket tidigare, man har kunnat spåra dem ända till forntida Egypten genom folksägnen⁹. Dartmouth konferensen handlade mycket om att marknadsföra området som kommit att kallas AI och på så sätt kunna få ökade anslag till forskning inom området.

Artificiell intelligens är ett tvärvetenskapligt ämne, d.v.s. en kombination av vetenskaper som datavetenskap, fysiologi och filosofi. För att kunna avgöra om en dator/system är intelligent eller inte måste vi definiera vad som är att betrakta för intelligent beteende.

3.1 Vad menar vi med intelligens?

När vi pratar om intelligens vad menar vi egentligen då? Är det en förmåga att lösa komplexa problem eller är det att göra generaliseringar och upptäcka samband. Det finns ett stort antal teorier om vad det är för egenskaper som särpräglar intelligent beteende. De flesta forskarna har sina egna teorier kring intelligens, det finns de som säger att förmåga att kommunicera är intelligens andra säger att det är självmedvetandet som avgör om en varelse är intelligent.

En exakt definition av intelligens är antagligen är omöjligt men det bästa försöket är antagligen det här:

- En förmåga att handskas med komplexa frågor och lösa problem som har någon typ av betydelse, vare sig det är lösningen på en komplicerad ekvation eller bara en rörelse för att sträcka sig efter lite frukt¹⁰.

Intelligens är enligt nationalencyklopedin definition¹¹:

- I dagligt tal detsamma som förstånd, begåvning eller tankeförmåga.

Den definitionen av intelligens betyder att de flesta handlingar vi människor utför under en dag kan betraktas som intelligenta även de som styrs av våra drifter eller vårt undermedvetna. Intelligens skapar snabba, tillfredställande och ibland överraskande lösningar på problem.

⁸ Konferens som anses som startskottet för det vi idag kallar för AI.

⁹ <http://library.thinkquest.org/2705/history.html>

¹⁰ <http://www.sciam.com/specialissues/1198intelligence/1198yam.html> , Philip Yam

¹¹ Nationalencyklopedin, internet upplagan www.evreka.com

Men om vi skall tala om den typen av intelligens som det kan finnas en anledning till att duplicera bör det vara en typ av intelligent beteende vilket uppfyller ett eller flera av kraven nedan¹².

- Lärande och förståelse från erfarenhet
- Förstå tvetydiga eller motsägelsefulla meddelanden
- Svara snabbt och rätt till en ny situation
- Använda ett förnuftigt resonemang till att lösa problem effektivt
- Handskas med överraskande situationer
- Förstå och dra slutsatser på ett vanligt rationellt sätt
- Använda sin kunskap till att påverka omgivningen
- Tänka och resonera
- Förstå olika objekts relativa betydelse i en situation

3.2 Artificiell intelligens

3.2.1 Turing's idéer

Alan Turing anses av många som en av de banbrytande forskarna inom AI trots att hans död inträffade innan termen AI fanns. När andra världskriget bröt ut var Turing var en ung och lovande forskare vid Cambridge Universitet. Han var mest känd för sina matematiska kunskaper. Han fick rollen som chifferexpert under kriget och lyckades skapa en apparat som kunde dechiffrera den tyska chifferingsapparaten Enigma's kod. Tack vare Turing's apparat kunde Engelmännen förhindra Tyskarna att nå framgångar i Storbritannien. Detta var det första steget mot det vi idag kallar AI.

Turing var övertygad om att Artificiell Intelligens skulle bli verklighet någon gång efter kriget var slut. Han menade om man kunde bryta ned hjärnans funktionssätt i dess minsta detaljer skulle en datormodell av den kunna utformas. Detta gav en klar syn på hur AI skulle kunna framställas.

Turing anses idag av de flesta forskare som litet för naiv i sin beskrivning av den mänskliga hjärnan och dess egenskaper eftersom han ansåg den vara likt en maskin uppbyggd av ett antal olika mindre och lätt särskiljbara processer.

Han hade också en idé om hur man skulle kunna veta när man hade lyckats framställa AI, det så kallade Turing testet. När en mänsklig intervjuare ställer frågor till en människa samt en dator som är placerad bakom ett skynke och genom de respektive svaren inte kan avgöra vilket av svaren som kommer ifrån människan respektive datorn kan datorn ses som intelligent.

¹² Turban E, Aronson J E, s. 398 fritt översatt

Även detta måste anses naivt. Bara för att en intervjuare inte kan skilja en dator från en människa i en intervju situation betyder ju inte att datorn är intelligent eller som den berömde AI forskaren Marvin Minsky har uttryckt sig angående Deep Blue¹³.

-Deep Blue kanske kan vinna i schack men den vet inte hur den skall ta sig in när det börjar regna ute¹⁴!

Till Turings försvar måste man säga att han utvecklade sina idéer under en tid då man varken hade den insynen i människans neurofysiologi man har idag och inte heller hade de första moderna datorerna konstruerats än. Därför hyser de flesta moderna AI forskare ändå största respekten för hans idéer och tankar.

3.2.2 Moderna synen på AI

Problemet med den tidiga synen på AI är att man försökte skapa allsidiga system. Tanken var att ett och samma system skulle klara av att lösa en mängd olika problem precis som en människa som exempelvis både kan snickra och laga mat. Nuförtiden vet man att all kunskap man vill att systemet skall innehålla måste vara tydligt förklarad för systemet. De första konstruktörerna av Intelligenta maskiner brydde sig inte om att skapa regler som ”En far som har en son, är äldre än sin son och kommer att vara det hela sitt liv”¹⁵. Detta är kanske inte så konstigt eftersom detta är självklart för oss som människor och det är någonting vi tar för givet att alla vet och därför aldrig behöver förklara det för någon.

När detta problem blev bekant tog man ett jättekiv i utvecklingen av IS. Istället för att skapa system som var universella problemlösare började man utveckla system som imiterar en vald del av ett mänskligt beteende. Det har lett till att AI forskning idag är uppdelat i ett stort antal domäner, några exempel är¹⁶:

- o **Vardagliga domäner**
 - *Varseblivning*
 - Syn
 - Hörsel
 - *Naturligt språk*
 - Förståelse
 - Generering
 - Översättning
 - *"Sunt förnuft"*
 - *Robotkontroll*
- o **Formella domäner**
 - *Spel*
 - Schack
 - Backgammon

¹³ Dator som lyckades slå världsmästaren Gary Kasparov i schack

¹⁴ Kurzweil R, s. 90 fritt översatt

¹⁵ <http://psych.utoronto.ca/%7Eereingold/courses/ai/>, fritt översatt

¹⁶ exemplen är tagna från http://www.csd.uu.se/kurs/ait/ait96/lectures/vad_ar_AI.shtml

- Dam
- Go
- *Matematik*
 - Geometri
 - Logik
 - Integralkalkyl
 - Bevisning av programegenskaper
- o **Expertdomäner**
 - *Ingenjörsuppgifter*
 - Design
 - Fellokalisering
 - Tillverkningsplanering
 - *Vetenskaplig analys*
 - *Medicinsk diagnos*
 - *Finansanalys*

De flesta AI forskare koncentrerar sig på en eller ett par av dessa domäner eftersom man använder särskilda tekniker inom varje domän för inläring mm. Därför kommer antalet AI domäner säkerligen växa allt eftersom man hittar nya användningsområden.

4. Intelligenta System

Ett av de största nyttoområdena för AI är Intelligenta System dvs. system som kan hjälpa oss människor att analysera och lösa komplicerade problem snabbare och bättre. Trots att AI's slutgiltiga mål är att imitera mänsklig intelligens är dagens Intelligent system långt ifrån att vara i närheten av detta. Dock skall man komma ihåg att AI ständigt utvecklas och har redan ökat produktiviteten och kvaliteten genom att automatisera många uppgifter som normalt kräver någon form av mänsklig intelligens¹⁷.

Låt oss då se på vad man behöver för att skapa ett Intelligent System¹⁸.

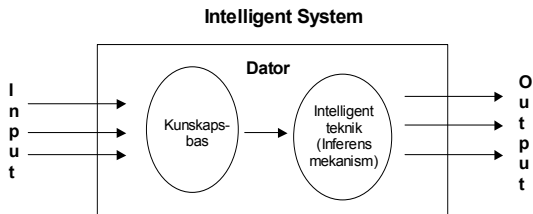
1. **Hårdvara för beräkning.** Någon typ av dator som kan utföra de avancerade beräkningarna som en intelligent teknik kräver.
2. **Kunskap** om problemområde. Denna kan samlas in automatiskt med hjälp av en intelligent teknik(neurala nätverk, genetiska algoritmer) eller genom manuella metoder som intervjuer av en expert inom området
3. **Intelligent teknik**¹⁹. Det vill säga en metod för systemet att sortera intrycken den har blivit programmerad att bevaka och dra slutsatser.

¹⁷ Turban E, s. 399

¹⁸ Kurzweil R, s. 101 fritt översatt

¹⁹ beskrivs i stycke 4.2 av uppsatsen

Naturligtvis krävs också ett användargränssnitt för eventuella användare men det känns som självklart och kommer inte beskrivas mer nedan.



4.1 Kunskapsbas

Kunskap som har insamlats för att användas i ett intelligent system kallas för en kunskapsbas och intelligenta system behöver en kunskapsbas. Kunskapsbasen innehåller den nödvändiga kunskapen för att förstå, formulera och lösa problem. Kunskap kan representeras på flera sätt och det finns ett antal olika kunskapskategorier som man bör skilja på bl.a.²⁰

- Procedurkunskap, - regler, strategier, procedurer
- Deklarativ kunskap, - koncept, objekt, fakta
- Metakunskap, - kunskap om kunskap och hur den skall användas
- Heuristisk kunskap, - tumregler

De olika typerna av kunskap används i olika typer av system. Vissa system använder två eller flera av kategorierna ovan. Kunskapen kan antingen inhämtas manuellt eller automatiskt. Med manuell inhämtning menar man när en s.k. kunskapsingenjör samlar in data (ex. från experter på området). Automatisk inhämtning är när ett intelligent system sköter inhämtningen av kunskapen.

Det viktigaste för kunskapsinhämtningen är att den s.k. kunskapsingenjören har de egenskaperna som gör att han/hon lyckas skapa den kunskapsbasen som är nödvändig för att lösa det specifika problemet.

Kunskapen kan inhämtas ifrån mängder av källor som böcker, filmer, databaser, bilder, diagram, sensorer mm. Kunskapen i kunskapsbasen kan delas upp i två områden, för det första fakta och teorier kring problemområdet, för det andra regler som vägleder användningen av kunskapen för att lösa ett specifikt problem inom ett problemområde.

²⁰ Föreläsninganteckningar till kurs i Beslutstödsystem, Faramarz Agahi

4.2 Intelligenta tekniker

Hjärnans neurofysiologi kommer att vara ett område som gäckar forskare och vetenskapsmän under lång tid framöver men forskarna av idag har utvecklat ett antal tekniker som var och en på sitt sätt simulerar ett mänskligt tänkande. Dessa teknikerna har olika styrkor och svagheter såväl som sina egna specialområden för problemlösning.

När väl kunskapen är insamlad behövs ett program som använder sig av kunskaperna och drar slutsatser. Ett sådant här program består bl.a. av en algoritm som styr den s.k. resonemangsprocessen och kallas normalt inferens mekanismen (slutledningsmekanism). Dessa algoritmer försöker imitera människans sätt att resonera och lösa problem²¹.

Det här är några av de egenskaper intelligenta tekniker använder sig har²²:

- Sökning
Ingen eller bara en partiell, direkt metod för att lösa problemet finns, så i stället söker sig problemet genom en lösningsrymd
- Kunskapsanvändning
Löser komplexa problem genom att utnyttja strukturen hos de ingående objekten
- Abstraktion
Tillhandahåller ett sätt att separera ut de faktorer som är intressanta för lösningen från alla de detaljer som annars skulle överväldiga en process

4.2.1 Neurala nätverk

Den mänskliga hjärnan är uppbyggd av en massa celler som kallas neuroner. Dessa celler är grupperade och uppdelade i så kallade nätverk. Varje nätverk innehåller tusentals neuroner som är sammankopplade. Därför kan hjärnan ses som en samling nätverk. Dessa celler dör inte som andra celler gör och det är antagligen därför vi har ett minne. Hjärnan innehåller 50 till 150 miljarder sådana här celler och de finns ett hundratal varianter på dem²³.

Neurala nätverk är en intelligent teknik som är baserad på idén att försöka imitera funktionen av nervcellerna (neuroner) som finns i den mänskliga hjärnan. I kontrast till vanliga datorprogram där all problemlösning sker i en steg för steg process kan ett neuralt nätverk tränas till att lösa specifika problem.

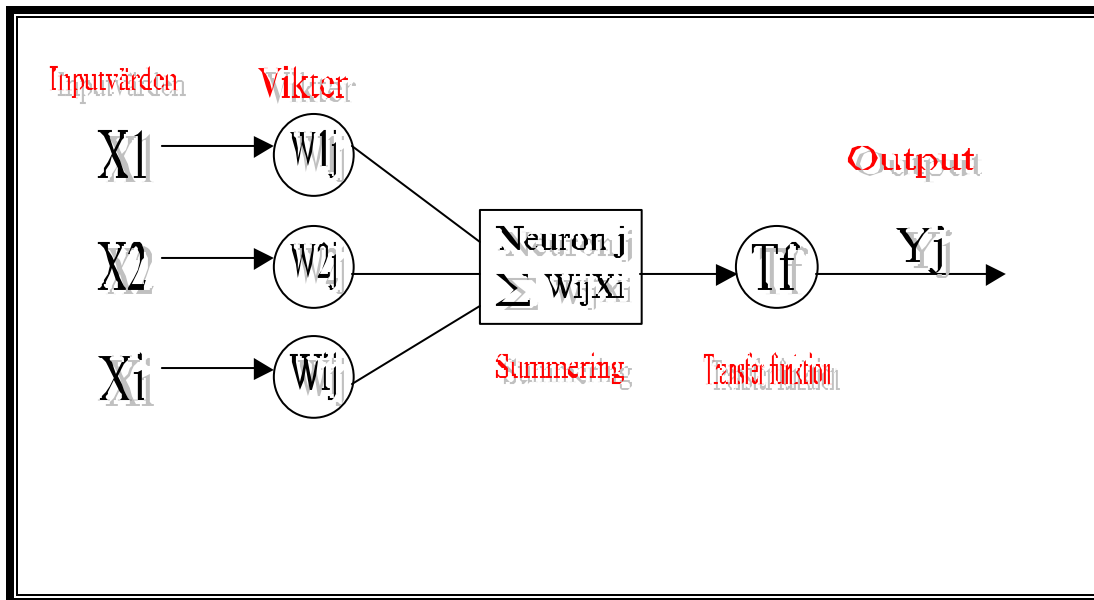
²¹ Turban E, s. 511

²² http://www.csd.uu.se/kurs/ait/ait96/lectures/vad_ar_AI.shtml

²³ Turban E, s. 607

När ett neuralt nätverk får ett problem att lösa får den i sig ett antal input värden vart och ett av de här inputvärdena svarar för ett speciellt attribut (likt de intryck en människa tar till sig när den skall lösa ett problem). De här inputvärdena processas genom att viktas och skickas vidare till en s.k. nod (likt de elektrokemiska processer som pågår i hjärnan när vi löser problem, noden kan liknas vid en neuron cell).

I noden summeras värdena och vi får ett outputvärde som jämförs med ett ”önskat” värde om det är lika så stämmer det in i det ”önskade” mönstret och nätverket känner igen sig. Se bild nedan²⁴:



Likt den mänskliga hjärnan kan man lära upp det neurala nätverket till att ta riktiga beslut. Lärandet sker genom en algoritm. Det finns hundratals algoritmer att välja på som alla har sina styrkor respektive svagheter.

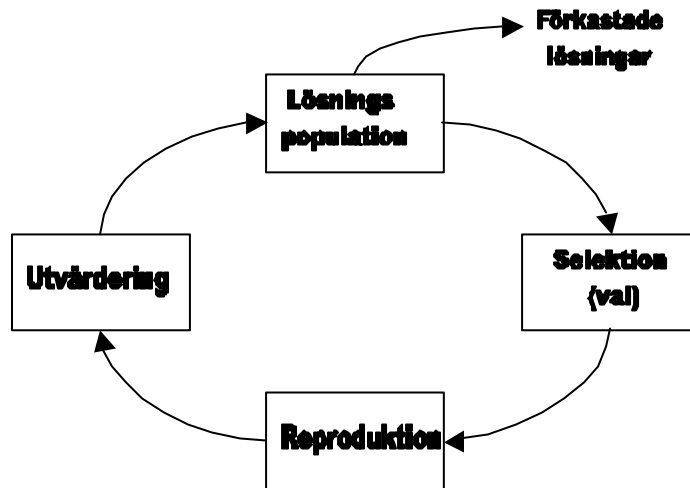
Neurala nätverk är speciellt effektivt när det kommer till mönsterigenkänning, lärande och klassificering. Dessa egenskaper gör att neurala nätverk är speciellt lämpat att använda för att t.ex. upptäcka kreditkortsbedrägeri(ändrade köpvanor), sortera post(känna igen handstilar) m.m.

Det finns också en hel del nackdelar med neurala nätverk varav den viktigaste är att de saknar förmågan att visa/förklara hur de lyckas komma fram till respektive lösning.

²⁴ Inspirerad av Turban E, s.610

4.2.2 Genetiska algoritmer

Genetiska algoritmer är inspirerade av läran om evolutionen av biologiska mekanismer och är en effektiv problemlösningsmetod. Idén bygger på att man startar med en samling av lösningar till ett problem, behåller de bra lösningarna och förkastar de dåliga och använder sedan de bra lösningarna till att skapa nya generationer av lösningar, och så börjar utsorteringsprocessen om igen. Se bilden nedan²⁵:



Den Genetiska Algoritm Cyclen

Denna process liknar mycket Darwins teorier om ”den starkes överlevnad” dvs. att endast de bäst anpassade klarar av att överleva i vissa givna förhållanden. Den första generationen är normalt inte alls bra på att lösa problemet!

Reproduktionen av de nya generationerna görs genom de bra lösningarna från den tidigare generationen blir ”föräldrar” till nästa generation genom att en del av en lösning från föräldergenerationen sätts samman med en del av en annan lösning från föräldragenerationen och därmed har man dels skapat en ny generation av lösningar varav vissa har fått föräldrarnas goda egenskaper medan andra fått de dåliga.

Genetiska algoritmer har visat sig vara särskilt användbart vid genomsökning av stora mängder av data, speciellt för att hitta optimala lösningar t.ex. vid schemaläggning. Ex. planera när man skall vilka program visa för att maximera reklamintäkter för TV-kanaler eller optimering av investeringsportfölj. Genetiska algoritmer kan lära sig komplexa samband och kan användas som ett verktyg vid ”data-mining”²⁶ för att upptäcka tidigare okända mönster.

²⁵ Inspirerad av Goonatilake S, s.12

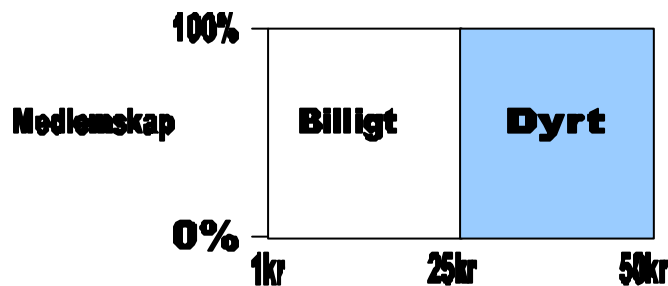
²⁶ när man letar tidigare okänd kunskap i databaser.

Nackdelar med denna teknik är att den konsumerar tid genom sin omfattande iterativa process.

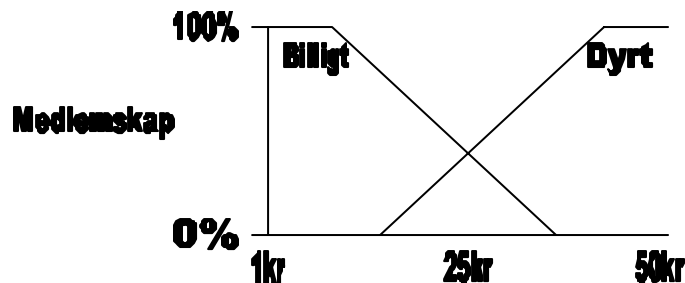
4.2.3 Fuzzy Logik

Fuzzy logik är en intelligent teknik framtagen för att kunna behandla data med oexakta lösningar, dvs. att det inte går att besvara ett problem med ja/nej, rätt/fel, sant/falskt osv. Exempel på sådana data är stor - liten, ung - gammal, hög - låg mm. Vad bestämmer t.ex. om en person är gammal eller ung, lång eller kort....

Tekniken är användbar för att likna mänsklig perception och beskriva ”grå” områden när allt inte går att besvara svart på vitt. Ta som exempel när du är i butiken och tittar på kilo priset på tomater: - oj vad dyrt de kostar 25 kr idag, tänker du. Medan nästa person som plockar buffrar för den tyckte det var billigt. Så var då tomaterna dyra eller billiga? Det går inte en skarp gräns på 25 kr så att om priset på tomater ligger under är billiga och över är dyra. Se bild nedan²⁷:



Båda två håller dock säkert med om att tomater för 20 kr/kg är billigt och 30 kr/kg är dyrt men det finns inte en klar linje mellan billigt och dyrt därför hamnar allting i mellan i en gråzon. Det är sådana här begrepp som Fuzzy logik reder ut. Så tomater för 22 kr/kg skulle vara mer medlem till klassen billig än dyr och tomater för 28 kr/kg tvärtom skulle vara mer dyr än billig men inte fullvärdiga medlemmar av någon klass. Se bilden nedan²⁸:



²⁷ Inspirerad av Goonatilake S, s.15

²⁸ Inspirerad av Goonatilake S, s.16

Vad Fuzzy logik gör är att den klassificerar begreppen billigt och dyrt och de tomater som i det här fallet inte direkt kan klassificeras in i de begreppen blir medlem i båda grupperna till olika grader.

Fuzzy resonemang är den process som sker när det intelligenta systemet försöker dra slutsatser ur s.k. Fuzzyfierad data dvs. data som inte direkt tillhör någon klass.

Styrkan med Fuzzy logik är precis det som beskrivits ovan, förmågan att behandla oprecis data detta är speciellt användbart vid beslutsfattande system t.ex. system som letar köpvärda aktier eller system som reglerar temperaturen.

En av de stora nackdelarna med tekniken är att man måste specificera alla funktioner och regler som bestämmer medlemskap manuellt. Detta kan vara mycket tidsödande.

4.2.4 Expert system

Expert system var de tidigaste och är idag den mest etablerade typen av intelligenta system. Expert system är nerkortning på termen ”kunskaps baserade expert system”²⁹. Som namnet anspelar på är det ett försök att likna en mänsklig experts resonemang vid lösandet av vissa specifika problem.

Expert system kan användas av både noviser som experter inom systemets expert område. Experterna kan använda det som en ”kunskapsassistent” men normalt används det av noviser som ersättning för en experts kunskap. Sådana här system fungerar ofta bättre än en mänsklig expert, mycket på grund av att sådana här system håller en hög jämn nivå i sitt beslutsfattande till skillnad från en expert som kan bli trött och ta felaktiga beslut.

Expertsystem syftar till att använda problem specifik kunskap som finns samlad i en kunskapsbas genom en samling regler i s.k. IF-THEN form, dvs. om en visst omständighet råder så gäller en viss slutsats, ett enkelt exempel: IF rinnande näsa AND feber THEN förkyld. Regel basen byggs upp med hjälp av kunskap från en eller flera experter.

Experts systemens styrkor är bl.a. snabbare beslutsprocess än mänskliga experter, jämnare kvalitet på slutsatserna och möjlighet att få tillgång till normalt svårtillgänglig och dyr kunskap samt att man kan följa resonemanget som ledde till slutsatsen. Systemen är speciellt användbara vid diagnostisering (medicinsk, mekanisk m.m.) och planering (personalrekrytering, strategisk planering m.m.).

Svagheterna är att expertsystemen inte har någon inlärningsmekanism därför ställer det höga krav på experterna samt systemutvecklarna som skall få in experternas kunskap i

²⁹ Turban E, s.402

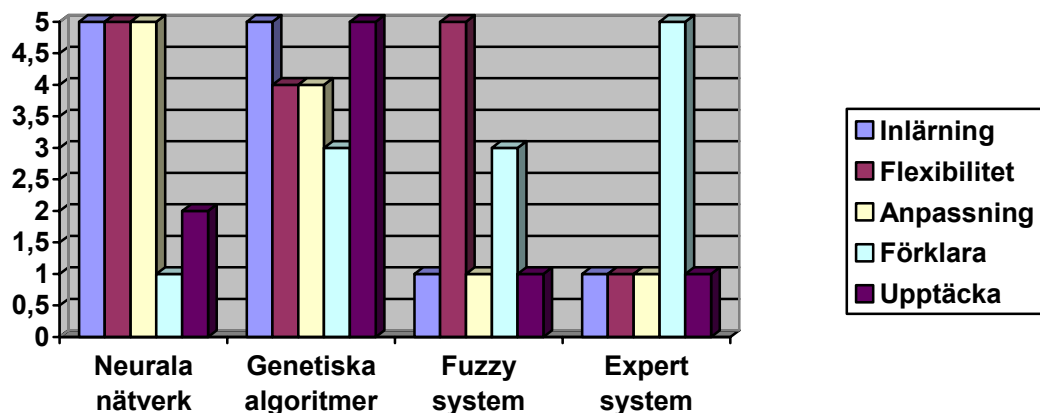
systemet. Eftersom en stor del av experternas kunskap bygger på erfarenhet och intuition är det svårt att få in den typen av kunskap i systemet. Systemet måste också vara lätt att använda så att inte användarens brist på kunskap blir systemets begränsning.

4.2.5 Intelligenta hybrid system

De intelligenta teknikerna som diskuterats ovan har alla sina speciella styrkor och svagheter. De har lyckosamt använts inom områden de lämpligen har kunnat appliceras till. Men vissa problem är för komplexa för att ett angreppssätt och därmed en teknik skall vara tillräcklig. Sådana här problem i kombination med teknikernas begränsningar har varit drivande för utvecklingen av ”Intelligenta Hybrid System”.

Vad är då Intelligenta hybrid system? Jo, det är en kombination av två eller flera av de teknikerna som har diskuterats ovan t.ex. vid ett komplicerat problem. Vid ex. planering och budgetering för ett företag har vi tre specifika problem, som framtidsprognos, optimering och vad vissa händelser får för påverkan då kan ett neuralt nätverk, en genetisk algoritm och ett experts system hantera respektive problem. De olika delsystemen kommunicerar mellan varandra för att få fram det slutliga resultatet.³⁰

Här nedan visar jag en tabell som visar på de intelligenta teknikernas betyg inom de fem nyckelområdena av AI (skala 1-5)³¹.



Hybrid systemens styrka är som tidigare sagts att de kan lösa problem som inte någon av de individuella teknikerna kan göra men vad har de för svagheter? En av de största svagheter är att det ställer mycket stora krav på utvecklarna som skall klara av att hantera

³⁰ Goonatilake S, s.21

³¹ Goonatilake S, s.21

flera olika tekniker som är komplicerade i sig, samt i vissa fall få dessa att kommunicera emellan varandra. Detta kan ställa till stora problem eftersom det antagligen bara finns ett fåtal personer som klarar av detta idag.

4.3 Användningsområden för intelligenta system

Idag har vi interaktion med intelligenta system dagligen även om vi inte är medvetna om det och jag skulle kunna skriva flera sidor bara om användningsområden för dessa system, men jag tänkte bara ta några väl spridda exempel för att skapa en uppfattning om bredden i dessa system.

4.3.1 Intelligenta System och Finansmarknaden

Jag tror att alla som investerar i aktier drömmer om att ha mer information än sina motspelare, något som ger dem övertaget att kunna kamma hem de stora pengarna. Det finns nog inte någon normalt sinnad person som köper sina aktier slumpvis utan alla försöker analysera företaget på något sätt och bedöma företagets möjligheter till vinst och tillväxt. Idag finns en relativt ny teknik för att förutsäga finansmarknaderna detta är neurala nätverk.

Dessa system skulle kunna ersätta dagens s.k. teknisk analys som man gör för att bedöma en aktie- eller en marknadstrend. Neurala nätverk är experter på att känna igen mönster och är därmed perfekta för finansiella tillämpningar eftersom de flesta marknadsexperter anser att marknaden grovt rör sig enligt vissa mönster.

Förutom att företagens investeringar ger bättre avkastning när timingen i handeln blir säkrare, så medför dessa datorsystem möjligheter till rationaliseringar på personalsidan.

4.3.2 Intelligenta System och Sjukvård

Datorer som läkare, låter det som ett otäckt framtidsscenario. Redan idag finns det beslutssystem som är utvecklade för att ge läkare råd om diagnos och behandling. Beroende på medicinskt problem används olika avancerade algoritmer för att bedöma testdata från en patient.

Dessa metoder måste anpassas till de medicinska krav som ställs i den specifika situationen. Hög tillförlitlighet är alltid av största vikt men i vissa situationer kan det mest betydelsefulla vara att undvika falskt positiva resultat eller vice versa. Det kan t.ex. vara fråga om analys av blodplasma prover från en patient med misstänkt cancer.

Sådana här system kan spara läkarna mycket tid vid diagnosställandet och en läkare kan ta sig en fler patienter. Dessa system blir mer och mer användbara allteftersom läkarbristen ökar.

4.3.3 Intelligenta och Industrin

Industrijobb har igenom alla tider varit farliga för människan men idag har vi kunnat ersätta människorna på några av de mest slitsamma arbetsuppgifterna med robotar. De industrirobotar som finns idag är mest av relativt ointelligent karaktär och kan bara arbeta under strukturerade förhållanden. Men det har börjat komma mer intelligenta varianter som även klarar av att arbeta i en mer oregelbunden miljö dessa robotar använder sig av olika sensorer för att orientera sig.

Om inte så länge kommer vi ha ökat användningen av intelligenta robotar inom industrin. Många forskare tror att dessa kommer att styras av mänsklig röst och att de kommer att kunna ”se”. Tillämpningsområden där man kan tänka sig att använda dessa robotar är kärnkraftverk och gruvor.

5. Resultat

Jag lyckades få tag i ett antal ledande personer inom ämnet AI för mina intervjuer. Var och en av personerna har lång erfarenhet av forskning inom området och vet vad som är möjligt och omöjligt inom området. Frågorna och intervjuerna finns i sina original som Appendix I och II.

Personerna som lät sig intervjuas var:

- **Bibel Wolfgang(WB)**, Professor Department of Computer Science of the Darmstadt Institute of Technology in Germany
- **Hayes Pat(PH)**, PhD in Artificial Intelligence, Professor John Pace Jr. Eminent Scholar at the University of West Florida
- **Michie Donald(DM)**, Professor Emeritus of Machine Intelligence at the University of Edinburgh
- **Nilsson Nils (NN)**, Professor of Engineering (Emeritus) in the Department of Computer Science at Stanford University
- **Schubert Len(LS)**, Professor Computer Science Department at the University of Rochester

5.2 Resultat av intervjuerna

- **Är det möjligt att få maskiner att bete sig som om de vore intelligenta?**

PH – Du måste förklara vad du menar med intelligent. Jag skulle säga att det redan idag finns maskiner som är intelligenta, eftersom om de utfördes av en människa skulle vi

betrakta det som en intelligent handling. Men jag antar att du har lite striktare kriterier vilket dagens maskiner inte kan nå upp till.

DM – Psykologen och lingvisten Howard Gardner identifierade sju former av intelligens och bekräftade att det antagligen finns ännu fler former av det vi kallar intelligens. Så till du specificerar vilken form av intelligens du menar är det inte möjligt att tolka frågan.

LS – Ja så småningom om 15- 30 år.

NN – Ja, jag tror det. Det grundar sig på min tro att människan är en typ av ”maskin”, så därför borde vi klara av att kopiera mänskliga funktioner till maskiner.

WB – Ja naturligtvis, schack datorer har gjort drag som till och med överraskade världsmästaren Kasparov (som förlorade mot en av dessa maskiner, Deep Blue), artificiella matematiker som lyckades bevisa teorem som inte mänskliga matematiker lyckades med osv. detta är intelligent beteende, eller?

- **Är det möjligt att konstruera maskiner som är medvetna och har en självständig intelligens?**

PH – Jag skulle säga ja. Jag ser i alla fall inget hinder för det. Men jag skulle vilja tillägga att vi idag inte vet hur vi skulle kunna lyckas, så detta är endast en kvalificerad gissning.

DM – Vi vet inte om människan är självmedveten, bara att vi påstår oss vara det och uppträder på ett sätt som talar för det.

LS – Ja. Detta är ett av mina aktuella forskningsområden.

NN – Jag tror att vi kommer att lyckas göra maskiner som påstår att de är självmedvetna precis som vi människor påstår att vi är självmedvetna.

WB – Det beror på vad du menar med “självmedvetenhet”. Vi har redan idag system som stämmer av sitt eget uppförande och ur den synvinkeln är självmedvetna. Om du menar självmedvetenhet som den upplevs av människan så tror jag att det kommer vara möjligt att göra men inte i den närmsta framtiden.

- **Bör man sträva efter det? (syftar på frågan ovan)**

PH – Jag tycker att det är värt att försöka. Det skulle t.ex. kunna ge oss mycket insikt om vår egen självmedvetenhet som idag är ganska utforskad och jag skulle vilja förstå den bättre.

DM – Jag har diskuterat det här problemet i mina böcker.

LS – Ja det skulle göra dator - människa interaktionen mycket mer naturlig och systemet skulle vara medvetet om vad det kan, inte kan och vad som nyligen skett i interaktionen med användaren.

NN – Jag tycker att vi skall det för det kan hjälpa oss att förstå vad ”självmedvetande” är för något.

WB – Människor vill förstå vad som pågår i världen.....allt som händer i den, inklusive medvetenhet. Eftersom det inte kommer att ske inom den närmsta framtiden vill jag inte slösa bort min tid genom att väga fördelar mot nackdelar.

● **Kan en intelligens skapa en annan intelligens ännu intelligentare än sig själv?**

PH – Jag förväntar mig det. Jag kan inte se någon anledning till motsatsen. Men ju mer vi lär oss om intelligens desto mindre greppbart och enkelt verkar området så meningar som mer intelligent klingar inte väl i mina öron. Jag önskar ställa en motfråga, - mer intelligent, ”på vilket sätt” eller ”ur vilket perspektiv”.

DM – Det är något mänskliga föräldrar ofta lyckas med

LS – Ja, naturen lyckades. Schaeffer’s schackprogram spelar mycket bättre schack än han själv dessutom spelar Deep Blue bättre schack än sina skapare. Teori bevisande system har i vissa fall lyckats bevisa teorem som inte mänskliga matematiker lyckats med.

NN – Ja, det tror jag vi kommer lyckas med.

WB – Ja, de människorna som utvecklade Deep Blue är alla sämre schackspelare än maskinen de skapade.

● **Vad anser du att bristerna i dagens Intelligent System ligger?**

PH – I stort sett allting, med det menar jag att dagens AI system är bara små delar av ett totalt intelligent system (jämförbart med mänsklig eller djurs intelligens). Inom vissa små begränsade (användbara) områden kanske de kan vara bättre en mänsklig intelligens. Här är en analogi jag tycker är bra: fråga dig om flygplan verkligen kan flyga som en fågel. Ur vissa perspektiv flyger de bättre, de är snabbare, de flyger högre och klarar lyfta mer vikt men i vissa fall är det totalt underlägset, det kan inte landa i trän eller fånga fisk ur havet. Om mänsklig intelligens är som en flygande fågel så är AI system mer som ett flygplan byggt för att klara av ett begränsat område av arbetsuppgifter, men gör de onaturligt bra!

DM – Förmågan att lära sig och förmågan att visa sin förståelse genom diskussion.

LS – (1) Sunt förnuft (vilket behövs när man för vissa typer av resonemang som vi än inte lyckats automatisera och på stora mängder kunskap som vi inte lyckats införa i system). (2) Kreativitet i skapande, lärande och problemlösning. (3) Självmedvetenhet. (4) Sammanlänkade motiv och ha drivkraft i det planerade beteendet.

NN – Många, många saker, bland annat: Sunt förnuft, perceptuell begåvning (ha visioner), inlärningsbegåvning, språkförståelse, generationsförståelse.....

WB – Självklart, våra maskiner har fortfarande en outvecklad intelligens (oftast bara ett litet expertisområde) till och med i jämförelse med de mest korkade människorna.

- **Kan man verkligen låta Intelligenta System utföra analys av komplicerade problem utan att granska lösningen kritiskt!(ett system kan trots allt göra misstag)**

PH – Självklart, det gör vi ofta redan! Som du säger kan alla system göra fel, detta yttrande stämmer dock bättre in på människor än maskiner.

DM – Programmerarna av Deep Blue lät ett ointelligent system lösa komplexa problem, genom en massiv sökmetod och en anpassad hårdvara. De granskade inte lösningarna kritiskt eftersom ingen av dem var speciellt bra på schack. Om Deep Blue hade varit intelligent hade de kunnat granska lösningarna kritiskt genom att be den att förklara sina drag.

LS – Inte förrän vi lyckats implementera sunt förnuft och sammanlänkning av motiv och förmåga att planera beteende samt att de sammanlänkade motiven är välvilligt inställda till oss.

NN – Jag är inte säker på att jag förstår vad du menar men jag tror att vi kommer att lyckas utveckla intelligenta maskiner vars beteende vi kanske inte kommer att förstå oss på.

WB – Kontroll av en lösnings korrekthet är en ganska enkelt. Finna lösningen är det som är den svåra delen. Se bara på matematisk bevisning som ett exempel.

- **Samma fråga som ovan men anta att det kan innebära allvarliga konsekvenser om systemet gör fel.**

PH – Vet inte om jag förstår vad du menar här. Det är uppenbart att om maskiner besitter och använder kunskap så måste vi försöka se till att den är korrekt. Det här är en verklig fråga just nu och inte en framtida fantasi, se t.ex. på den felaktiga bombningen av den kinesiska ambassaden för några år sedan.....³²

³² Har inte lyckats ta reda på vad som hände man antar att det var ett systemfel eftersom Philip Hayes uttrycker sig så.

DM – Ett meteorologiskt datasystem förutser inte vädret med perfekt pålitlighet! Ur den synvinkeln gör systemet fel vilket skulle kunna få allvarliga konsekvenser t.ex. för de som arbetar på havet och måste ha orkan-/stormvarningar. Om det vore intelligent så skulle kunna förklara sina prognoser och erbjuda ursäkter för sina misstag precis som en mänsklig meteorolog gör. Skulle en sjöfarare kunna låta en intelligent ”människa” lösa det komplexa problemet, vilket ge prognos vädret är och ta sina beslut grundade på ”den” informationen. Om samme sjöfarare inte hade råd att betala för de tjänster som dagens meteorologiska system erbjuder, vilket ger betydligt bättre prognoser än en mänsklig meteorolog utan hjälpmedel gör skulle han antagligen föredra det ointelligenta datasystemet framför den intelligenta människan. Hans bästa alternativ om pengar inte är något problem är att låta en mänsklig meteorolog som använder sig av det allra senaste och exakta meteorologiska datasystemet som finns tillgängligt oavsett om det är intelligent eller inte.

LS – Jag trodde det var den frågan jag nyss svarade på men din fokus här kanske ligger mer på en förmåga att sammanlänka motiv och planera sitt beteende.

NN – Vi har redan sådana system (luftfarts kontroll, flygreservationssystem, personal- och kunddatabaser, ...), de gör fel och vi är tvungna att lära hur vi skall skydda oss mot dessa fel.

WB – Jag förstår inte denna fråga. Maskiner likt människor gör fel. Mer intelligenta maskiner gör färre fel.

- **Anser Ni att det finns någon fara i att förlita sig på intelligenta system för mycket?**

PH – Inte någon överliggande fara, jag tror att ju mer intelligens vi kan trycka in i våra maskiner, ju mer användbara kommer de att bli. Det kan finnas en fara i att tro att de är mer kompetenta än de verkligen är men vi har det problemet med varandra hela tiden.

DM – Inte lika farligt som att fortsätta förlita sig på ointelligenta system.

LS – Ja så länge de kommer sakna de saker jag pratade om i frågan om ”bristerna i dagens Intelligenta System”, eller saknar välvilliga motiv.

NN – Ja, precis som jag ser det som ett problem att vara för beroende av människan.

WB – Självklart, buggar kommer alltid att vara en viktig fråga. Men jämför med dumheten hos mänskliga politiker och misstagen de begår (Usama bin Ladin som ett extremt exempel).

- **Ni känner till superdatoren HAL ifrån boken och filmen 2001- Ett Rymdäventyr. En dator med självmedvetande ligger långt in i framtiden men en retorisk fråga: - vem ansvarar för när HAL dödar? Användarna, utvecklarna eller kanske HAL själv?**

PH – Tja, det beror på om vi tänker på HAL som en riktig ”person” med rättigheter och skyldigheter. Jag tror inte vi kommer att se våra maskiner på det här sättet inom en överkomlig framtid. Så för att svara på din fråga skulle jag säga att det beror på omständigheterna som de flesta juridiska frågor. Om ett bildäck exploderar och någon dör, vem skall man då anklaga? Kanske ingen i slutändan.

DM – I laglig mening måste sådana här frågor liknas vid en hund som dödar ett får eller en person. Det vanligaste utslaget är att hundens ägare är lagligt och moraliskt ansvarig. I HALs fall är den första frågan som bör fastställas för de rättsliga prejudikaten, ägarskapsfrågan.

LS – Alla de ovanstående, om HAL verkligen besatt mänsklig intelligensnivå. Utvecklarna var för arroganta och oförsiktiga, användarna litade för mycket på systemet och HAL var för övermodig i sina beslut.

NN – Jag tror inte att någon s.k. expert har bättre förmåga än vem som helst att besvara den här typen av fråga. Vi måste klura ut den här typen av frågor ur ett samhällsperspektiv.

WB – Det ligger för långt borta för att kunna spekulera ur ett sådant här scenario. Om tiden kommer så får vi hitta passande lösningar då. I vilket fall som helst kan vi specificera HAL så att den inte kommer att döda och att vi alltid kommer att kunna bryta strömmen om någonting går fel.

6. Analys, Diskussion och slutsats

6.1 Analys och diskussion

Jag liksom de flesta andra människor mötte termen AI för första gången genom film. Där möter man ofta intelligenta system som klarar av allt möjligt och knappast har några begränsningar, robotar som kan gå, kommunicera, känna känslor mm.

I stort sett alla intervjuade ansåg att det går att skapa system med ett intelligent beteende även om vissa vill vrida och vända på termen intelligens. Speciellt Nils Nilsson synsätt är intressant i det här fallet eftersom han liknar människan vid en maskin och ser därför inga problem med att kopiera mänskliga funktioner till maskiner. Vissa av de intervjuade anser t.o.m. att vi har intelligenta system redan idag Wolfgang Bibel använde sig av

Schack maskinen Deep Blue som exempel och ansåg den vara intelligent för att den gjorde drag som överraskade och vann över världsmästaren i schack vid tidpunkten. Men jag anser att det är att ta i för mycket för då kan man säga att en miniräknare är intelligent för att den kan lösa en ekvation snabbare än en kvalificerad matematiker.

En viktig del av vår egen intelligens är att vi är självmedvetna dvs. vi vet att vi är och därför är vi intresserade av att lära saker och det är kanske först då vi kan betrakta ett system som intelligent. När vi lyckas skapa ett system som ger oss meddelanden som inte på något sätt är förprogrammerade eller kan förutses ex. ”Jag känner mig ensam och uttråkad, håll mig sällskap”³³.

Och de flesta av de intervjuade är optimistiska till möjligheten att kunna skapa en artificiell ”medveten” intelligens, även om de är överens om att det i så fall ligger en bra bit fram i tiden.

Philip Hayes sa att vi idag inte vet hur vi skall kunna lyckas skapa, och var skall man börja om man skall skapa en s.k. medveten intelligens? Så bör man då sträva efter att försöka skapa det?

Ja, självklart tyckte de intervjuade. De menade på att det är en möjlighet för oss att utforska mer om vårt egna s.k. självmedvetande. Len Schubert tyckte att det var bra för att skapa en enklare människa – maskin interaktion.

Själv är jag lite mera skeptisk eftersom om du skapar en självmedveten intelligens så skapar du också ett väsen med en vilja och det betyder att den måste vilja samma sak som användaren för att interaktionen skall förenklas och förbättras.

Men kan en intelligens skapa en intelligens som är intelligentare än själv? Ja, är det enade svaret av alla intervjuade och vissa använder liknelser som ”det är något mänskliga föräldrar ofta lyckas med” eller ”Deep Blue spelar bättre schack än sina skapare”. Men jag tror knappast att Deep Blue’s schack kunskaper kommer från systemets skapare utan att de använt sig av experter och det genetiska lotteriet som ger föräldrar barn som är intelligentare än sig själva kan lika gärna slå åt andra hållet.

Som problemet ser ut idag kan är det ju inga problem att baka in mer kunskap om ett område i ett intelligent system men det betyder ju inte att systemet är mer intelligent än en människa med en mindre kunskapsbas utan vad dagens IS saknar är ju ett resonemang som fungerar liknande människans (ex. en människa kan bl.a. se samband och utnyttja sin kunskap i ett område som ligger utanför problemområdet för att lösa ett problem).

Idag när intelligenta system alltmer berör våran vardag, exempelvis finns idag bilar med intelligenta system som rätar upp bilen om vi får sladd och intelligenta system som bedömer din kreditkortsansökan osv. är det ju intressant att utreda vad bristerna ligger i dagens system.

³³ Kurzweil R, s.51-52

Alla verkar vara överens om att det finns brister i dagens system i alla fall. Något som verkar vara genomgående tema är ”sunt förnuft” och det måste man ju hålla med om eftersom det är detta som gör den mänskliga intelligensen så speciell. Dessutom tycker flera av de intervjuade att det är viktigt att utveckla områden som kreativitet, förmågan att lära sig, språkförståelse och självmedvetenhet.

Här pratar också ett par av de intervjuade om att dagens intelligenta system är alldeles för specialiserade. Framför allt Philip Hayes är inne på det spåret att trots att vi idag har IS som kan sköta sitt lilla expertisområde bättre än en människa så är det här bara fragment av ett totalt IS.

Wolfgang Bibel säger också något jag tycker stämmer in på min tankegång och det är att som IS ser ut i dag så är de dummare än de mest korkade människorna. Den här synpunkten har han för att han tycker precis som Philip Hayes att de är för begränsade med sina små expertisområden och har ingen större generell intelligens.

När man vet att dagens system saknar en hel del bitar, kan man då låta dem utföra analyser eller hantera vissa situationer för oss utan vår insyn och lita på att de gör rätt?

Ja, verkar det gemensamma svaret vara. De intervjuade menar att maskiner likt människor kan göra fel och det är större risk att människor gör fel än maskiner. Det är endast Len Schubert som vill att systemen skall ha ett sunt förnuft innan man låter dem göra något oönskat.

Och om det kan innebära allvarliga konsekvenser om systemet gör fel? Här går åsikterna mer isär än på frågan innan. Philip Hayes menar att ”det är självklart att om maskiner besitter och använder kunskap så bör man se till att den är korrekt” medan han på frågan innan tyckte det var självklart att man kunde låta systemen arbeta själv med motiveringen maskiner likt människor gör fel.

Nils Nilsson är väl den av de intervjuade som har bäst syn på detta problem tycker jag. Han säger som så att dessa system är en del av vår vardag och ”vi måste lära hur vi skall skydda oss mot dessa fel”.

Något som verkar vara genomgående för gruppen intervjuade är att de har synen: att förlita sig på ett intelligent system är precis lika riskfyllt som att förlita sig på en människa. Det kan de mycket väl ha rätt i men normalt så brukar man ifrågasätta om man tycker att en människa säger något felaktigt medan man ofta tar för givet att ett datasystem utför sin uppgift korrekt.

Jag använde mig av superdatoren HAL från filmen och boken ”2001 - ett rymdäventyr” för att exemplifiera ett system i den sista frågan. Jag tyckte det var konstigt att ett par av de intervjuade hade en strikt juridisk syn på ansvarsfrågan om vem man skall anklaga när systemet gör fel mer än att hävda sitt eget ansvar i sammanhanget. Jag vet inte om de tog frågan för bokstavligt kanske och ansåg HAL vara ett riktigt väsen.

Len Schubert för däremot ett resonemang som jag anser vara riktigt dvs. ”utvecklarna var för arroganta”, ”användarna litade för mycket på systemet” samtidigt som HAL fattade dumdristiga beslut. Det är här ”Pudelns kärna³⁴” ligger anser jag!

6.2 Slutsats

Som Marvin Minsky uttrycker det! ”Redan nu kan vi en hel del rörande att skapa användbara och specialiserade system. Men vi vet fortfarande inte hur vi skall få dem att utveckla sig själva på ett sätt som är intressant. Men när vi försöker lösa de frågorna är vi tvungna att ställa oss en klurigare och konstigare fråga.

–När vi lyckas skapa ett system med självständig intelligens, skall vi då skapa ett system som är ”smartare” än oss själva?

Vi har ju ändå tur som fortfarande kan lämna vidare det här valet till senare generationer. De kommer säkert inte att bygga ett sådant system om de inte har goda skäl till det³⁵”.

Jag tolkar detta som om att han menar att det kanske inte alltid är så bra att skapa någonting man inte förstår sig på!

Min syn på Artificiell Intelligens och Intelligenta system är och kommer förmodligen alltid att vara att datorers intelligens aldrig kommer att vara jämförbar med människans intelligens.

Jag trodde att jag skulle få lite mer stöd för min skepsis mot Intelligenta Systems suveränitet över oss människor. Detta kan låta som om jag är emot utvecklingen men jag tycker inte vi skall vara rädda för att utnyttja och utveckla tekniken men aldrig låta systemen arbeta oövervakade om de sysslar med något som om det gör fel kan innebära allvarliga konsekvenser för en individ eller en grupp.

Ta det exempel som Philip Hayes tar upp i intervjun ovan, roboten som för några år sedan felaktigt slog ner i den Kinesiska ambassaden i Jugoslavien. Konsekvenserna var i alla fall flera döda och många skadade samt uppstod en politisk kris som om det hade velat gå illa kunnat leda till krig.

Nu vet inte jag om det berodde på att informationssystemet gav felaktig output om det planerade målets placering, robotens målsökningssystem hade fått felaktiga koordinater inmatade eller om det uppstod något annat systemfel. Dock tror jag det att om det utförts en noggrann mänsklig kontroll av systemens input, output och rimligheten i dessa hade bedömts hade den här situationen kunnat undvikas.

Det var bara en eller två av de intervjuade som gav mig stöd för mitt sett att se på saken dvs. att se på de intelligenta systemen som en resurs men inte som ett människan överlägset väsen.

³⁴ Uttrycket "pudelns kärna" myntade Goethe i dramat "Faust" ung. den verkliga innebörden av något.

³⁵ <http://www.media.mit.edu/people/minsky/papers/ComputersCantThink.txt>

Vissa ansåg att vi idag redan hade intelligenta system medan några andra tyckte att man hade långt kvar till dess. Själv anser jag likt några av de intervjuade att intelligens inte bara är en förmåga att utföra en enda form av intelligent handling utan också en förmåga att utvärdera konsekvenserna av sitt handlande. Men att inte kontrollera systemen och helt förlita sig på dem även om de har denna förmåga anser jag vara riskfyllt, för även om intelligensen blir mer avancerad är den ändå likgiltig!

Naturligtvis är det inte så att man behöver övervaka ett system som Deep Blue som om det gör ett misstag som värsta konsekvens får att den förlorar sin schackmatch utan det jag menar är att system som ex. genomför en diagnos av en patient och därmed tredje part kan drabbas av misstagen bör övervakas av en mänsklig expert som bedömer rimligheten i diagnosen, i fall att!

Jag kan inte påstå att min uppsats ger upphov till några nya vetenskapliga slutsatser, däremot kanske den ger oss en tankeställare om hur vi vill att våra intelligenta system skall se ut i framtiden och hur vi skall hantera användandet av dem.

6.3 Förslag till vidare studier

Jag skulle tycka att det skulle vara intressant om någon skulle ta reda på och skriva om ett ”real case scenario” där vi som människor förlitat oss för mycket på ett intelligent system, exempelvis den tragiska bombningen av Kinesiska ambassaden i Jugoslavien.

Ta reda på var någonstans gick det fel. Var det konstiga inputvärden, systembugg eller ett tolkningsfel av användarna?

8. Referenslista

Turban E, Aronsson J E, *Decision Support Systems and Intelligent Systems*, 6th edition, Prentice Hall International Inc., 2001, New Jersey

Kurzweil R, *The age of the spiritual machine*, Penguin Group, 1999, New York

Goonatilake S, Treleaven P, *Intelligent Systems for finance and business*, John Wiley & Sons Ltd., 1995, Chichester

Wiedersheim-Paul F, Eriksson LT, *Att utreda, forska och rapportera*, Liber-Hermods, 1991, Malmö

Rothfeder J, *Den tänkande datorn*, Bonnier Fakta Bokförlag, 1986, Stockholm

Michie D, Johnston R, *The knowledge machine*, William Morrow and Company, 1985, New York

Hartman Jan, *Vetenskapligt tänkande -från kunskapsteori till metodteori*, Studentlitteratur, 1998, Lund

Knight K, Rich E, *Artificial intelligence* 2nd edition, McGraw-Hill Inc., 1991, New York

Rowe A J, Davis SA, *Intelligent Information Systems*, Quorum Books, 1996, Westport-Connecticut

Andra medier

Föreläsningsanteckningar till kurs i Beslutstödsystem, Agahi F

<http://www.sciam.com/specialissues/1198intelligence/1198yam.html>

<http://library.thinkquest.org/2705/history.html>

<http://www.aaai.org/>

<http://http.cs.berkeley.edu/~russell/ai.html>

<http://psych.utoronto.ca/%7Eeringold/courses/ai/>

http://www.csd.uu.se/kurs/ait/ait96/lectures/vad_ar_AI.shtml

<http://www.media.mit.edu/people/minsky/papers/ComputersCantThink.txt>

Övrig litteratur

Granholm A, Hörnfeldt E, *Digital Guldrusch*,
Bonnier Icon, 1999, Falun

Stork DG, *Hal's Legacy – 2001's Computer as dream and reality*,
MIT Press, 1997, Cambridge-Massachusetts

Appendix I

Questions

- Is it possible to make machines act as if they are truly intelligent?
- Is it possible to make intelligent systems with self-awareness?
- Should we strive for that? (refers to the question above)
- Can an intelligence create another intelligence more intelligent than itself?
- What do you regard today's intelligent systems is lacking?
- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)
- Same question as above with the knowledge it could mean serious consequences if the system made an error.
- Do you see any danger to rely on intelligent system too much!
- I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away but I want to ask you a rhetorical question: -who's to blame when HAL kills? The users, the developers or maybe HAL itself?

Appendix II

Pat Hayes

- Is it possible to make machines act as if they are truly intelligent?

You really need to say what you mean by 'truly intelligent'. I would say that machines currently exhibit intelligence, in the sense that they do things that if people did them we would rate them as evidence of intelligence in the human. But I suspect that you have some other, stricter, criterion in mind, and that current performance would not measure up to what you mean by 'truly'.

- Is it possible to make intelligent systems with self-awareness?

I would say, yes. That is, I see no barrier in principle to doing this. However I would add that we do not currently know how to do it, so this can be nothing better than an educated guess.

- Should we strive for that? (refers to the question above)

I think it is worth trying to do, yes. It would probably give us a lot of insight into our own sense of self-awareness, which is currently rather mysterious and I would like to understand better, for example.

- Can an intelligence create another intelligence more intelligent than itself?

I expect so. That is, I see no principled reason why not. However, the more we find out about 'intelligence', the less of a unitary, simple property it seems to be, so phrases like 'more intelligent' seem to ring false. I am tempted to reply, more intelligent *in what way* or *in what respect*?

- What do you regard today's intelligent systems as lacking?

Oh, almost everything :-). That is, current AI systems are only small pieces of a fully intelligent system (comparable to a human or even most animal intelligences, for example). In some ways they may surpass human performance in some limited area, but only in some very small (but often useful) area.

Here's an analogy I often find helpful: ask yourself if aircraft can really fly in the same way that birds do? In some ways they fly better: faster, higher, lifting more weight - but in many ways they are hopelessly inferior: they cannot land in trees, or catch fish from the ocean, or dart around at dusk catching insects. If human intelligence is like bird flight, AI systems are more like aircraft: built to do a limited range of things, but to do them unnaturally well.

- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)

Sure, we do this all the time already. As you say, any system can make an error: that applies to humans more forcibly than to machines, however.

- Same question as above with the knowledge it could mean serious consequences if the system made an error.

Not sure what you mean here. Obviously if machines have and use knowledge then we must take care to try to ensure that it is accurate. This is a real-life issue right now, not a future fantasy, c.f. the accidental bombing of the Chinese embassy a few years ago..

- Do you see any danger to rely on intelligent system too much!

Not really. I think that the more intelligence we can put into our machines, the more useful they will be, on the whole. There may be some dangers in thinking they are more competent than they really are, but we have that problem with each other all the time.

- I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away but I want to ask you a rhetorical question: -who's to blame when HAL kills? The users, the developers or maybe HAL itself?

Well, that depends on whether we think of HAL as a real 'person' with rights and responsibilities. I don't think that we will think of our machines in this way in the foreseeable future, so I would say that the answer to your question will depend on circumstances, like most legal questions. If a car tyre bursts and this causes a death, who is to blame? Maybe no-one, in the end.

Donald Michie

- Is it possible to make machines act as if they are truly intelligent?

The psychologist and linguist Howard Gardner wrote a book called "The Theory of Multiple Intelligences", in which he identified and discussed seven distinct forms of intelligence, while acknowledging that many more forms probably shelter under this same umbrella term "intelligent". Until you specify which form of intelligence you have in mind, it is not possible to interpret your question.

- Is it possible to make intelligent systems with self-awareness?

We do not know that *humans* are self-aware, but only that they claim to be and that they behave in ways consistent with this claim.

- With regard to machines:

I have discussed these matters in "Turing's Test and conscious thought", *Artif. Intell.* 60, 1-22, 1993 and in "Consciousness as an engineering issue" which appeared in two Parts in *J. of Consciousness Studies*, 1, 182-195, 1994 and 2, 52-66, 1995.

- Should we strive for that? (refers to the question above)

Of course we should, supposing that we want machine intelligences with which we can easily exchange the facts and fruits of experience.

- Can an intelligence create another intelligence more intelligent than itself?

This is quite often achieved by human parents.

- What do you regard today's intelligent systems as lacking?

Teachability and the capacity to demonstrate understanding through discourse.

- Can you really let intelligent systems solve complex problems without having a critical view on the solution! (knowing that any system can make an error)

The programmers of Deep Blue let an unintelligent system solve complex problems by brute force search and special purpose hardware. They did not have a critical view on the solutions, since none of them were master-strength chess players. If Deep Blue had been an intelligent system, they could have obtained a critical view by asking it to explain its moves.

- Same question as above with the knowledge it could mean serious consequences if the system made an error.

A meteorological computing system does not forecast the weather with perfect accuracy. In that sense it makes errors, some of which mean serious consequences, for example to sea-farers who rely on it for hurricane warnings. If it were intelligent it would be able to explain its forecasts and offer plausible excuses for its errors, just as a human weather forecaster does. Could a sea-farer let an intelligent *human* solve complex problems of weather forecasting and take decisions based on the offered solutions? If he could not afford the fee charged for access to today's computer-based systems, which forecast much more accurately than unaided human meteorologists can, I expect that he probably would prefer the unintelligent computer to the intelligent human as a source of advice.

Perhaps his best bet, if cost were not a factor, would be to put his problems to a human meteorologist who uses the help of the most accurate weather-forecasting computing system available, regardless of whether the system was intelligent or not.

- Do you see any danger to rely on intelligent system too much!

Not as great a danger as continued reliance on unintelligent systems.

- I presume you remember about HAL the supercomputer from the movie 2001- A Space Odyssey . A computer with self-awareness is a long time away but I want to ask you a rhetorical question: - who s to blame when HAL kills? The users, the developers or maybe HAL itself?

In the legal context, similar questions have to be adjudicated in the context of a dog which kills sheep, or kills a person. The usual ruling is that the dog's owner is legally, and presumptively morally, liable. In the case of HAL, the first question to be determined, according to the legal precedents, is the question of ownership.

Len Schubert

- Is it possible to make machines act as if they are truly intelligent?
Yes, eventually (15 - 30 years ;-)

- Is it possible to make intelligent systems with self-awareness?
Yes. It's one of my current research interests.

- Should we strive for that? (refers to the question above)
Yes. It will make interaction more natural, and the system will show awareness of what it knows (or can do) and what it doesn't know, and what has recently transpired in its interactions with users.

- Can an intelligence create another intelligence more intelligent than itself?
Yes. "Nature" did. Schaeffer's checkers program plays much better than he does, and Deep Blue plays better chess than its designers. Theorem provers have proved some theorems that stymied human mathematicians.

- What do you regard today's intelligent systems is lacking?
(1) Common sense (which depends on certain kinds of reasoning we haven't successfully automated, and on vast amounts of knowledge we haven't successfully imparted to machines); (2) Creative concept formation and discovery in learning and problem solving (except in rather weak senses); (3) self-awareness; (4) overarching motives, driving planned behaviour.

- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)
Not until we have endowed them at least with (1) and (4) above, and the overarching motives are benign towards us.

- Same question as above with the knowledge it could mean serious consequences if the system made an error.

Well, I thought that was the question I was answering above; but maybe here your focus is on (4), rather than (1).

- Do you see any danger to rely on intelligent system too much!

Sure, as long as they lack the above features (1-4), or lack benign motives.

- I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away but I want to ask you a rhetorical question: -who's to blame when HAL kills? The users, the developers or maybe HAL itself?

All of the above, if HAL were really of human-level intelligence. The developers were too arrogant and incautious, the users too trusting, and HAL too presumptuous in the decisions it made.

Hector Levesque

- Is it possible to make machines act as if they are truly intelligent?

Nobody knows.

- Is it possible to make intelligent systems with self-awareness?

Ditto.

- Should we strive for that? (refers to the question above)

Yes, I think so.

- Can an intelligence create another intelligence more intelligent than itself?

Nobody knows.

- What do you regard today's intelligent systems is lacking?

I know of no intelligent systems.

- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)

I'm not sure I understand the question. If you mean should we be skeptical of a system solving a complex problem, then yes.

- same question as above with the knowledge it could mean serious consequences if the system made an error.

Yes, this is true for any advanced technology.

- Do you see any danger to rely on intelligent system too much!

Ditto.

- I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away

but I want to ask you a rhetorical question: -who's to blame when HAL kills?
The users, the developers or maybe HAL itself?

I think the story was that the developers gave HAL conflicting instructions:
protect the crew members, safeguard the mission.

Nils Nilsson

- Is it possible to make machines act as if they are truly intelligent?

Yes, I believe so. My belief is founded on my other belief that humans are a kind of machine, so therefore we ought to be able to duplicate human functions in machines.

- Is it possible to make intelligent systems with self-awareness?

I think we will make machines that will claim they are self-aware, just as we claim to be self-aware.

- Should we strive for that? (refers to the question above)

I think we should because it will help us understand what "self-awareness" is.

- Can an intelligence create another intelligence more intelligent than itself?

Yes, I think we will.

- What do you regard today's intelligent systems is lacking?

Many, many things. Among these are: commonsense knowledge, perceptual abilities (especially vision), learning abilities, language understanding and generation,

- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)

I'm not sure what you mean, but I believe we will be able to "evolve" intelligent machines whose workings we might not understand.

- Same question as above with the knowledge it could mean serious consequences if the system made an error.

We already have such machines (air traffic control, airline reservation systems, employee and customer databases, ...), and they do make errors, and we are learning how to protect ourselves against these errors.

- Do you see any danger to rely on intelligent system too much!

Yes---just as I see danger in relying on humans too much.

I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away but I want to ask you a rhetorical question: -who's to blame when HAL kills? The users, the developers or maybe HAL itself?

I don't think so-called "experts" have any better understanding about how to answer a question like that than anyone else. As a society, we'll have to figure out how we want to answer these kinds of questions.

Wolfgang Bibel

- Is it possible to make machines act as if they are truly intelligent?

Yes of course. Chess machines generate moves which have astonished even the worldmaster Kasparov (who lost against one of these machines, deep blue). Or artificial mathematicians (called theorem provers) proved theorems which human mathematicians tried hard but unsuccessfully to prove. And so forth. This is intelligent behavior isn't it?

- Is it possible to make intelligent systems with self-awareness?

This depends a lot on the definition of what you mean by "self-awareness". In one sense we have of course systems which (on a meta-level) watch their own behavior and in this respect are "aware" of themselves. If you mean self-consciousness as experienced by humans then I also believe that this will be possible but not in the very near future.

- Should we strive for that? (refers to the question above)

Humans want to understand the world -- all of it, including consciousness. Since it will not come around in the near future, I will not waste my time with weighing advantages and disadvantages against each other.

- Can an intelligence create another intelligence more intelligent than itself?

Yes. Those people who developed deep blue, all are less good chess players than the machine created by them.

- What do you regard today's intelligent systems as lacking?

Of course. Our machines are still only rudimentary intelligent (eg. mostly only in a tiny window of expertise) in comparison even with the most stupid people.

- Can you really let intelligent systems solve complex problems without having a critical view on the solution!(knowing that any system can make an error)

Checking the correctness of a solution is mostly rather trivial or straightforward. Finding the solution is the hard part. Just think of a mathematical proof as an example.

- Same question as above with the knowledge it could mean serious consequences if the system made an error.

I don't understand this question. Machines like people make errors. More intelligent machines make less errors.

- Do you see any danger to rely on intelligent system too much!

Of course, bugs will always remain an important issue. But compare the stupidity of human politicians and the grand mistakes they make (eg. Osama Bin Ladin as an extreme example)

- I presume you remember about HAL the supercomputer from the movie 2001-"A Space Odyssey". A computer with self-awareness is a long time away but I want to ask you a rhetorical question: -who's to blame when HAL kills? The users, the developers or maybe HAL itself?

It's too far away to speculate on such a scenario. If time comes men will find appropriate legislation to cope for such a problem. In any case we can specify HAL so that it will not kill and can always cut the power supply if things go wrong.