Advanced topics in data structures:

bibliography list on string matching

Algorithms to construct suffix trees in linear time are in [15, 12, 14, 4]. Suffix arrays were introduced in [11]. The linear time construction algorithm for suffix arrays is from [8]. The simple construction of the LCP array from the suffix array is from [9]. The mismatched counting problem is solved in [1]. The k-mismatch problem is discussed in [5, 10, 2]. Some of the stuff appears in the book by Gusfield [6]. Lowest common ancestor algorithms are in [7, 13, 3].

References

[1] Karl Abrahamson. Generalized string matching. *SIAM J. Comput.*, 16(6):1039–1051, 1987.

[2] Amihood Amir, Moshe Lewenstein, and Ely Porat. Faster algorithms for string matching with k mismatches. *J. Algorithms*, 50(2):257–275, 2004.

[3] Michael A. Bender, Richard Cole, Erik D. Demaine, Martin Farach-Colton, and Jack Zito. Two simplified algorithms for maintaining order in a list. In Rolf H. Möhring and Rajeev Raman, editors, *ESA*, volume 2461 of *Lecture Notes in Computer Science*, pages 152–164. Springer, 2002.

[4] Martin Farach-Colton, Paolo Ferragina, and S. Muthukrishnan. On the sorting-complexity of suffix tree construction. *J. ACM*, 47(6):987–1011, 2000.

[5] Z. Galil and R. Giancarlo. Data structures and algorithms for approximate string matching. *J. Complex.*, 4(1):33–72, 1988.

[6] Dan Gusfield. *Algorithms on Strings, Trees and Graphs.* Cambridge Univ. Press, Cambridge, 1997.

[7] Dov Harel and Robert Endre Tarjan. Fast algorithms for finding nearest common ancestors. *SIAM J. Comput.*, 13(2):338–355, 1984.

[8] J. Kärkkäinen and P. Sanders. Simple linear work suffix array construction. In *Proc. 13th International Conference on Automata, Languages and Programming.* Springer, 2003.

[9] Toru Kasai, Gunho Lee, Hiroki Arimura, Setsuo Arikawa, and Kunsoo Park. Linear-time longest-common-prefix computation in suffix arrays and its applications. In *CPM '01: Proceedings of the 12th Annual Symposium on Combinatorial Pattern Matching*, pages 181–192, London, UK, 2001. Springer-Verlag.

[10] G.M. Landau and U. Vishkin. Efficient string matching in the presence of errors. In *Proc. 26th IEEE Symposium on Foundations of Computer Science*, pages 126–136, Los Alamitos CA, USA, 1985. IEEE Computer Society.

[11] Udi Manber and Gene Myers. Suffix arrays: a new method for on-line string searches. *SIAM J. Comput.*, 22(5):935–948, 1993.

[12] Edward M. McCreight. A space-economical suffix tree construction algorithm. *J. ACM*, 23(2):262–272, 1976.

[13] Baruch Schieber and Uzi Vishkin. On finding lowest common ancestors: simplification and parallelization. *SIAM J. Comput.*, 17(6):1253–1262, 1988.

[14] Esko Ukkonen. On-line construction of suffix trees. *Algorithmica*, 14(3):249–260, 1995.

[15] Peter Weiner. Linear pattern matching algorithms. In *FOCS*, pages 1–11. IEEE, 1973.