KISMET
En social robot
Abstract
Idén bakom Kismet är att göra en robot som med hjälp av ett antal underliggande system kan kommunicera och interagera med en människa. Detta arbete redogör för hur Kismet är uppbyggd, och vad som gör att man kallar den en social robot.
För att få en människa att tro att roboten har känslor och intentioner i det den gör måste den efterlikna människan så mycket som möjligt för att människan den interagerar med reaktivt ska kunna tolka informationen. Detta uppnås genom att bygga in ett antal system i roboten, som var för sig har hand om en del av den mänskliga perceptionen av världen och hur människan uttrycker och visar sina känslor och intentioner.
1 Inledning ..................................................................................................................................1
2 Social interaktion med människan ..........................................................................................2
3 Kismets uppbyggnad .............................................................................................................2
   3.1 Yttre system ......................................................................................................................2
      3.1.1 Visuella systemet .......................................................................................................2
      3.1.2 Auditiva systemet ......................................................................................................2
      3.1.3 Vokala systemet .........................................................................................................3
      3.1.4 Uttryckssystemet .......................................................................................................3
   3.2 Inre systemet .....................................................................................................................3
      3.2.1 Särdragssystemet .......................................................................................................3
      3.2.2 Det perceptuella systemet ..........................................................................................4
      3.2.3 Uppmärksamhetssystem ............................................................................................4
      3.2.4 Motivationssystemet ..................................................................................................4
      3.2.5 Beteendesystemet ......................................................................................................4
      3.2.6 Motoriska systemet ....................................................................................................6
         3.2.6.1 Okulär kontroll ...................................................................................................6
         3.2.6.2 Ansiktsuttryck ...................................................................................................6
         3.2.6.3 Auditivt uttryck ..................................................................................................7
4 Diskussion ...............................................................................................................................9
1 Inledning

2 Social interaktion med människan

Breazeal påstår i sina artiklar att för man ska kunna kalla en robot social, måste roboten ge
sken av att ha en intention med alla sina handlingar, att kunna efterlikna de varelser vi är vana
att möta i vår sociala interaktion, och att roboten kan förmedla sitt interna tillstånd, med
känslor och ansiktsuttryck. En social robot måste också kunna hantera oförutsedda händelser
och kunna hitta ett bra sätt att reagera på dessa. För att få en människa att tro att roboten har
känslor och intentioner i det den gör måste den efterlikna människan så mycket som möjligt
för att människan den interagerar med reaktivt ska kunna tolka informationen, så som man
tolkar information från en annan medmänniska. Roboten efterliknar då människans
egenskaper vid konversationer, genom att kunna rikta blicken, betona talet beroende på
känsla, och visa med ansiktsuttryck vad den menar.

Kismet är utvecklad för att likna ett litet barn, och har många likheter med dessa. Man vill
kunna lära roboten vilka känslouttryck som krävs för att få människan man interagerar med
att göra en viss sak som tillfredsställer robotens behov. Småbarn fungerar på precis samma
sätt. Eftersom bebisen inte kan uttrycka vad dess behov är, testar den olika uttryck för att se
vilken responsen blir från föräldern. Genom att göra detta får barnet föräldern att tro att barnet
har en bakomliggande mening med sitt beteende, och därmed utför det barnet önskar. Genom
att prova många olika beteenden lär sig barnet så småningom vilket som är det bästa, och i
likhet med detta vill man att en social robot ska kunna lära sig beteenden, och utveckla sin
sociala förmåga allt eftersom man interagerar med människor i sin miljö.

3 Kismets uppbyggnad

För att robot ska kunna agera så att den verkar intelligent, och dessutom ska kunna följa
sociala regler krävs en hel uppsättning med system som tar in och hanterar information på
bästa möjliga sätt. För dessa system redogörs nedan.

3.1 Yttre system

De system som gör att Kismet över huvud taget kan ta in input, och ge ifrån sig output kan
callas de yttre systemen.

3.1.1 Visuella systemet

Ett antal kameror är monterade på Kismet. Det sitter en kamera i båda pupillerna. Dessa
används för t.ex. ögonigenkännning som beskrivs senare i rapporten. Det är även två kameror
monterade mitt i ansiktet, där man kan tänka att en näsa skulle vara placerad. Dessa kameror
används för att avgöra avstånd till objekt i rummet, och för att ha en lite mer övergripande
bild över rummet för att kunna se om något nytt tillförs som uppmärksamheten bör riktas till.
Det krävs 9 datorer som arbetar tillsammans för att hantera informationen från det visuella
systemet.

3.1.2 Auditiva systemet

Det finns en mikrofon kopplad till Kismet. Så som det är i nuläget får personen som pratar
med Kismet sätta på sig mikrofonen så att den är nära munnen, eftersom det tas upp alldeles
för mycket störande ljud om mikrofonen skulle vara placerad på Kismet. Man använder
information om röstläge och prosodi, alltså betoning av ord för att avgöra vilka sociala
ledtrådar människan skickar ut. Man kan även ta upp information om vilka exakta fonem som
uttalas, men man använder inte den informationen än då man främst vill fokusera på tonläge.
3.1.3 Vokala systemet

3.1.4 Uttrycksystemet

3.2 Inre systemet
Det inre systemet utgörs av sex subsystem, som tillsammans hanterar informationen som kommer in från de yttre systemen. På bilden nedan ges en översikt för hur systemet är sammansatt.

3.2.1 Särdragssystemet
Det auditiva delen avgör vilka ljud som kommer från en människa som talar, och vilka ljud som produceras på ett annat sätt. Detta är avgörande för hur Kismet ska agera. År det en människa som talar ska han slå på det sociala systemet, är det t.ex. en leksak som rasslar ska han agera på ett annat sätt.

### 3.2.2 Det perceptuella systemet

Det perceptuella får in de utvalda särdragen från särdragssystemet, och omvandlar dessa till ”riktiga” percept, som roboten sedan kan ta nytta av när den avgör vilka motoriska funktioner som ska utföras som output.

### 3.2.3 Uppmärksamhetssystem

För att Kismet ska kunna efterlikna en människa så mycket som möjligt krävs det att det går att fokusera robotens uppmärksamhet på samma sätt som man fokuserar en människas uppmärksamhet på något. Man vill också få någon slags respons tillbaka, som visar att roboten har uppmärksamheten på det man vill, och i Kismets fall visas det genom att ögonen riktas mot det objekt han har uppmärksamheten på.

Kismets uppmärksamhetssystem bygger på att percepten från perceptsystemet viktas, alltså blir olika viktiga att uppmärksamma, beroende på hur motivationen och habitueringen är, vilket beskrivs närmare senare i rapporten. De saker som generellt fångar uppmärksamhet är objekt som rör sig, är färgade, eller ser ut som ett ansikte. Är t.ex. roboten motiverad till att leka, kommer uppmärksamheten lättare riktas mot en färgglad leksak, än mot ett ansikte. Detta nämns också mer om senare i rapporten.

### 3.2.4 Motivationssystemet


### 3.2.5 Beteendesystemet

På nivå 0, den översta markerat i lila, avgörs vilken av Kismets tre grundläggande behov som ska tillfredsställas. För att avgöra detta får varje möjligt beteende input från dess motsvarande drift, och det beteende som har störst drift är det som blir valt att tillfredsställas. Även yttre faktorer från världen påverkar valet.


På nivå 2 läggs ytterligare en aspekt till vilket beteende som ska väljas. Har man i det tidigare steget valt att dra sig bort från stimulit då det är ett ”dåligt” stimuli avgörs lite mer noggrant vilket ansiktsuttryck som ska visas beroende vilken slags dåligt stimuli det är.

3.2.6 Motoriska systemet

När informationen nu har processerats genom de olika systemen som i slutänden beslutar vilket beteende som ska visas krävs också ett motoriskt system som gör att dessa beteenden kan uttryckas genom att fysiskt ändra på ansiaktsuttryck, och att kunna uttrycka ljud. Till detta används tre undersystem. Ett system som styr ögonrörelserna, ett system som styr ansiktsuttrycken och ett auditivt system, som gör att Kismet kan ge ifrån sig ljud.

3.2.6.1 Okulär kontroll


3.2.6.2 Ansiktsuttryck

Ansiaktsuttryck är mycket viktiga vid kommunikation, då det hjälper till att förstärka och förydliga den bakomliggande meningen av det man säger. Ett snabbt ansiaktsuttryck kan visa vilket känslosmässigt tillstånd man är i, och det blir mycket lättare för den man konverserar med att ta till sig informationen på rätt sätt. Ansiaktsuttrycken hjälper även till att ge ledtrådar vid turtagning i konversationen.

Kismet ligger inte på en fullt utvecklad nivå när man ser till användandet av ansiaktsuttryck, utan ligger mer på nivån av ett litet barn, som har nämnts tidigare. Fyra grundfunktioner har programmerats in i roboten. Dels de känslosmässiga ansiaktsuttrycken, som visar vilket känslosmässigt tillstånd roboten befinner sig i, dels de ansiaktsuttryck som visar på sociala ledträd i konversationen, så kallade facial displays. Kismet kan även använda ansiaktsuttryck för att försöka skydda sig från hotande stimuli, t.ex. genom att stänga ögonen. Slutligen synkroniseras läpprörelserna med robotens tal, detta för att ge ett naturligt intryck.

Fördelen med att använda detta system är att varje ansiktsuttryck svarar mot en underliggande känsla. Det går även att se ett kontinuerligt byte mellan ansiktsuttryck.

3.2.6.3 Auditivt uttryck

När man designat Kismets talsystem har man tagit hänsyn till hur talet ska fungera uttrycksmässigt och hur det underlättar för social interaktion. Som det är nu klarar Kismet av proto-dialoger, dvs. med hjälp av sociala ledtrådar och ett system som gör att betoning i talet kan vara olika kan han bedriva en konversation, även fast det inte är några riktiga meningar som sägs. Talet i sig kan säkert vara bättre, men i och med att Kismets ansiktstuttryck är med och förstärker känslor och intentioner bakom talet fås meningen fram i de auditiva yttrandena.

Precis som många andra system i Kismets design har utvecklingen påverkas av hur motsvarande system fungerar hos människan. Forskning har visat på att när känslorna hos människan ändras, ändras även förutsättningar i kroppen såsom puls, blodtryck, andning och annat. Detta påverkar hur talet låter, och man försöker få Kismet att fungera så det verkar som att han har samma bakomliggande system.

Talsystemet är även utformat så att Kismet kan förmedla och ta emot intentioner i tal, dvs. kunna känna igen om talaren vill förmedla tröst, uppmuntran, förbud eller dylikt. Detta är, som många andra system, byggt för att efterlikna det mänskliga systemet. Det ligger en tanke bakom att Kismet är designat för att likna en liten person, eller ett djur, nämligen att människan automatiskt ändrar sitt tonfall och röstläge till ett lite ”gulligare” prat. Detta medför att vi överdriver intentionerna i vårt språk, och detta underlättar för Kismets förmåga att förstå och lära sig intentioner i tal, som han sedan kan använda sig av för att själv förmedla intentioner i sitt eget språk.
4 Diskussion


Arbetet med Kismet, och forskningen kring robotar in allmänhet är ett mycket fascinerande område och jag tror att man kommer gå långt inom det, men hur intelligent maskinerna än verkar så är det människan som har skapat dem. Därmed kommer människan alltid vara mer intelligent, och detta tror jag medför att robotarna inte kommer få så stor betydelse i det sociala livet som man kanske kan hoppas, utan att deras främsta användningsområde kommer vara kvar i de miljöer där de kan underlätta och hjälpa människan i fysiska arbeten.
5 Källor

Cynthia Breazeal, Brian Scassellati (1999) ”How to build robot that make friends and influence people”

Cynthia Breazeal, Brian Scassellati (1999) “A context-dependent attention system for a social robot”

http://www.ai.mit.edu/projects/sociable/overview.html
http://www.ai.mit.edu/projects/sociable/kismet.html
http://www.ai.mit.edu/projects/sociable/visual-attention.html
http://www.ai.mit.edu/projects/sociable/emotions.html