

## Opponering på *Empty Rectangles* (grupp 21)

Simon Edwardsson

Olle Fredriksson

Daniel Gustafsson

Dan Rosén

May 25, 2010

För det första skulle vi vilja säga att vi tycker att arbetet är välskrivet, med en indelning som gör det lätt att överblicka och ett bra språk. De anmärkningar och funderingar vi har är mestadels småsaker.

Något som är lite oklart är huruvida algoritmerna ger den bästa lösningen eller en approximation. I de fall då de inte ger en optimal lösning skulle det vara intressant att ha med en jämförelse med en optimal lösning (om det är möjligt). Är det korrekt av oss att anta att era algoritmer ger en bättre lösning än Edmonds, och att det är anledningen till att man skall använda er? Är problemet NP-komplett? Isåfall kanske det borde poängteras mer.

En annan sak som vi har funderingar på är vart den åttonde algoritmen tagit vägen. I introduktionen och abstraktet står det åtta, men så vitt vi har förstått är det sju algoritmer som används i resultatdelen. Antalet algoritmer är stort, varför det skulle vara tydligare att istället fokusera på ett mindre antal algoritmer som istället skiljer sig åt mer. Det skulle även vara trevligt att redan från början ha kortare namn på algoritmerna så att det är lättare att komma ihåg vilken som är vilken när de återkommer i avsnitten.

Pseudokoden för algoritmerna skulle vara tydligare om variabler som är globala initieras före de används. Variablerna verkar även användas på sätt som är lite mystiska för oss; exempelvis används *currentSolution* både som en mängd och som en numerisk variabel. Om typerna skrevs ut explicit och variablerna användes på ett mer uniformt sätt skulle koden vara lättare att tyda.

Graferna i resultatavsnittet har vi några funderingar på: Vad beror sick-sack-formen på i figur 4.3? Allmänt skulle graferna som behandlar tid vara bättre med tiden i logaritmisk skala, eftersom man då lättare kan se hur komplexiteten hos olika algoritmer skiljer sig.

Ni nämner att det inte är intressant att titta på stora  $s$  inom datamining. Det är int helt klart för oss varför det är så. Ett konkret exempel hade hjälpt förståelsen.

Trots att vi kanske förväntas komma ihåg skillnaden mellan begreppen  $\mathcal{O}$ ,  $\Omega$  och  $\Theta$  skulle det ändå vara trevligt med en kort förklaring av uttrycken för att fräscha upp våra minnen (kanske som en fotnot).

Vidare till lite mer strukturella tankar:

Avsnittet "Experimenting with rectangles" har en lite informell titel, som

inte innefattar hela dess innehåll, trots att det är en av de viktigaste delarna i rapporten. Kanske skulle det namnet kunna ändras. På tal om titlar känns även rapportens titel lite tråkig med variabeln  $s$  som någon som inte läst rapporten bara kan gissa sig till betydelsen för.

En annan tanke är att man innan sektionerna börjar i ett nytt kapitel kan ha en liten övergångstext som berättar för läsaren var i texten hon är och vad kommande kapitel handlar om. Dessa texter skulle kunna knyta an till sektion 1.5, "Outline", i er rapport.

Avsnittet "Implementations" känns malplacerat och inte helt relevant för läsaren. Det skulle kunna vara med i en bilaga. (Dessutom är det inte helt faktakorrekt, då även C# vanligtvis kompileras till en virtuell maskin.)

Här kommer några mindre viktiga anmärkningar som vi ändå skulle vilja ta upp: Vi är osäkra på om  $\log^2 n$  betyder  $\log n * \log n$  eller  $\log(\log n)$ . Det vore bra att skriva vad det verkligen betyder första gången det används. Algoritm 3.3 har variabelnamn som ser konstiga ut (troligen L<sup>A</sup>T<sub>E</sub>Xs mathmode).

Det framgår inte klart om ni använder samma seed för era benchmarks, vilket det borde göra om det är så.

På sidan 41 skriver ni att exekutionstider inte presenteras eftersom de inte är intressanta, men de nämns ändå på sidan 42.

Tack för trevlig läsning!