

En schack AI som använder case-based reasoning (översättning?)

Undertitel (?)

A Case-Based Reasoning approach to chess AI

Subtitle (?)

Examensarbete inom huvudområdet Datavetenskap

Grundnivå 30 högskolepoäng

Vårtermin 2015

Johannes Qvarford

Handledare: Namn Namnsson

Examinator: Namn Namnsson

Sammanfattning

[Direkt efter titelsidan ska det finnas en sammanfattning, som omfattar högst 150 ord. Denna sida ingår inte i ordinarie sidnumrering. Sammanfattningen ska i kort form sammanfatta hela arbetet inklusive bakgrund, problemformulering, genomförande/implementation, resultat och framtida arbete. Den ska därmed vara det sista som skrivs i rapporten. Den behöver därför endast vara med i den rapport som lämnas in för opponering (och den slutgiltiga rapporten). I tidigare inlämningar kan denna text lämnas oförändrad.]

**Nyckelord:** [minst tre nyckelord men inte allt för många, helst inte fler än sex stycken]

Innehållsförteckning

1 Introduktion 1

2 Bakgrund 2

2.1 Underrubrik bakgrund 2

2.1.1 Underrubrik till underrubrik 2

2.2 Exempel på hur figurer kan användas 2

2.3 Tabeller 4

2.4 Citat 4

3 Problemformulering 5

3.1 Metodbeskrivning 5

4 Genomförande/Implementation/ Projektbeskrivning 6

4.1 Research / Förstudie 6

4.2 Progressionsexempel: modellering 6

5 Utvärdering 8

5.1 Presentation av undersökning 8

5.2 Analys 8

5.3 Slutsatser 8

6 Avslutande diskussion 9

6.1 Sammanfattning 9

6.2 Diskussion 9

6.3 Framtida arbete 9

Referenser 10

[En korrekt innehållsförteckning ska genereras inför varje inlämning. Den ska ha rubriken ”Innehållsförteckning” och vara centrerad, och 18 punkter stor. Innehållsförteckningen ska inte ingå i den huvudsakliga sidnumreringen. Innehållsförteckningen ska själv inte vara med i innehållsförteckningen (denna behöver ibland tas bort manuellt beroende på hur den genereras). En innehållsförteckning ska maximalt innehålla tre undernivåer (till exempel 1.1.2). Fyra nivåer (till exempel 1.1.2.4) är därför inte tillåtet i rapportstrukturen.

Innehållsförteckningen ska ha samma typsnitt som rubrikerna i rapporten men kan med fördel vara av lite mindre storlek än rubrikerna.

Rubrikerna sammanfattning, innehållsförteckning och appendix ska *inte* finnas med i Innehållsförteckningen.]

# Introduktion

**(Schacks ursprung är komplicerat, och det finns ett flertal myter om hur det uppkom i land X år Y. Jag kan heller inte få tag på den definitiva förstahandskällan som alla använder.)**

**(En del schackmaskinhistoria bör kanske flyttas till bakgrund?)**

Schack har en lång historia som sträcker sig ända till början av 600-talet e.Kr.. Schack har traditionellt spelats mellan två människor, men under det senaste århundradet har även maskiner utvecklats för att spela spelet. Dessa maskiner har visat sig kunna mäta sig med mänskliga spelare, och redan under 70-talet utvecklades en maskin som kunde besegra en stormästare (Levy, 1978). Sedan dess har utvecklingen bara fortsatt, och 1997 besegrades den då regerande världsmästaren Garri Kasparov av en schackspelande maskin vid namn ”Deep Blue” (Cambell, Hoane & Hsu, 2001).Detta innebär inte att artificiell intelligens (AI) för schack inte har rum att utvecklas; Deep Blue var en maskin med specialanpassad hårdvara, så det är inte säkert att dess framgång går att replikera på persondatorer och spelkonsoler.

Ett problem med att utveckla artificiell intelligens (AI) som kan spela schack är att det är väldigt beräkningskrävande att räkna ut det bästa draget att utföra i ett givet läge med hjälp av naiva algoritmer. Om AI:n kan utföra drag a, b eller c, och motståndaren kan utföra drag d, e eller f efteråt, måste alla lägen ad, ae, af, bd, be, bf, cd, ce och cf undersökas för att se vilket drag som resulterade i bäst situation för AI:n, givet att motståndaren gör sitt bästa. För varje extra steg in i framtiden som AI:n vill se måste x gånger så många fall undersökas, där x är antalet drag som kan utföras i varje läge. XXX(källa) estimerar att antalet möjliga slutlägen är minst 10^YY, och det är i nuläget omöjligt för en dator att undersöka så många lägen under en rimlig tid.

**(kan diskussion om tidigare forskning inbakad i denna paragraf, eller separat?)**

Syftet med det här arbetet är att presentera ett alternativ till ”brute force” tekniken. Case-based reasoning (CBR) är en teknik för att utveckla AI:s som bygger på att lösa problem baserat på lösningar av tidigare, liknande problem. När en AI ska göra ett drag kan den härma vad en expert gjort i samma fall, genom att konsultera en falldatabas. Alla möjliga fall kan inte lagras eftersom det inte finns tillräckligt med minne för att representera dem, så AI:n måste basera sitt val på det expertfall som är mest likt det nuvarande fallet.

I det här arbetet ska ett system skapas för att producera AI-agenter för schack som använder CBR. Karaktärernas falldatabas kommer kunna baseras på tidigare spelade partier dokumenterade i algebraisk schacknotation (AN). Systemet ämnar att skapa AI-agenter som spelar på en likvärdig nivå som de spelare vars matcher använts för att bygga agenternas falldatabaser.**(Det känns mer ambitiöst och användbart för forskningen att skapa ett generellt system istället för en specifik AI-agent.) (Hur jag ska testa systemet/specifika producerade AI-agenter vet jag inte.)**

[Introduktionen ska innehålla en sammanfattning av bakgrund, problem, metod och genomförande/implementation. Inför varje inlämning behöver bara de relevanta delarna av introduktionen lämnas in. Tänk på att lämna plats i introduktionen för stycket om genomförande vid första inlämningen. Introduktionen bör revideras vid varje inlämning eftersom arbetet förändras över tid.

För att separera innehållet i introduktionen ska den indelas i onumrerade stycken med hjälp av radbrytningar (i likhet med dokumentmallen). Styckena ska inte vara kortare än 4-5 rader text och innehålla ett sammanhållet ämne - till exempel bakgrund om datorspel.

Introduktionen ska *inte* ha numrerade underkapitel eller figurer och vara ungefär en sida lång (inte under en halv sida och inte mycket mer än en sida).

Introduktionen ska ha sidnummer 1.]

# Bakgrund

## Schack

### Regler

**(Regler till schack med bilder osv.)**

### Algebraisk notation

**(Om algebraisk notation)**

## Case-based Reasoning

**(Vet inte om det finns tillräckligt mycket att prata om som inte tas upp i metoden)**

[Bakgrundskapitlet innehåller teoretisk bakgrund som är intressant för problemområdet. Det inleds ofta med mera generell bakgrund för att mot slutet av bakgrundskapitlet avsluta med underkapitel som är mera specifika för problemet. Bakgrunden ska innehålla referenser.

Bakgrunden ska inledningsvis hållas generell för att i senare avsnitt gå in på mer specifika frågor och närma sig problemformuleringen.

Referenser i texten ska anges enligt harvardsystemet (exempel på referenslista sist i mallen):

Artiklar med en författare kan refereras på följande sätt:

… Ernefeldt (2008) menar att...

… I en tidigare studie (Ernefeldt, 2008) argumenteras det för att...

Två författare:

… Nya synsätt (Salen & Zimmerman, 2005) påverkar hur …

... Salen och Zimmerman (2005) menar att …

Om flera författare: första gången man anger en referens med flera författare kan man ange samtliga för att senare använda m.fl. Om det är väldigt många författare är det tillåtet att endast använd m.fl. varianten.

… Lundell, Lings och Syberfeldt (2011) visade att…

… Lundell m.fl. (2011) visade att…]

## Underrubrik bakgrund

[Ett numrerat underkapitel ska inte vara kortare än 5-6 rader. Mycket korta numrerade kapitel (bara någon rad per kapitel) slås samman till ett större kapitel utan separat numrerade underkapitel. På motsvarande sätt behöver väldigt långa kapitel, som sträcker sig över mer än två sidor, delas upp i mindre numrerade delar så att sidnumreringen i innehållsförteckningen inte sträcker sig över ett väldigt stort antal sidor.]

### Underrubrik till underrubrik

[En underrubrik kan i sin tur innehålla underrubriker men rapporten ska inte ha djupare nivåer än så. Det ska alltså inte finnas en fjärde nivå, t ex 2.1.1.1. I övrigt gäller samma principer om längd som för underrubriker.]

## Exempel på hur figurer kan användas

[Figurer ska refereras tydligt i texten med figurnummer. Figurer ska vara i högupplöst format som ska gå att skriva ut i hög upplösning och i svart-vitt**,** se figur 1. Det finns därför verktyg som oftast är olämpliga för att rita figurer med, till exempel *Microsoft Paint* och liknande applikationer. Undantaget från denna regel är olika former av skärmdumpar (figur 2) eller fotografier som har pixelformat i sitt original (figur 3).

Skärmdumpar från datorprogram och spel kan användas för att illustrerar någon viktig aspekt för arbetet. En källangivelse måste anges som tydligt pekar ut programmet och utgivaren. Om figurer och skärmdumpar inte är egenproducerade måste det tydligt framgå vad syftet är med att ha med just denna bild. Går samma syfte att uppnå med en egenskapad bild så ska detta alternativ väljas.

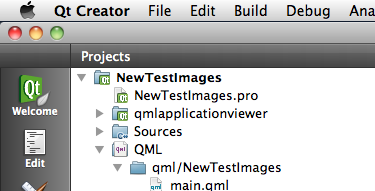
Figurer ska ha numrerad figurmarkör som är centrerad och i **fetstil (bold)** under själva figuren. Figurtexten ska ge en tydlig förklaring till vad figuren illustrerar. Alla figurer i rapporten ska vara refererade i texten.

Alla figurer ska vara centrerade och de får inte vara inbakade i texten, det ska med andra ord inte finnas text till vänster eller höger om en figur. Figurer ska helst sträckas ut så att de upptar sidans hela bredd. I de fall figuren har ett innehåll som är väldigt högt och smalt så kan en smal figur användas.]



1. Exempel på en figur med bra kvalité

Genom att ange stilen (*Style*) "figur" till själva bilden så kommer ordbehandlaren hålla ihop figur och figurtext (som ska ha stilen "figurtext") så att de alltid hamnar på samma sida.



1. Exempel på en skärmdump i pixelformat

## Tabeller

[Tabeller ska hanteras separat från figurer. De ska numreras oberoende av figurerna (Tabell 1, Tabell 2 osv). Tabeller ska ha numrerad markör som är centrerad och i **fetstil (bold)** *över* själva tabellen, se tabell 1. Tabelltexten ska ge en tydlig förklaring till vad den illustrerar. Alla tabeller i rapporten ska vara refererade i texten.]

1. Exempel på hur en tabell kan se ut

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Typ** | **Test A** | **Test B** | **Test C** | **Test D** |
| Grön | 45% | 30% | 44% | 68% |
| Röd | 55% | 70% | 56% | 32% |

## Citat

Citat används när den exakta ordalydelsen i refererade texter behöver lyftas fram. Det är väldigt viktigt att det tydligt framgår var citatet är hämtat (källhänvisning med sidnummer) och texten får inte ändras på något sätt. Kortare citat kan komma löpande i texten och skrivs då omgivet med citattecken.

Ett exempel på kortare citat: Roberts, Furst, Dorn och Isbell (2009, s. 25) beskriver sitt ramverk som ett "web-based mixed-media choose-your-own-adventure-style interactive storytelling system".

Längre citat ska anges som ett eget stycke som är indraget. I mallen kan stilen "Citat" användas. Det kan till exempel se ut så här:

Our framework for evaluating our approach is a web-based mixed-media choose-your-own-adventure-style interactive storytelling system. Our system displays a sequence of authored text and videos that comprise narrative units, or events, that are linked together by explicit decision points for the player.

Roberts, Furst, Dorn & Isbell, 2009, s. 25

Källhänvisningen placeras direkt under citatet och ska inte vara kursiv. I mallen kan stilen "CitatRef" användas.

# Problemformulering

## Problembeskrivning

Syftet med arbetet är att skapa ett system som gör det möjligt att skapa AI-agenter med hjälp av ett antal dokumenterade matcher. Det finns två steg i denna process, att generera falldatabaserna, och att utveckla CBR-tolken av AI-agenten som kan applicera CBR på falldatabaserna.

### Generera falldatabaser

För att kunna skapa stora och varierande falldatabaser måste en stor mängd data samlas in från experter. Lyckligtvis har många matcher dokumenterats genom åren, antingen genom att indikera sekvensen av drag som spelarna utförde, eller pjäsernas koordinater efter varje drag. Det finns många olika sorters notationer, men den vanligaste är algebraisk notation. Drag i algebraisk notation är väldigt kortfattade, och inte mer explicita än nödvändigt. Detta kan underlätta för människor att läsa, men jobbigt för datorer. För att minska beräkningskraften som krävs för att tolka dessa matcher under körtid, kommer ett program skrivas som översätter ett antal matcher till ett mer lättolkat format i förväg.

För att underlätta för datorn används ett mer entydigt format. Positionen för en pjäs representeras av positionens nummerkoordinat respektive bokstavskoordinat. Ett drag som flyttar en pjäs från en position till en annan representeras av från- och tillpositionen. Detta format har fördelen att alla drag tar lika mycket utrymme.

Ett fall representeras med en tvådimensionell matris, där ett element M\_ij representerar vilken sorts pjäs som finns på positionen med koordinaterna (i, j), eller att det inte finns någon.**(detta är antagligen inte den bästa representationen, men jag behöver undersöka hur liknelsealgoritmen kan implementeras innan jag kan ta ett beslut.)**

### CBR-tolken

CBR-tolkens uppgift är att bestämma ett drag att utföra i ett visst läge, givet dess falldatabas. Den gör detta genom att undersöka alla fall i listan, och göra det drag i fallet vars läge är mest likt det givna läget. En liknelsefunktion används för att gradera hur pass lika två lägen är. Den är utformad som följande: **(ekvationen är förstorad för att det inte gick att läsa texten under sigma.)**

Där a och b är två lägen och P är en funktion som ger 1 om båda pjästyperna är lika, och noll om inte (om båda saknar pjäs så är de lika). **(bra med ekvationer? Mer/mindre?)**

[Problemformuleringen ska detaljerat redogöra för det problem examensarbetet är baserat på. Problemet ska först beskrivas övergripande för att senare i problemformuleringen beskrivas mer detaljerat.]

## Metodbeskrivning

[Problemformulering ska även innefatta en metodbeskrivning som ger en tydlig bild av hur frågeställningen ska besvaras, undersökas och arbetet utvärderas.]

# Genomförande/Implementation/ Projektbeskrivning

[Kapitlet kan rubriksättas *Genomförande* eller *Implementation* eller *Projektbeskrivning* eller liknande. Innehåll kan variera beroende på projektets innehåll men nedan finns förslag på upplägg.

Viktigt i detta kapitel är beskrivningen av progression och designval, och då även att diskutera och problematisera alternativ som valdes bort, varför detta skedde och stickspår till den valda lösningen.

För ett grafikarbete kan detta exempelvis betyda att ofärdiga, kommenterade skisser presenteras i rapporten. Detta för att visa på progressionen och designval under arbetets gång. För en programmerare kan detta innebära att algoritmer och lösningar som förkastats vägs mot den valda lösningen.]

## Research / Förstudie

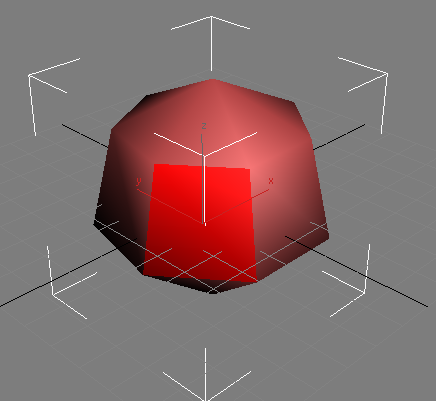
[Ett research- eller förstudiekapitel kan vara viktigt. Om man inspirerats av något, exempelvis ett program, en film, en bok eller något annat så kan detta listas under denna rubrik.

Källor som underbygger genomförandet men inte problemet eller metoden kan ligga i bakgrunden, men de passar ofta bättre i ett separat kapitel i början av genomförandedelen än som en del av bakgrunden.

Så mycket som möjligt av det som inspirerat genomförandet ska listas i detta kapitel]

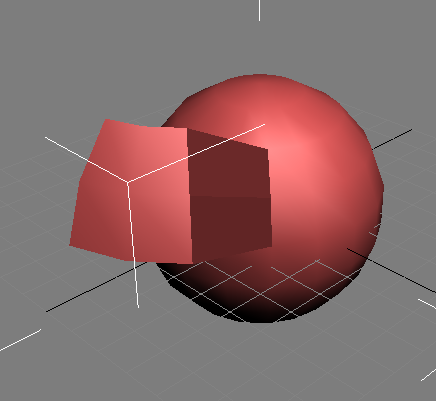
## Progressionsexempel: modellering

[Exempel: progression och designval inom området grafik. Till en början användes ett sfäriskt utgångsobjekt med få hörn. Detta gav tyvärr ett alltför kantigt utseende, se figur 3.



1. Ett sfäriskt objekt med få hörn

Alternativet blev då att använda ett objekt som hade fler hörn, se figur 4. De främsta nackdelarna med att använda så många hörn är att man får poler där hörnen går samman i toppen och botten av objektet samt att det ger lägre skärmuppdateringshastighet på grund av fler hörn i modellen.]



1. Ett sfäriskt objekt med många hörn

# Utvärdering

[Kapitlet *utvärdering* ska innehålla en presentation av den genomförda undersökningen, en analys av utfallet och de slutsatser som kan dras därav.

Kapitlet kan variera ifråga om struktur beroende på projektets utformning.

]

## Presentation av undersökning

## Analys

## Slutsatser

# Avslutande diskussion

## Sammanfattning

[

Här sammanfattas rapporten som helhet från frågeställning till slutsats. En läsare som vill få en snabb överblick av uppsatsen ska kunna gå direkt från kapitel ett till kapitel sex.

]

## Diskussion

[I diskussionskapitlet sätts problemet och resultatet i ett större sammanhang utanför examensarbetets specifika problemformulering. Det är viktigt att koppla till andras arbete - till exempel centrala artiklar som används i bakgrundskapitlet.

Diskutera resultatets trovärdighet. Lyft fram faktorer som påverkar trovärdigheten. Diskutera utifrån den genomförda studien.

Diskussionskapitlet är även en lämplig plats för att ta upp de saker som täcks av kriteriet ”Rapporten innehåller en diskussion kring relevanta samhälleliga och etiska aspekter på arbetet”. Välj några av nedanstående aspekter och diskutera dem i relation till ditt arbete. Beroende på typen av arbete varierar vilken typ av aspekter som är relevanta att undersöka, om man till exempel gjort ett spel eller en illustration som innehåller människor kan genus och kulturella aspekter vara intressant medan samhällelig nytta blir intressant om man har utvecklat en algoritm eller ett program.

* Etiska aspekter
* Forskningsetiska aspekter hos arbetet eller undersökningsmetoden
* Samhällelig nytta hos arbetet
* Genus och/eller kulturella aspekter

]

## Framtida arbete

[I det framtida arbetet ska en hypotetisk fortsättning på examensarbetet diskuteras. Detta gäller både i det korta perspektivet - om arbetet skulle fortsätta några extra dagar eller månader. Projektet ska även sättas in i ett större sammanhang och ses ur ett bredare perspektiv. Till exempel om ett företag skulle kunna fortsätta på examensarbetet och driva resultatet vidare för att bygga ett komplett spel. Detta kan med fördel relateras till det större sammanhanget som diskuteras i diskussionsdelen.]

Referenser

**(Den första är referensen från google scolars citationsverktyg och den andra är handskriven enligt de APA referenser jag kan hitta. Något att kommentera?)**

Campbell, M., Hoane, A. J., & Hsu, F. H. (2002). Deep blue. *Artificial intelligence*, *134*(1), 57-83.

Campbell, M. Hoane, A. J., Jr, & Hsu F. (2001) Deep Blue. Artificial intelligence, 134(1-2)

Hapgood, F (23–30 **(Den släpps en gång i veckan, men APA har bara regler för år och månad+år?)** December 1982). "Computer chess bad-human chess worse". New Scientist. pp. 827–830.

Levy, D. (1978) Man beats machine, Chess Life & Review, November 1978, pp. 600–603, at pp. 601–03.

case based reasoning. (n.d.). The Free On-line Dictionary of Computing. Retrieved February 01, 2015, from Dictionary.com website: http://dictionary.reference.com/browse/case based reasoning

Electronic Arts (2011) *Battlefield 3* (Version: 1.0) [Datorprogram]. Electronic Arts. Tillgänglig på Internet: http://www.battlefield.com/battlefield3.

Ernefeldt, E. (2008) *Phun - a 2D physics playground - GameDev.net*. 13 February 2008. Phun -a 2D physics playground. Tillgänglig på Internet: http://www.gamedev.net/topic/482775-phun---a-2d-physics-playground/ [Hämtad December 13, 2011].

Lundell, B., Lings, B. & Syberfeldt, A. (2011) Practitioner perceptions of Open Source software in the embedded systems area. *Journal of Systems and Software*. 84 (9), s. 1540–1549.

Roberts, D.L., Furst, M.L., Dorn, B. & Isbell, C.L. (2009) Using influence and persuasion to shape player experiences. *Proceedings of the 2009 ACM SIGGRAPH Symposium on Video Games. Sandbox ’09. New York*, NY, USA, ACM. s. 23–30.

Salen, K. & Zimmerman, E. (red.) (2005) *The game design reader: a Rules of play anthology*. Cambridge, MA: MIT Press.

[Kontrollera noga med sökfunktionen att alla referenser som anges i referenslistan faktiskt finns i brödtexten. Kontrollera på samma sätt med sökfunktionen att alla referenser i brödtexten finns i referenslistan. Var noga med att samtliga referenser formateras enligt Harvard. Ovan finns exempel på ett antal referenser som följer skolans format på referenslistor enligt Martin G Eriksson (se kursens filsamling).

Referenserna ovan är av typen: datorprogram, webbsida, journalartikel, konferensartikel samt bok.

Datorspel ska refereras till som datorprogram även om de spelas på spelkonsol. I automatiserade referenshanteringssystem så används oftast versionsnumret för att skilja på böcker och datorprogram. Spelen bör därför ha ett versionsnummer, speciellt om spelen finns i flera olika versioner.

Referenslistan ska vara uppställd i bokstavsordning, med utgångspunkt från författarens/redaktörens efternamn. Referenserna ska utgöras av en enhetlig uppställning och *inte* delas in i sektioner beroende på typ, webbsidor, böcker och artiklar av olika slag.

Notera att det bara är referenser som *endast* finns på internet som ska vara listade som Tillgänglig på internet. Alltså ska referenserna inte ha ”Tillgänglig på internet” om referensen finns i tryckt form. URL:er ska inte vara understrukna eller blå.]

1. Designdokument etc.

[Appendix ska fungera som referenslistan - dvs det ska finnas referenser till den från texten. Appendix ska inte vara numrerade utan ska namnges med: Appendix A, Appendix B osv. De ska vara sidnumrerade (I, II, III ...) men de ska inte finnas med i innehållsförteckningen. Varje nytt appendix ska börja på toppen av sidan.]