

Oppositionsrapport på Datorassistans vid lyrikskrivande*

Rapportförfattare: Fredrik Bystam, Jonas Sköld

Christopher Teljstedt, 901118-5775, CDATE 3

Examensarbete inom datalogi, grundnivå, DD143X

24 april 2013

Sammanfattning

Detta är en oppositionsrapport som ämnar att lyfta fram styrkor och svagheter i kandidatexamensarbetet Datorassistans vid lyrikskrivande. I rapporten analyseras kandidatexamensarbetet samt förlag ges för hur den helhetliga kvaliteten i examensarbetet kan förbättras.

1 Introduktion

Denna rapport kommer att ta fram styrkor och svagheter med kandidatexamensarbetet Datorassistans vid lyrikskrivande (Lyrics Writing Assistant). Avhandlingen fokuserar på att undersöka hur väl ett datorprogram kan assistera vid skrivande av lyrik samt avgöra om ett assisterande program är användbart. För att undersöka detta skapades en testmjukvara baserat på definitioner av rim samt poetiska kriterier från bland annat *Barron's SAT Subject Test Literature 2009* och från hemsidan *The Poetry Archive*[3, 4, 5].

Egen efterforskning av liknande studier - vars mål generellt var att automatisera generering av lyrik och poesi på olika vis[6, 7, 8] - resulterade i fyndet "An Electronic Assistant for Poetry Writting". Rapporten hade ändamålet att undersöka assistering i skrivandet av poesi. I rapporten förklaras utvecklingen av ett program som utforskar möjligheten att assistera genom klassifiering av poem, förslag till verser samt avslutande ord till meningar med hjälp av n-gram som språkmodell. Relativt lite forskning är publicerat inom detta område och därför är det lite anmärkningsvärt att inga referenser fanns till denna rapport. Trots en snarlik titel och innehåll har detta kandidatexamensarbete implementerat en egen mjukvara och tillskillnad från den funna rapporten har denna avhandling dessutom utfört tester. Vidare har svenska ord använts som ordbank i testen vilket inte har gjorts i några rapporter som hittades i efterforskningen.

Pågrund av detta har studien fått en unik nisch som får mig att anse att denna avhandling ändå har lite nyhetsvärde¹

*Lyrics Writing Assistant

¹Trots att svenska ord som ordbank inte har vägts in i diskussionen av testresultaten.

2 Viktiga anmärkningar

- En komplett rapport
- Väldigt bra språkbruk
- Arbetet har nyhetsvärde
- Beskrivning av metologin är inte riktigt komplett
- Testupplägget och kvalitén i resultaten kan ifrågasättas

3 Innehåll, grammatik & formalia

I sin helhet är det en väldigt trevlig läsning med näst intill inga språkmissar. Språket håller en lagom nivå och författarna är noga med att förklara nya begrepp så att läsaren inte är lämnad i förvirring. Det är lättlästa och koncisa meningar utan felaktiga tempusväxlingar. Däremot förekom en och annan förkortning, till exempel "d.v.s." som bör skrivas ut.

Avhandlingen är generellt väldigt välformaterad med en beskrivande titel samt sammanfattning som reflekterar innehållet väl. Introduktionen är uppstrukturerad med en inledande text, problemformulering samt en presentation av rapportens upplägg. Bakgrunden tar på ett väldigt bra vis upp definitioner av begrepp, viktiga kriterier samt förklarar språkfenomen som har relevans för rapporten.

Vidare från introduktionen finns en hyfsad röd tråd genom hela avhandlingen med några undantag i kapitlet "Metod". Det saknas en tydligen beskrivning av tillvägagångssättet för att besvara frågeställningen. Som läsare antar jag att det bör ha funnits tankar hos författarna om hur utvecklingen av programmet skulle ske, vilka funktioner som skulle inkluderas och hur de skulle testas. Detta bör i en tydligare form ha presenterats före sektionen "Mjukvara" som mer i detalj berättar hur programmet är implementerat. Det skulle också knyta ihop ge en bättre tråd till förklaringen av utförandet av tester som inleder resultaten.

Författarna stöder fakta och påståenden som presenteras trovärdiga källor på ett korrekt vis. Somliga kanske kritiserar referenserna till men jag anser att poesi och definitioner av rim är så pass subjektiva att denna websida kan godkännas som referens.

I övrigt saknades även en förklarande text till tabellen i resultatsektionen. För att vara helt korrekt bör tabelltexten placeras ovanför tabellen och de resterande figurtexterna i avhandlingen bör flyttas ner under figurerna.

4 Upplägg av test

Målet med undersökningen var att avgöra vad det är som gör ett program som assisterar vid tillverkning av lyrik användbart. Det var således viktigt för författarna att ta reda på vilka faktorer som påverkar vad i deras program.

I avhandlingen testades tio personer att skriva tre julklappsrim var. Varav de två rim var under mer strikta förhållanden där programmet byggde i stort sett hela meningar och rimen var tvugna att handla om en "kniv" och "bok". Medan det sista rimmet var mer fritt och personerna fick skriva om en egen julklapp, bygga meningar och välja passande rim efter egen känsla. Vidare användes två olika typer av ordbanker under testning av programmet, en innehållandes svenska tidningsartiklar och en med dikter av svenska kända

poeter. Till första två rimmen slumpades en av ordbankerna ut medan i det fria testet fick personerna själva välja vilken ordbank de ville använda sig av.

På grund av upplägget blir det således svårt att dra vissa slutsatser kring träningsdatan samt "assistansinställningarnas" signifikans när inte utförs under samma förutsättningar i testen. I Tabell 1 går det inte att avgöra hur mycket som de friare förutsättningarna påverkar rimmen då testpersonerna har möjlighet att välja från båda ordbankerna. Det är till exempel möjligt att ett resultat med enbart dikter som ordbank under friare förutsättningar är sämre än att använda enbart tidningsartiklar under strikta förhållanden. Det går inte heller att konstatera om de friare förutsättningarna med valfri ordbank är bättre än med enbart tidningsartiklar som ordbank. Vad som gör det ytterligare svårt att jämföra resultaten är att under det friare testet fick rimmen handla om vad som helst till skillnad från det striktare förhållandet.

Detta kan förbättras genom att utföra testerna på ett mer systematiskt tillvägagångssätt. Det vill säga att vardera ordbank testas med samma mängd av "julkappar" som ska rimmas under både strikta och fria förhållanden. På så sätt får man mer säkra och signifikanta resultat.

5 Presentation av resultat

Läsaren måste kunna tyda testresultaten samt förstå dess signifikans. Därav är det viktigt att författarna presenterar hur mätningarna utförts samt vad de baseras på. Detta är särskilt viktigt när det handlar om subjektiva bedömningar som i denna avhandling. Om jag uttrycker mig lite barskt anser jag att testresultaten i den här avhandlingen är minst sagt "klent" presenterade och skulle kunnat få mer kärlek.

Testresultaten presenteras i en tabell som två genomsnittliga betygssättningar från dels de tio testpersonerna samt från författarna för varje typ av träningsdata, se Tabell 1 nedan. Bedömningen görs på julkappsrimns kvalitet på en skala 1 till 10 utan någon förklaring till hur dåligt eller bra dessa steg i skalan anses vara. Det blir på så sätt väldigt otydligt vad skalan betyder när det inte framgår vad skalan baseras på explicit.

I slutet av sektionen "Undersökning" i kapitlet "Metod" precis innan resultaten förklaras det att rimmen *bedöms av författarna efter tre kriterier*; hur väl de håller sig till ämnet, hur språkligt korrekta samt hur sammanhängande texterna är. Trots förberedelserna framkommer det inte explicit i resultaten att dessa har använts. Dessutom nämns det att författarna gör sin bedömning efter kriterier medan testpersonerna gör en egen bedömning på en skala 1 till 10 och utöver det svarar på frågor. I resultaten jämförs således data med två olika betygssgrunder, se Tabell 1.

Träningsdata \Betygsättare	Testpersonen	Författarna
Dikter	4,8	3,8
Tidningsartiklar	6,5	4,3
Fritt test	6,8	5,1

Tabell 1: Testresultat från avhandlingen. Se sektion 4.2 i Lyrics Writing Assistant

Följande är fyra förslag för att förbättra presentationen av resultaten:

- Ge testerna mer betydelse genom att presentera testresultaten i kategorier baserat på de nämnda kriterierna.
- Presentera vad varje typ av träningsdata fick för resultat på samtliga kriterier.
- Klargöra vad stegen i skalan baseras på - gör läsaren mindre kritiskt till rådande subjektivitet i evalueringen.
- Använd samma betygsgrunder för testpersonernas bedömning och författarnas bedömning - ger mer trovärdighet i resultaten.

Genom att göra dessa förtydliganden blir jag som läsare mindre kritisk till resultaten då bedömningarna i grunden är väldigt subjektiva. Dessutom leder det även till att en mer strukturerad diskussion av resultaten kan föras.

6 Diskussion och slutsatser

I avhandlingen diskuteras bland annat om att testet med fria förhållanden gav bättre rim på grund av att testpersonerna kan "på ett bättre sätt forma meningarna så de får den tilltänkta betydelsen". Problemet från testresultaten kvarstår dock, trots att detta är en rimlig och rationell diskussion så är inte resultaten som presenteras för läsaren en stabil grund för dra denna slutsats. Fler exempel på detta finns i diskussionen och mycket går inte att göras åt saken på grund av testernas utformning. Trots problematiken med resultaten förs ändå en bra diskussion och många bra slutsatser konstateras. Diskussionen förs sedan med en smidig övergång till slutsatser där problemställningen slutligen besvaras.

Referenser

- [1] C. Myers-Shaffer. *Barron's SAT Subject Test Literature 2009*. Barron's SAT Subject Test Literature. Barron's Educational Series, Incorporated, 2008. ISBN: 9780764138713. URL: <http://books.google.se/books?id=G5V4NcWUGu4C>.
- [2] *The Poetry Archive*. 2013. URL: <http://www.poetryarchive.org/>.
- [3] Poetryarchive. *The Poetry Archive. Definition of poetic form*. 2013. URL: <http://www.poetryarchive.org/poetryarchive/glossaryItem.do?id=8094>.
- [4] *The Poetry Archive. Definition of stress*. 2013. URL: <http://www.poetryarchive.org/poetryarchive/glossaryItem.do?id=8075>.
- [5] *The Poetry Archive. Definition of limerick*. 2013. URL: <http://www.poetryarchive.org/poetryarchive/glossaryItem.do?id=8082>.
- [6] Nuno J. Mamede m.fl. "An Electronic Assistant for Poetry Writing". I: *Advances in Artificial Intelligence - IBERAMIA 2004, 9th Ibero-American Conference on AI, Puebla, México, November 22-26, 2004, Proceedings*. Utg. av Christian Lemaître, Carlos A. Reyes och Jesus A. González. Vol. 3315. Lecture Notes in Computer Science. Springer, 2004, s. 286–294. ISBN: 3-540-23806-9. DOI: <http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=3315&page=286>.

-
- [7] Hisar Maruli Manurung m.fl. "A Flexible Integrated Architecture For Generating Poetic Texts". I: *In Proceedings of the Fourth Symposium on Natural Language Processing, Chiang Mai*. 2000, s. 7–22.
- [8] Nuno Mamede m.fl. *Poetry Assistant*.