58093 String Processing Algorithms (Autumn 2014)
Study Groups 3 (November 27)

Advance reading material. The following article describes a large number of algorithms for approximate string matching and summarizes their theoretical and experimental properties.

> G. Navarro: **A Guided tour to approximate string matching**.
> ACM Computing Surveys 33(1), 2001, pp. 31 - 88.
> `http://dx.doi.org/10.1145/375360.375365`

Topics for discussion. Read about the algorithm(s) assigned to your group before the study group meeting. In the meeting discuss at least the following topics and prepare to summarize the discussion to the members of the other groups.

- Describe the algorithm(s). How are they related to the algorithms on the lectures?

- How do the algorithm(s) compare against other algorithms in theory and/or in practice. Experimental and theoretical comparisons are summarized in Figures 28, 30 and 31 in the article.

You may also discuss additional topics, for example other algorithms in the same category.

Group A: Theoretical worst case

Members

- Greer, Maximillian
- Heino, Lauri
- Hoya Quecedo, Jose
- Hulkko, Heidi
- Leppänen, Jarno
- Wang, Ping

Algorithms

- Galil–Park (Sect. 5.2.6, see also 5.2.2 and 5.2.3)

Group B: Theoretical average case

Members

- Escoter I Torres, Llorenç
- Hurme, Teemu
- Kruglaia, Anna
- Pitkänen, Teemu
- Puuska, Samir
- Virolainen, Herkko

Algorithms

- Chang–Marr (Sect. 8.3.4, see also 8.3.1)
- Additionally, explain the significance of the bound $\alpha \leq 1 - e/\sqrt{\sigma}$ (see Sect. 4.1).

Group C: Practical nonfiltering

Members

- Chinnasamy, Mohan
- Goryachev, Vladimir
- Hopp, Joshua
- Lagus, Jarkko
- Nidia, Obscura
- Siciliano, Gianvito

Algorithms

- Wu–Manber (Sect. 7.2.1)
- Baeza-Yates–Navarro (Sect. 7.2.2)

Group D: Practical filtering

Members

- Faghihi Berenjegani, Farbod
- Kukkola, Ville
- Longi, Krista
- Paasiniemi, Markus
- Radev, Martin
- Viding, Jasu

Algorithms

- Baeza-Yates–Navarro (Sects. 8.2.4 and 8.2.5)