References

[1] I. J. Aalbersberg and G. Rozenberg. Theory of traces. *Theor. Comput. Sci.*, 60(1):1–82, 1988.

[2] I. J. Aalbersberg and E. Welzl. Trace languages defined by regular string languages. *RAIRO Inform. Théor.*, 20(2):103–119, 1986.

[3] S. Abdeddaïm. *Comparaison de séquences biologiques.* Thèse de doctorat, Université Paris 6, 1996.

[4] S. Abdeddaïm. Comparaison de séquences biologiques. Rapport Th96/09, LITP, Universités Paris 6 et 7, 1996.

[5] S. Abdeddaïm. On incremental computation of transitive closure and greedy alignment. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 167–179, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[6] S. Abdeddaïm. Algorithms and experiments on transitive closure, path cover, and multiple sequence alignment. In *Proceedings of the 2nd Algorithm Engineering and Experiments*, pages 157–169, San Francisco, CA, 2000.

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

[8] S. De Agostino. A work-optimal parallel implementation of lossless image compression by block matching. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 1–8, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[9] A. V. Aho. Pattern matching in strings. In R. Book, editor, *Formal Language Theory: Perspectives and Open Problems*, pages 325–347. Academic Press, London, U. K., 1980.

[10] A. V. Aho. Algorithms for finding patterns in strings. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science, Algorithms and complexity*, volume A, chapter 5, pages 255–300. Elsevier, Amsterdam, The Netherlands, 1990.

[11] A. V. Aho and M. J. Corasick. Efficient string matching: an aid to bibliographic search. *Commun. ACM*, 18(6):333–340, 1975.

[12] A. V. Aho and M. Ganapathy. Efficient tree pattern matching: an aid to code generation. In *Proceedings of the 12th Annual ACM Symposium on Principles of Programming Languages*, pages 334–340, New Orleans, LO, 1985.

[13] A. V. Aho, D. S. Hirschberg, and J. D. Ullman. Bounds on the complexity of the longest common subsequence problem. *J. Assoc. Comput. Mach.*, 23(1):1–12, 1976.

[14] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The design and analysis of computer algorithms.* Addison-Wesley, Reading, MA, 1974.

[15] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. On finding lowest common ancestors in trees. *SIAM J. Comput.*, 5(1):115–132, 1976.

[16] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *Data structures and algorithms.* Addison-Wesley, Reading, MA, 1983.

[17] A. V. Aho, B. W. Kernighan, and P. J. Weinberger. *The AWK programming language.* Addison-Wesley, Reading, MA, 1988.

[18] A. V. Aho and T. G. Paterson. A minimum distance error-correcting parser for context-free languages. *SIAM J. Comput.*, 1(4):305–312, 1972.

[19] A. V. Aho, R. Sethi, and J. D. Ullman. *Compilers – Principles, Techniques and Tools.* Addison-Wesley, Reading, MA, 1986.

[20] A. V. Aho and J. D. Ullman. *Principles of compiler design.* Addison-Wesley, Reading, MA, 1977.

[21] T. Akutsu. A linear time pattern matching algorithm between a string and a tree. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 1–10, Padova, Italy, 1993. Springer-Verlag, Berlin.

[22] T. Akutsu. Approximate string matching with don't care symbols. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 240–249, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[23] T. Akutsu. Approximate string matching with don't care characters. *Inf. Process. Lett.*, 55(5):235–239, 1995.

[24] T. Akutsu, K. Kanaya, A. Ohyama, and A. Fujiyama. Matching of spots in 2D electrophoresis images. point matching under non-uniform distortions. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 212–222, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[25] C. N. Alberga. String similarity and misspellings. *Commun. ACM*, 10(5):302–313, 1967.

[26] K. S. Alexander. Shortest common superstring for strings of random letters. In M. Crochemore and D. Gusfield, editors, *5th Annual Symposium on Combinatorial Pattern Matching*, volume 807 of *Lecture Notes in Computer Science*, pages 164–172, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[27] F. Alizadeh, R. M. Karp, L. A. Newberg, and D. K. Weisser. Physical mapping of chromosomes: A combinatorial problem in molecular biology. *Algorithmica*, 13(1/2):52–76, 1995.

[28] Julien Allali, Pavlos Antoniou, Costas S. Iliopoulos, Pascal Ferraro, and Manal Mohamed. Validation and decomposition of partially occluded images with holes. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 129–136, Czech Technical University in Prague, Czech Republic, 2009.

[29] C. Allauzen, M. Crochemore, and M. Raffinot. Factor oracle: a new structure for pattern matching. In J. Pavelka, G. Tel, and M. Bartosek, editors, *SOFSEM'99, Theory and Practice of Informatics*, number 1725 in Lecture Notes in Computer Science, pages 291–306, Milovy, Czech Republic, 1999. Springer-Verlag, Berlin.

[30] C. Allauzen and M. Raffinot. Simple optimal string matching algorithm. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 364–374, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[31] C. Allauzen and M. Raffinot. Simple optimal string matching algorithm. *J. Algorithms*, 36(1):102–116, 2000.

[32] L. Allison. A fast algorithm for the alignment of three strings. *J. Theor. Biol.*, 164:261–269, 1993.

[33] L. Allison. Using Hirschberg's algorithm to generate random alignments of strings. *Inf. Process. Lett.*, 51(5):251–255, 1994.

[34] L. Allison. Information-theoretic sequence alignment. Technical Report 98/14, School of Computer_Science and Software_Engineering, Monash University, 1998.

[35] L. Allison and T. I. Dix. A bit string longest common subsequence algorithm. *Inf. Process. Lett.*, 23(6):305–310, 1986.

[36] L. Allison, T. Edgoose, and T. I. Dix. Compression of strings with approximate repeats. In *Proceedings International Conference on Intelligent Systems for Molecular Biology*, pages 8–16, 1998.

[37] L. Allison, D. Powell, and T. I. Dix. Compression and approximate matching. *Comput. J.*, 42(1):1–10, 1999.

[38] L. Allison and C. S. Wallace. An information measure for the string to string correction problem with applications. *Aust. Comput. Sci. Commun.*, 16(1):659–668, 1994.

[39] L. Allison, C. S. Wallace, and C. N. Yee. Induction inference over macro–molecules. Report 90/148, Monash University, 1990.

[40] L. Allison, C. S. Wallace, and C. N. Yee. Finite-state models in the alignment of macromolecules. *J. Mol. Evol.*, 35:77–89, 1992.

[41] L. Allison and C. N. Yee. Minimum message length encoding and the comparison of macromolecules. *Bull. Math. Biol.*, 52:431–453, 1990.

[42] A. Alstrup, G. S. Brodal, and T. Rauhe. Pattern matching in dynamic texts. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 819–828, San Francisco, CA, 2000.

[43] H. Alt, K. Mehlhorn, and J. I. Munro. Partial match retrieval in implicit data structures. *Inf. Process. Lett.*, 19(2):61–65, 1984.

[44] S. F. Altschul. Gap costs for multiple sequence alignment. *J. Theor. Biol.*, 138:297–309, 1989.

[45] S. F. Altschul. Amino acid substitution matrices from an information theoretic perspective. *J. Mol. Biol.*, 219:555–565, 1991.

[46] S. F. Altschul. A protein alignment scoring system sensitive to all evolutionary distances. *J. Mol. Evol.*, 36:290–300, 1993.

[47] S. F. Altschul, R. J. Carroll, and D. J. Lipman. Weights for data related by a tree. *J. Mol. Biol.*, 207:647–653, 1989.

[48] S. F. Altschul and B. W. Erickson. A nonlinear measure of subalignment similarity and its signifiance level. *Bull. Math. Biol.*, 48(5):617–632, 1986.

[49] S. F. Altschul and B. W. Erickson. Optimal sequence alignment using affine gap costs. *Bull. Math. Biol.*, 48(5):603–616, 1986.

[50] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. A basic local alignment search tool. *J. Mol. Biol.*, 215:403–410, 1990.

[51] S. F. Altschul and D. J. Lipman. Trees, stars, and multiple biological sequence alignment. *SIAM J. Appl. Math.*, 49:197–209, 1989.

[52] S. F. Altschul and D. J. Lipman. Protein database searches for multiple alignments. *Proc. Natl. Acad. Sci. U.S.A.*, 87:5509–5513, 1990.

[53] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. Gapped Blast and Psi-Blast: a new generation of protein database search programs. *Nucleic Acids Res.*, 25:3389–3402, 1997.

[54] A. Amir. Multidimensional and relative approximate pattern matching. Reports UMIACS-TR-87-41 and CS-TR-1912, Department of Computer Science and Institute for Advance Computer Studies, University of Maryland, 1987.

[55] A. Amir. Multidimensional pattern matching. Manuscript, 1991.

[56] A. Amir, A. Apostolico, and M. Lewenstein. Inverse pattern matching. *J. Algorithms*, 24(2):325–339, 1997.

[57] A. Amir, Y. Aumann, Landau G, M. Lewenstein, and N. Lewenstein. Pattern matching with swaps. In *Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science*, pages 144–153, Miami Beach, FL, 1997.

[58] A. Amir, Y. Aumann, Landau G, M. Lewenstein, and N. Lewenstein. Pattern matching with swaps. *J. Algorithms*, 37(2):247–266, 2000.

[59] A. Amir and G. Benson. Two-dimensional periodicity and its applications. In *Proceedings of the 3rd ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 440–452, Orlando, FL, 1992.

[60] A. Amir and G. Benson. Two-dimensional periodicity in rectangular arrays. *SIAM J. Comput.*, 27(1):90–106, 1998.

[61] A. Amir, G. Benson, and M. Farach. Alphabet independent two dimensional matching. In *Proceedings of the 24th ACM Symposium on the Theory of Computing*, pages 59–68, Victoria, Canada, 1992. ACM Press.

[62] A. Amir, G. Benson, and M. Farach. The truth, the whole truth and nothing but the truth: alphabet independent 2-d witness table construction. Report GIT-CC-92-51, Georgia Tech, 1992.

[63] A. Amir, G. Benson, and M. Farach. Let sleeping files lie: Pattern matching in Z-compressed files. Report GIT-CC-93-42, Georgia Tech, 1993.

[64] A. Amir, G. Benson, and M. Farach. Optimal parallel two-dimensional pattern matching. In *Proceedings of the 5th Annual ACM Symposium on Parallel Algorithms and Architectures*, pages 79–85. ACM Press, 1993.

[65] A. Amir, G. Benson, and M. Farach. An alphabet independent approach to two-dimensional pattern matching. *SIAM J. Comput.*, 23(2):313–323, 1994.

[66] A. Amir, G. Benson, and M. Farach. Let sleeping files lie: Pattern matching in Z-compressed files. In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 705–714, Arlington, VA, 1994.

[67] A. Amir, G. Benson, and M. Farach. Optimal two-dimensional compressed matching. In S. Abiteboul and E. Shamir, editors, *Proceedings of the 21st International Colloquium on Automata, Languages and Programming*, volume 820 of *Lecture Notes in Computer Science*, pages 215–226, Jerusalem, Israel, 1994.

[68] A. Amir, G. Benson, and M. Farach. Let sleeping files lie: Pattern matching in Z-compressed files. *J. Comput. Syst. Sci.*, 52(2):299–307, 1996.

[69] A. Amir, G. Benson, and M. Farach. Optimal two-dimensional compressed matching. *J. Algorithms*, 24(2):354–379, 1997.

[70] A. Amir, G. Benson, and M. Farach. Optimal parallel two dimensional text searching on a CREW PRAM. *Inf. Comput.*, 144(1):1–17, 1998.

[71] A. Amir, A. Butman, and M. Lewenstein. Real scaled matching. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 815–816, San Francisco, CA, 2000.

[72] A. Amir and G. Călinescu. Faster scaled matching. Report GIT-CC-93-40, Georgia Tech, 1993.

[73] A. Amir and G. Călinescu. Alphabet independent and dictionary scaled matching. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 320–334, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[74] A. Amir and G. Călinescu. Alphabet independent and dictionary scaled matching. *J. Algorithms*, 36(1):34–62, 2000.

[75] A. Amir and M. Farach. Adaptative dictionary matching. In *Proceedings of the 32nd IEEE Annual Symposium on Foundations of Computer Science*, pages 760–766, San Juan, Puerto Rico, 1991. IEEE Computer Society Press.

[76] A. Amir and M. Farach. Efficient 2-dimensional approximate matching of non-rectangular figures. In *Proceedings of the 2nd ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 212–223, San Francisco, CA, 1991.

[77] A. Amir and M. Farach. Two-dimensional dictionary matching. *Inf. Process. Lett.*, 44(5):233–239, 1992.

[78] A. Amir and M. Farach. Efficient 2-dimensional approximate matching of half-rectangular figures. *Inf. Comput.*, 118(1):1–11, 1995.

[79] A. Amir and M. Farach. Two dimensional matching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 9, pages 267–292. Oxford University Press, 1997.

[80] A. Amir, M. Farach, Z. Galil, R. Giancarlo, and K. Park. Fully dynamic dictionary matching. *J. Comput. Syst. Sci.*, 49(2):208–222, 1994.

[81] A. Amir, M. Farach, R. Idury, J. La Poutré, and A. Schäffer. Improved dynamic dictionary matching. In *Proceedings of the 4th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 392–401, San Francisco, CA, 1993.

[82] A. Amir, M. Farach, R. Idury, J. La Poutré, and A. Schäffer. Improved dynamic dictionary matching. *Inf. Comput.*, 119(2):258–282, 1995.

[83] A. Amir, M. Farach, and Y. Matias. Efficient randomized dictionary matching algorithms. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 262–275, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[84] A. Amir, M. Farach, and S. Muthukrishnan. Alphabet dependence in parameterized matching. *Inf. Process. Lett.*, 49(3):111–115, 1994.

[85] A. Amir, M. Farach, and Y. Muthukrishnan. Alphabet dependence in parametrized matching. Report GIT-CC-93-44, Georgia Tech, 1993.

[86] A. Amir, D. Keselman, G. M. Landau, M. Lewenstein, N. Lewenstein, and M. Rodeh. Text indexing and dictionary matching with one error. *J. Algorithms*, 37(2):309–325, 2000.

[87] A. Amir, G. Landau, M. Lewenstein, and N. Lewenstein. Efficient special cases of pattern matching with swaps. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 209–220, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[88] A. Amir, G. Landau, M. Lewenstein, and N. Lewenstein. Efficient special cases of pattern matching with swaps. *Inf. Process. Lett.*, 68(3):125–132, 1998.

[89] A. Amir and G. M. Landau. Fast parallel and serial multidimensional approximate array matching. *Theor. Comput. Sci.*, 81(1):97–115, 1991.

[90] A. Amir, G. M. Landau, and D. Sokol. Inplace run-length $2d$ compressed search. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 817–818, San Francisco, CA, 2000.

[91] A. Amir, G. M. Landau, and U. Vishkin. Efficient pattern matching with scaling. In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 344–457, San Francisco, CA, 1990.

[92] A. Amir, G. M. Landau, and U. Vishkin. Efficient pattern matching with scaling. *J. Algorithms*, 13(1):2–32, 1992.

[93] A. Amir, M. Lewenstein, and N. Lewenstein. Pattern matching in hypertext. In F. K. H. A. Dehne, A. Rau-Chaplin, J.-R. Sack, and R. Tamassia, editors, *Proceedings of the 5th Workshop on Algorithms and Data Structures*, number 1272 in Lecture Notes in Computer Science, pages 160–173, Halifax, Nova Scotia, Canada, 1997. Springer-Verlag, Berlin.

[94] A. Amir, M. Lewenstein, and N. Lewenstein. Pattern matching in hypertext. *J. Algorithms*, 35(1):82–99, 2000.

[95] A. Amir, M. Lewenstein, and E. Porat. Faster algorithms for string matching with k mismatches. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 794–803, San Francisco, CA, 2000.

[96] A. Anderson, T. Hagerup, J. Hastad, and O. Petersson. The complexity of searching a sorted array of strings. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 317–325, Montréal, Canada, 1994. ACM Press.

[97] A. Andersson, N. J. Larsson, and K. Swanson. Suffix trees on words. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 102–115, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[98] A. Andersson, N. J. Larsson, and K. Swanson. Suffix trees on words. *Algorithmica*, 23(3):246–260, 1999.

[99] A. Andersson and S. Nilsson. Improved behavior of tries by adaptive branching. *Inf. Process. Lett.*, 46(6):295–300, 1993.

[100] A. Andersson and S. Nilsson. Faster searching in tries and quadtrees—an analysis of level compression. In J. van Leeuwen, editor, *Proceedings of the 2nd Annual European Symposium*, volume 855 of *Lecture Notes in Computer Science*, pages 82–93, Utrecht, The Netherlands, 1994. Springer-Verlag, Berlin.

[101] K. Ando, K. Iwasaki, M. Fuketa, and J.-I. Aoe. A fast morphological analysis using the extended AC machine for oriental languages. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 1–13, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[102] K. Ando, M. Shishibori, and J.-I. Aoe. An efficient multi-attribute pattern matching machine. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 1–14, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[103] M. V. A. Andrade. Métodos eficientes para reconhecimento de padrões em texto (in portuguese). Master thesis, UNICAMP, Campinas, Brazil, 1993.

[104] G. Andrejková. Systolic systems for the longest common subsequence problem. *Computers and Artificial Intelligence*, 5(3):199–212, 1986.

[105] G. Andrejková. The longest restricted common subsequence problem. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 14–25, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[106] G. Andrejková. The closest common subsequence problems. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '99*, pages 1–15, Czech Technical University, Prague, Czech Republic, 1999. Collaborative Report DC–99–05.

[107] G. Andrejková. The set-set closest common subsequence problem. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 1–12, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[108] R. C. Angell, G. E. Freund, and P. Willet. Automatic spelling correction using a trigram measure. *Inf. Process. Manage.*, 19:255–261, 1993.

[109] D. Angluin. Finding patterns common to a set of strings. In *Proceedings of the 11th ACM Symposium on the Theory of Computing*, pages 130–141, Atlanta, Georgia, 1979. ACM Press.

[110] D. Angluin. Finding patterns common to a set of strings. *J. Comput. Syst. Sci.*, 21(1):46–62, 1980.

[111] E. L. Anson and E. W. Myers. ReAligner: A program for refining DNA sequence multi-alignments. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 9–16, Santa Fe, NM, 1997. ACM Press.

[112] Pavlos Antoniou, Maxime Crochemore, Costas S. Iliopoulos, Inuka Jayasekera, and Gad M. Landau. Conservative string covering of indeterminate strings. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 108–115, Czech Technical University in Prague, Czech Republic, 2008.

[113] J.-I. Aoe. An efficient digital search algorithm by using a double-array structure. *IEEE Trans. Softw. Eng.*, 15(9):1066–1077, 1989.

[114] J.-I. Aoe. An efficient implementation of static string pattern matching machines. *IEEE Trans. Softw. Eng.*, 15(8):1010–1016, 1989.

[115] J.-I. Aoe. An efficient implementation of string pattern matching machines for a finite number of keywords. *SIGIR Forum*, pages 22–33, 1989.

[116] J.-I. Aoe, editor. *String pattern matching strategies*. IEEE Computer Society Press, 1994.

[117] J.-I Aoe, K. Morimoto, and T. Sato. An efficient implementation of trie structure. *Softw. Pract. Exp.*, 22(9):695–721, 1992.

[118] J.-I. Aoe, Y. Yamamoto, and R. Shimada. A method for improving string pattern matching machines. *IEEE Trans. Softw. Eng.*, 10(1):116–120, 1984.

[119] J.-I. Aoe, Y. Yamamoto, and R. Shimada. An efficient implementation of static string pattern matching machines. In *Proceedings of the 1st International Conference on Supercomputing*, pages 491–498, St. Petersburg, FL, 1985.

[120] J.-I. Aoe, Y. Yasutome, and T. Sato. An efficient digital search algorithm by using a double-array structure. In *Proceedings of the 12th International Conference on Computer Software and Applications*, pages 472–479, Chicago, IL, 1988.

[121] A. Apostolico. Linear pattern matching and problems of data compression. In *IEEE International Symposium on Information Theory*, 1979.

[122] A. Apostolico. On context-constrained squares and repetitions in a string. *RAIRO Inform. Théor.*, 18(2):147–159, 1984.

[123] A. Apostolico. The myriad virtues of suffix trees. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 85–96. Springer-Verlag, Berlin, 1985.

[124] A. Apostolico. Improving the worst-case performance of the Hunt-Szymanski strategy for the longest common subsequence of two strings. *Inf. Process. Lett.*, 23(2):63–69, 1986.

[125] A. Apostolico. Parallel log–time construction of suffix trees. Report CSD-TR-632, Purdue Universsity, 1986.

[126] A. Apostolico. Remark on the Hsu-Du new algorithm for the longest common subsequence problem. *Inf. Process. Lett.*, 25(4):235–236, 1987.

[127] A. Apostolico. Optimal parallel detection of squares in string. Reports CSD-TR-932 and CSD-TR-1012, Purdue University, 1990.

[128] A. Apostolico. Fast parallel detection of squares in strings. *Algorithmica*, 8(4):285–319, 1992.

[129] A. Apostolico. Efficient CRCW-PRAM algorithms for universal substring searching. *Theor. Comput. Sci.*, 108(2):331–344, 1993.

[130] A. Apostolico. On-line string searching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 3, pages 89–122. Oxford University Press, 1997.

[131] A. Apostolico. String editing and longest common subsequences. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 2 Linear Modeling: Background and Application, chapter 8, pages 361–398. Springer-Verlag, Berlin, 1997.

[132] A. Apostolico. General pattern matching. In Mikhail J. Atallah, editor, *Algorithms and Theory of Computation Handbook*. CRC Press Inc., Boca Raton, FL, 1998.

[133] A. Apostolico and M. J. Atallah. Parallel computations of Levenshtein distances. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 5, pages 143–184. Oxford University Press, 1997.

[134] A. Apostolico, M. J. Atallah, L. L. Larmore, and S. McFaddin. Efficient parallel algorithms for string editing and related problems. *SIAM J. Comput.*, 19(5):968–988, 1990.

[135] A. Apostolico, M. E. Bock, and S. Lonardi. Linear global detectors of redundant and rare substrings. In J. A. Storer and M. Cohn, editors, *Proceedings Data Compression Conference*, pages 168–177, Snowbird, UT, 1999. IEEE Computer Society Press.

[136] A. Apostolico, M. E. Bock, S. Lonardi, and X. Xu. Efficient detection of unusual words. *J. Comput. Bio.*, 7(1/2):71–94, 2000.

[137] A. Apostolico, M. E. Bock, and X. Xu. Annotated statistical indices for sequence analysis. In B. Carpentieri, A. De Santis, U. Vaccaro, and J. Storer, editors, *Sequences*, pages 215–229, Positano, Italy, 1998. IEEE Computer Society Press.

[138] A. Apostolico and D. Breslauer. An optimal $O(\log n)$ time parallel algorithm for detecting all squares in a string. Report CUCS-040-92, Computer Science Department, Columbia, University, New York NY, 1992.

[139] A. Apostolico and D. Breslauer. Optimal parallel algorithms for periods, palindromes and squares. In O. Nurmi and E. Ukkonen, editors, *Proceedings of the 3rd Scandinavian Workshop on Algorithm Theory*, number 621 in Lecture Notes in Computer Science, pages 296–307, Helsinki, Finland, 1992. Springer-Verlag, Berlin.

[140] A. Apostolico and D. Breslauer. An optimal $O(\log \log N)$ time parallel algorithm for detecting all squares in a string. *SIAM J. Comput.*, 25(6):1318–1331, 1996.

[141] A. Apostolico and D. Breslauer. Of periods, quasiperiods, repetitions and covers. In J. Mycielski, G. Rozenberg, and A. Salomaa, editors, *Structures in Logic and Computer Science, A Selection of Essays in honor of A. Ehrenfeucht*, number 1261 in Lecture Notes in Computer Science, pages 236–248. Springer-Verlag, Berlin, 1997.

[142] A. Apostolico, D. Breslauer, and Z. Galil. Optimal parallel algorithms for periods, palindromes and squares. In *Proceedings of the 19th International Colloquium on Automata, Languages and Programming*, pages 296–307. Springer-Verlag, Berlin, Wien, Austria, 1992.

[143] A. Apostolico, D. Breslauer, and Z. Galil. Parallel detection of all palindromes in a string. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 497–506, Caen, France, 1994. Springer-Verlag, Berlin.

[144] A. Apostolico, D. Breslauer, and Z. Galil. Parallel detection of all palindromes in a string. *Theor. Comput. Sci.*, 141(1–2):163–173, 1995.

[145] A. Apostolico and V. E. Brimkov. Fibonacci arrays and their two-dimensional repetitions. *Theor. Comput. Sci.*, 237(1–2):263–273, 2000.

[146] A. Apostolico, S. Browne, and C. Guerra. Fast linear-space computations of longest common subsequences. *Theor. Comput. Sci.*, 92(1):3–17, 1992.

[147] A. Apostolico and M. Crochemore. Optimal canonization of all substrings of a string. Rapport LITP 89.75, Universités Paris 6 et 7, France, 1989.

[148] A. Apostolico and M. Crochemore. Fast parallel Lyndon factorization with applications. Report CSD-TR-931, Purdue University, 1990.

[149] A. Apostolico and M. Crochemore. Optimal canonization of all substrings of a string. *Inf. Comput.*, 95(1):76–95, 1991.

[150] A. Apostolico and M. Crochemore. Fast parallel Lyndon factorization with applications. *Math. Syst. Theory*, 28(2):89–108, 1995.

[151] A. Apostolico and A. Ehrenfeucht. Efficient detection of quasiperiodicities in strings. Report 90.5, The Leonardo Fibonacci Institute, Trento, Italy, 1990.

[152] A. Apostolico and A. Ehrenfeucht. Efficient detection of quasiperiodicities in strings. *Theor. Comput. Sci.*, 119(2):247–265, 1993.

[153] A. Apostolico, M. Farach, and C. S. Iliopoulos. Optimal superprimitivity testing for strings. *Inf. Process. Lett.*, 39(1):17–20, 1991.

[154] A. Apostolico and A. Fraenkel. Robust transmission of unbounded strings using Fibonacci representations. *IEEE Trans. Inf. Theory*, 33(2):238–245, 1987.

[155] A. Apostolico and Z. Galil, editors. *Combinatorial algorithms on words*, volume 12. Springer-Verlag, Berlin, 1985.

[156] A. Apostolico and Z. Galil, editors. *Pattern matching algorithms*. Oxford University Press, 1997.

[157] A. Apostolico and R. Giancarlo. Pattern matching implementation of a fast test for unique decipherability. *Inf. Process. Lett.*, 18(3):155–158, 1984.

[158] A. Apostolico and R. Giancarlo. The Boyer-Moore-Galil string searching strategies revisited. *SIAM J. Comput.*, 15(1):98–105, 1986.

[159] A. Apostolico and R. Giancarlo. Sequence alignment in molecular biology. *J. Comput. Bio.*, 5(2):173–196, 1998.

[160] A. Apostolico and C. Guerra. A fast linear-space algorithm for computing longest common subsequences. In *Proceedings of the 23rd Allerton Conference*, pages ?–?, Monticello, 1985.

[161] A. Apostolico and C. Guerra. The longest common subsequence problem revisited. *Algorithmica*, 2:315–336, 1987.

[162] A. Apostolico, C. Iliopoulos, G. M. Landau, B. Schieber, and U. Vishkin. Parallel construction of a suffix tree with applications. *Algorithmica*, 3:347–365, 1988.

[163] A. Apostolico, G. M. Landau, and S. Skiena. Matching for run-length encoded strings. In *Proceedings Complexity and Compression of SEQUENCES97*, pages 137–146, Positano, Italy, 1997.

[164] A. Apostolico, G. M. Landau, and S. Skiena. Matching for run-length encoded strings. *J. Complexity*, 15:4–16, 1999.

[165] A. Apostolico and S. Lonardi. Compression of biological sequences by greedy off-line textual substitution. Technical Report TR-99-037, Dept. of Computer Sciences, Purdue University, 1999.

[166] A. Apostolico and S. Lonardi. Compression of biological sequences by greedy off-line textual substitution. In J. A. Storer and M. Cohn, editors, *Proceedings Data Compression Conference*, pages 143–152, Snowbird, UT, 2000. IEEE Computer Society Press.

[167] A. Apostolico and S. Lonardi. Off-line compression by greedy textual substitution. *?*, 88(11):1733–1744, 2000.

[168] A. Apostolico and F. P. Preparata. Optimal off-line detection of repetitions in a string. *Theor. Comput. Sci.*, 22(3):297–315, 1983.

[169] A. Apostolico and F. P. Preparata. Data structures and algorithms for the string statistics problem. Report CSD-TR-547, Purdue University, 1985.

[170] A. Apostolico and F. P. Preparata. Structural properties of the string statistics problem. *J. Comput. Syst. Sci.*, 31(3):394–411, 1985.

[171] A. Apostolico and W. Szpankowski. Self-alignment in words and their applications. *J. Algorithms*, 13(3):446–467, 1992.

[172] A. W. Appel and G. J. Jacobson. The world's fastest scrabble program. *Commun. ACM*, 31(5):572–

[173] M.D. Araújo, G. Navarro, and N. Ziviani. Large text searching allowing errors. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 2–20, Valparaiso, Chile, 1997. Carleton University Press.

[174] J.-P. Arcangeli, M. Crochemore, J.-N. Hourcastagnou, and J.-E. Pin. Compression for an effective management of telemetry data. In *Proceedings of the 2nd International Symposium on Ground Data Systems for Space Operations*, pages 823–830. Jet Propulsion Laboratory Publications, 1993.

[175] L. Arge, P. Ferragina, R. Grossi, and J.S. Vitter. On sorting strings in external memory (extended abstract). In *Proceedings of the 29th ACM Symposium on the Theory of Computing*, pages 540–548, El Paso, TX, 1997.

[176] P. Argos. A sensitive procedure to compare amino acid sequences. *J. Mol. Biol.*, 193:385–396, 1987.

[177] P. Argos and M. Vingron. Sensitivity comparison of protein amino acid sequences. *Methods Enzymol.*, 188:352–365, 1990.

[178] P. Argos, M. Vingron, and G. Vogt. Protein sequence comparison: methods and significance. *Protein Engineering*, 4:375–383, 1991.

[179] S. R. Arikati, A. Dessmark, A. Lingas, and M. V. Marathe. Approximation algorithms for maximum two-dimensional pattern matching. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 348–360, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[180] S. R. Arikati, A. Dessmark, A. Lingas, and M. V. Marathe. Approximation algorithms for maximum two-dimensional pattern matching. *Theor. Comput. Sci.*, 255(1–2):51–62, 2001.

[181] S. Arikawa. One-way sequential search systems and their powers. *Bull. Math. Stat.*, 19(3,4):69–83, 1981.

[182] S. Arikawa and T. Shinohara. A run-time efficient realization of aho-corasick pattern matching machines. *New Gener. Comput.*, 2:171–186, 1984.

[183] S. Arikawa, T. Shinohara, S. Shiraishi, and Y. Tamakoshi. An information system for researchers use. *Bull. Inform. Cybern.*, 20(1,2):97–114, 1982.

[184] H. Arimura, T. Shinohara, and S. Otsuki. Polynomial time inference of unions of two tree pattern languages. *IEICE Trans. Inform. Syst. D*, E75:426–434, 1992.

[185] H. Arimura, T. Shinohara, and S. Otsuki. Finding minimal generalizations for unions of pattern languages and its application to inductive inference fom positive data. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 649–660, Caen, France, 1994. Springer-Verlag, Berlin.

[186] T. Arita, M. Shishibori, and J.-I. Aoe. An efficient algorithm for full text retrieval for multiple keywords. *Inf. Sci.*, 104(3–4):345–363, 1998.

[187] C. Armen and C. Stein. Improved length for the shortest superstring problem. In S. G. Akl, F. K. H. A. Dehne, J.-R. Sack, and N. Santoro, editors, *Proceedings of the 4th Workshop on Algorithms and Data Structures*, number 955 in Lecture Notes in Computer Science, pages 494–505, Kingston, Ontario, Canada, 1995. Springer-Verlag, Berlin.

[188] C. Armen and C. Stein. Algorithm for the shortest superstring problem. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 87–101, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[189] C. Armen and C. Stein. Short superstrings and the structure of overlapping strings. *J. Comput. Bio.*, 2(2):307–332, 1996.

[190] D. B. Arnold and M. R. Sleep. Uniform random generation of balanced parenthesis strings. *ACM Trans. Program. Lang. Syst.*, 2:122–128, 1980.

[191] R. Arratia, L. Gordon, and M. S. Waterman. An extreme value theory for sequence matching. *Ann. Stat.*, 14:971–993, 1986.

[192] R. Arratia, L. Gordon, and M. S. Waterman. The Erdös-Rényi law in distribution, for coin tossing and sequence matching. *Ann. Stat.*, 18:539–570, 1990.

[193] R. Arratia and G. Reinert. Poisson process approximation for repeats in one sequence and its application to sequencing by hybridization. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 209–219, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[194] R. Arratia and M. S. Waterman. Critical phenomena in sequence matching. *Ann. Probab.*, 13(4):1236–1249, 1985.

[195] R. Arratia and M. S. Waterman. The Erdös-Rényi strong law for pattern matching with a given proportion of mismatches. *Ann. Probab.*, 17:1152–1169, 1989.

[196] R. Arratia and M. S. Waterman. A phase transition for the score in matching random sequences allowing deletions. *Ann. Appl. Probab.*, 4(1):200–225, 1994.

[197] A. N. Arslan and O. Egecioglu. Efficient algorithms for normalized edit distance. *J. Discret. Algorithms*, 1(1):3–20, 2000.

[198] L. Arvestad. Aligning coding DNA in the presence of frame-shift errors. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 180–190, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[199] I. Ashdown. Parallel pattern matching and fgrep. *Dr. Dobbs J.*, 10(9):46–67, 1985.

[200] M. J. Atallah, Y. Génin, and W. Szpankowski. Pattern matching image compression. In J. A. Storer and M. Cohn, editors, *Proceedings of the 6th Data Compression Conference*, page 421, Snowbird, UT, 1996. IEEE Computer Society Press.

[201] M. J. Atallah, Y. Génin, and W. Szpankowski. Structural pattern analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(7):614–627, 1999.

[202] M. J. Atallah, P. Jacquet, and W. Szpankowski. Pattern matching with mismatches: a probabilistic analysis and a randomized algorithm. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 27–40, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[203] M. J. Atallah, P. Jacquet, and W. Szpankowski. A probabilistic approach to pattern matching with mismatches. *Random Struct. Algorithms*, 4:191–213, 1993.

[204] M. J. Atallah, S. R. Kosaraju, L. L. Larmore, G. L. Miller, and S. H. Teng. Constructing trees in parallel. Report CSD-TR-883, Purdue University, 1989.

[205] T. K. Attwood and D. J. Parry-Smith. *Introduction to bioinformatics.* Addison Wesley Longman, 1999.

[206] L. Audoire, J.-J. Codani, D. Lavenier, and P. Quinton. Machines Spécialisées pour la Comparaison de Séquences Biologiques. *Tech. Sci. Inform.*, 14(1), 1995.

[207] B. Awerbuch, B. Berger, L. Cowen, and D. Peleg. Near-linear time construction of sparse neighborhood covers. *SIAM J. Comput.*, 28(1):263–277, 1998.

[208] S. Baase. *Computer algorithms – Introduction to design and analysis.* Addison-Wesley, Reading, MA, 1988.

[209] K. Baba, A. Shinohara, M. Takeda, S. Inenaga, and S. Arikawa. A note on randomized algorithm for string matching with mismatches. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 9–17, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[210] R. Backofen. An upper bound for number of contacts in the HP-model on the Face-Centered-Cubic Lattice (FCC). In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 277–292, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[211] D. J. Bacon and W. F. Anderson. Multiple sequence comparison. *J. Mol. Biol.*, 191:153–161, 1986.

[212] D. J. Bacon and W. F. Anderson. Multiple sequence comparison. *Methods Enzymol.*, 183:438–447, 1990.

[213] R. Baeza-Yates and G. Navarro. New models and algorithms for multidimensional approximate pattern matching. *J. Discret. Algorithms*, 1(1):21–49, 2000.

[214] R. A. Baeza-Yates. On the average case on string matching algorithms. Report CS-87-66, Computer Science Department, University of Waterloo, 1987.

[215] R. A. Baeza-Yates. The subsequence graph of a text. Report CS-88-34, Computer Science Department, University of Waterloo, 1988.

[216] R. A. Baeza-Yates. Algorithms for string searching: a survey. *SIGIR Forum*, 23(3,4):34–58, 1989.

[217] R. A. Baeza-Yates. *Efficient text searching.* Ph. D. Thesis, Department of Computer Science, University of Waterloo, Ontario, 1989.

[218] R. A. Baeza-Yates. Improved string searching. *Softw. Pract. Exp.*, 19(3):257–271, 1989.

[219] R. A. Baeza-Yates. String searching algorithms revisited. In F. Dehne, J. R. Sack, and N. Santoro, editors, *Proceedings of the 1st Workshop on Algorithms and Data Structures*, number 382 in Lecture Notes in Computer Science, pages 75–96, Ottawa, Canada, 1989. Springer-Verlag, Berlin.

[220] R. A. Baeza-Yates. The subsequence graph of a text. In J. Diaz and F. Orejas, editors, *Proceedings of the 3rd Theory and Practice of Software Development*, number 351 in Lecture Notes in Computer Science, pages 104–118, Barcelona, Spain, 1989. Springer-Verlag, Berlin.

[221] R. A. Baeza-Yates. Searching subsequences. *Theor. Comput. Sci.*, 78(2):363–376, 1991.

[222] R. A. Baeza-Yates. Notes on a very fast substring search algorithm. *Commun. ACM*, 35(4):132, 1992.

[223] R. A. Baeza-Yates. Text retrieval: theory and practice. In J. van Leeuwen, editor, *Proceedings of the 12th IFIP World Computer Congress*, pages 465–476, Madrid, Spain, 1992. North-Holland.

[224] R. A. Baeza-Yates. A unified view of string matching algorithms. In *Proceedings SOFSEM'96: Theory and Practice of Informatics*, number 1175 in Lecture Notes in Computer Science, pages 1–15. Springer-Verlag, Berlin, 1996.

[225] R. A. Baeza-Yates. A unified view to pattern-matching problems. In *Proceedings PANEL'96*, Bogotá, Colombia, 1996.

[226] R. A. Baeza-Yates, E. F. Barbosa, and N. Ziviani. Hierarchies of indices for text searching. *Inf. Syst.*, 21(6):497–514, 1996.

[227] R. A. Baeza-Yates, W. Cunto, U. Manber, and S. Wu. Proximity matching using fixed-queries trees. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 198–212, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[228] R. A. Baeza-Yates, R. Gavaldà, and G. Navarro. Bounding the expected length of longest common subsequences and forests. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 1–15, Recife, Brazil, 1996. Carleton University Press.

[229] R. A. Baeza-Yates, R. Gavaldà, G. Navarro, and R. Scheihing. Bounding the expected length of longest common subsequences and forests. *Theor. Comput. Syst.*, 32(4):435–452, 1999.

[230] R. A. Baeza-Yates and G. H. Gonnet. New algorithm for pattern matching with and without mismatches. Report CS-88-37, Department of Computer Science, University of Waterloo, 1988.

[231] R. A. Baeza-Yates and G. H. Gonnet. Boyer-Moore automata. Report, University of Waterloo, 1989.

[232] R. A. Baeza-Yates and G. H. Gonnet. Efficient text searching of regular expressions. In G. Ausiello, M. Dezani-Ciancaglini, and S. Ronchi Della Rocca, editors, *Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, number 372 in Lecture Notes in Computer Science, pages 46–62, Stresa, Italy, 1989. Springer-Verlag, Berlin.

[233] R. A. Baeza-Yates and G. H. Gonnet. A new approach to text searching. In N. J. Belkin and C. J. van Rijsbergen, editors, *Proceedings of the 12th International Conference on Research and Development in Information Retrieval*, pages 168–175, Cambridge, MA, 1989. ACM Press.

[234] R. A. Baeza-Yates and G. H. Gonnet. All-against-all sequence matching. Report, Department of Computer Science, Universidad de Chile, 1990.

[235] R. A. Baeza-Yates and G. H. Gonnet. A new approach to text searching. *Commun. ACM*, 35(10):74–82, 1992.

[236] R. A. Baeza-Yates and G. H. Gonnet. Fast string matching with mismatches. *Inf. Comput.*, 108(2):187–199, 1994.

[237] R. A. Baeza-Yates and G. H. Gonnet. Fast text searching for regular expressions or automaton searching on tries. *J. Assoc. Comput. Mach.*, 43(6):915–936, 1996.

[238] R. A. Baeza-Yates, G. H. Gonnet, and M. Régnier. Analysis of Boyer-Moore type string searching algorithms. In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 328–343, San Francisco, CA, 1990.

[239] R. A. Baeza-Yates and G. Navarro. A fast heuristic for approximate string matching. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 47–63, Recife, Brazil, 1996. Carleton University Press.

[240] R. A. Baeza-Yates and G. Navarro. A faster algorithm for approximate string matching. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 1–23, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[241] R. A. Baeza-Yates and G. Navarro. Block-addressing indices for approximate text retrieval. In *Proceedings ACM CIKM'97*, pages 1–8, 1997.

[242] R. A. Baeza-Yates and G. Navarro. Multiple approximate string matching. In F. K. H. A. Dehne, A. Rau-Chaplin, J.-R. Sack, and R. Tamassia, editors, *Proceedings of the 5th Workshop on Algorithms and Data Structures*, number 1272 in Lecture Notes in Computer Science, pages 174–184, Halifax, Nova Scotia, Canada, 1997. Springer-Verlag, Berlin.

[243] R. A. Baeza-Yates and G. Navarro. Fast two-dimensional approximate pattern matching. In C. L. Lucchesi and A. V. Moura, editors, *Proceedings of the 3rd Latin American Symposium*, number 1380 in Lecture Notes in Computer Science, pages 341–351, Campinas, Brazil, 1998. Springer-Verlag, Berlin.

[244] R. A. Baeza-Yates and G. Navarro. New and faster filters for multiple approximate string matching. Technical Report TR/DCC-98-10, Dept. of Computer Science, University of Chile, 1998.

[245] R. A. Baeza-Yates and G. Navarro. Faster approximate string matching. *Algorithmica*, 23(2):127–158, 1999.

[246] R. A. Baeza-Yates and C. H. Perleberg. Fast and practical approximate string matching. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 185–192, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[247] R. A. Baeza-Yates and C. H. Perleberg. Fast and practical approximate string matching. *Inf. Process. Lett.*, 59(1):21–27, 1996.

[248] R. A. Baeza-Yates and M. Régnier. Fast algorithms for two-dimensional and multiple pattern matching. In R. Karlsson and J. Gilbert, editors, *Proceedings of the 2nd Scandinavian Workshop on Algorithm Theory*, number 447 in Lecture Notes in Computer Science, pages 332–347, Bergen, Norway, 1990. Springer-Verlag, Berlin.

[249] R. A. Baeza-Yates and M. Régnier. Average running time of the Boyer-Moore-Horspool algorithm. *Theor. Comput. Sci.*, 92(1):19–31, 1992.

[250] R. A. Baeza-Yates and M. Régnier. Fast two-dimensional pattern matching. *Inf. Process. Lett.*, 45(1):51–57, 1993.

[251] R. A. Baeza-Yates and B. Ribeiro-Neto, editors. *Modern Information Retrieval*. Addison-Wesley, 1999.

[252] R. A. Baeza-Yates and R. Scheihing. New lower bounds for the expected length of longest common subsequences and forests. In *Proceedings of the 15th International Conference of the Chilean Computer Science Society*, pages 48–58, Arica, Chile, 1995.

[253] R.A. Baeza-Yates, C. Choffrut, and G.H. Gonnet. On Boyer-Moore automata. *Algorithmica*, 12(4/5):268–292, 1994.

[254] R.A. Baeza-Yates and L.O. Fuentes. A framework to animate string algorithms. *Inf. Process. Lett.*, 59(5):241–244, 1996.

[255] V. Bafna, E. L. Lawler, and P. A. Pevzner. Approximation algorithms for multiple sequence aligmnent. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 43–53, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[256] V. Bafna, E. L. Lawler, and P. A. Pevzner. Approximation algorithms for multiple sequence aligmnent. *Theor. Comput. Sci.*, 182(1–2):233–244, 1997.

[257] V. Bafna, S. Muthukrishnan, and R. Ravi. Computing similarity between RNA strings. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 1–16, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[258] V. Bafna and P. Pevzner. Genome rearrangments and sorting by reversal. *SIAM J. Comput.*, 25(2):272–289, 1996.

[259] V. Bafna and P. A. Pevzner. Genome rearrangements and sorting by reversals. In *Proceedings of the 34th Annual Symposium on Foundations of Comptuer Science*, pages 148–157, Palo Alto, CA, 1993. IEEE Computer Society Press.

[260] V. Bafna and P. A. Pevzner. Sorting by transpositions. *SIAM J. Discret. Math.*, 11(2):224–240, 1998.

[261] T. A. Bailey and R. G. Dromey. Fast string searching by finding subkeys in subtext. *Inf. Process. Lett.*, 11(3):130–133, 1980.

[262] W. Bains. MULTAN: a program to align multiple DNA sequences. *Nucleic Acids Res.*, 14:159–177, 1986.

[263] W. Bains. MULTAN (2), a multiple string alignment program for nucleic acids and proteins. *Comput. Appl. Biosci.*, 5:51–52, 1989.

[264] B. S. Baker. A program for identifying duplicated code. In *Proceedings of the 24th Symposium on the Interface: Computer Science and Statistics*, pages 18–21, College Station, TX, 1992. ACM Press.

[265] B. S. Baker. A theory of parameterized pattern matching: algorithms and applications. In *Proceedings of the 25th ACM Symposium on the Theory of Computing*, pages 71–80, San Diego, CA, 1993. ACM Press.

[266] B. S. Baker. Parameterized pattern matching by Boyer-Moore-type algorithms. In *Proceedings of the 6th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 541–550, San Francisco, CA, 1995.

[267] B. S. Baker. Parameterized pattern matching: algorithms and applications. *J. Comput. Syst. Sci.*, 52(1):28–42, 1996.

[268] B. S. Baker. Parameterized duplication in strings: algorithms and an application to software maintenance. *SIAM J. Comput.*, 26(5):1343–1362, 1997.

[269] B. S. Baker and R. Giancarlo. Longest common subsequence from fragments via sparse dynamic programming. In G. Bilardi, G. F. Italiano, A. Pietracaprina, and G. Pu cci, editors, *Proceedings of the 6th Annual European Symposium*, number 1461 in Lecture Notes in Computer Science, pages 79–90, Venice, Italy, 1998. Springer-Verlag, Berlin.

[270] K. A. Baker, G. F. McNulty, and W. Taylor. Growth problems for avoidable words. *Theor. Comput. Sci.*, 69(3):319–345, 1989.

[271] T. P. Baker. A technique for extending rapid exact-match string matching to arrays of more than one dimension. *SIAM J. Comput.*, 7(4):533–541, 1978.

[272] M. Balík. Implementation of DAWG. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 26–35, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[273] M. Balík. Condensation principle. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 13–23, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[274] A. M. Barber and J. V. Maizel, Jr. SequenceEditingAligner: a multiple sequence editor and aligner. *Gene Anal. Tech. Appl.*, 7:39–45, 1990.

[275] Md. Faizul Bari, M. Sohel Rahman, and Rifat Shahriyar. Finding all covers of an indeterminate string in $o(n)$ time on average. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 263–271, Czech Technical University in Prague, Czech Republic, 2009.

[276] M. P. Barnett. Primitive string transformations as reductions to normal form. *SIGPLAN Not.*, 25(5):29–33, 1990.

[277] G. Barth. An alternative for the implementation of Knuth-Morris-Pratt algorithm. *Inf. Process. Lett.*, 13(4/5):134–137, 1981.

[278] G. Barth. An analytical comparison of two string searching algorithms. *Inf. Process. Lett.*, 18(5):249–256, 1984.

[279] G. Barth. Relating the average-case cost of the brute-force and the Knuth-Morris-Pratt string matching algorithm. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 45–58. Springer-Verlag, Berlin, 1985.

[280] G. J. Barton. Protein multiple sequence alignment and flexible pattern matching. *Methods Enzymol.*, 188:403–428, 1990.

[281] G. J. Barton. An efficient algorithm to locate all locally optimal alignments between two sequences allowing for gaps. *Comput. Appl. Biosci.*, 9(6):729–734, 1993.

[282] G. J. Barton and M. J. E. Sternberg. A strategy for the rapid multiple alignment of protein sequences. *J. Mol. Biol.*, 198:327–337, 1987.

[283] Paweł Baturo, Marcin Piatkowski, and Wojciech Rytter. Usefulness of directed acyclic subword graphs in problems related to standard Sturmian words. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 193–207, Czech Technical University in Prague, Czech Republic, 2008.

[284] S. Batzoglou and S. Istrail. Physical mapping with repeated probes: The hypergraph superstring problem. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 66–77, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[285] S. Batzoglou and S. Istrail. Physical mapping with repeated probes: The hypergraph superstring problem. *J. Discret. Algorithms*, 1(1):51–76, 2000.

[286] D. R. Bean, A. Ehrenfeucht, and G. F. McNulty. Avoidable patterns in strings of symbols. *Pac. J. Math.*, 85:261–294, 1979.

[287] D. Beauquier, J. Berstel, and P. Chrétienne. *Éléments d'algorithmique.* Masson, Paris, France, 1992.

[288] D. Beauquier and J.-E. Pin. Factors of words. In G. Ausiello, M. Dezani-Ciancaglini, and S. Ronchi Della Rocca, editors, *Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, number 372 in Lecture Notes in Computer Science, pages 190–200, Stresa, Italy, 1989. Springer-Verlag, Berlin.

[289] G. Bejerano and G. Yona. Modeling protein families using probabilistic suffix trees. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 3rd Annual International Conference on Computational*

Molecular Biology (RECOMB), pages 15–24, Lyon, France, 1999. ACM Press.

[290] G. I. Bell, R. L. Bivins, J. D. Louck, N. Metropolis, and M. L. Stein. Properties of words on four letters from those on two letters with an application to DNA sequences. *Adv. Appl. Math.*, 14(3):348–366, 1993.

[291] T. Bell and D. Kulp. Longest-match string searching for Ziv-Lempel compression. *Softw. Pract. Exp.*, 23(7):757–771, 1993.

[292] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[293] A. Ben-Amram, O. Berkman, C. S. Iliopoulos, and K. Park. The subtree max gap problem with application to parallel string covering. In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 501–510, Arlington, VA, 1994.

[294] A. Ben-Dor, G. Lancia, J. Perone, and R. Ravi. Banishing bias from consensus sequences. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 247–261, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[295] A. Ben-Dor, I. Pe'er, R. Shamir, and R. Sharan. On the complexity of positional sequencing by hybridization. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 88–100, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[296] S. Ben-Yehuda and R. Y. Pinter. Symbolic layout improvement using string matching based local transformation. In *Proceedings of the Decennial Caltech Conference on VLSI*, pages 227–239, 1989.

[297] C. Benham, S. Kannan, and T. Warnow. Of chicken teeth and mouse eyes, or generalized character compatibility. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 17–26, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[298] G. Benson. A space efficient algorithm for finding the best non-overlapping alignment score. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 1–14, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[299] G. Benson. Sequence alignment with tandem duplication. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 27–36, Santa Fe, NM, 1997. ACM Press.

[300] G. Benson. An algorithm for finding tandem repeats of unspecified pattern size. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 20–29, New York, NY, 1998. ACM Press.

[301] G. Benson and M. S. Waterman. A method for fast database search for all k-nucleotide repeats. *Nucleic Acids Res.*, 22:4828–4836, 1998.

[302] G. E. Benson. *Two-dimensional periodicity and matching algorithms*. Ph. D. Thesis, Department of Computer Science, University of Maryland, 1992.

[303] K. Bent. A recursive function for calculating the number of legal strings of parentheses and for calculating catalan numbers. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 18–27, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[304] J. Bentley and D. McIlroy. Data compression using long common strings. In *Proceedings Data Compression Conference*, pages 287–295, 1999.

[305] J. L. Bentley and R. Sedgewick. Fast algorithms for sorting and searching strings. In *Proceedings of the 8th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 360–369, New Orleans, LO, 1997.

[306] L. Bergroth, H. Hakonen, and T. Raita. New approximation algorithms for longest common subsequences. In *Proceedings String Processing and Information Retrieval: A South American Symposium*, pages 32–40, Santa Cruz de la Sierra BoliviaSanta Cruz de la Sierra, Bolivia, 1998. IEEE Computer Society Press.

[307] O. Berkman, D. Breslauer, Z. Galil, B. Schieber, and U. Vishkin. Highly parallelizable problems. In *Proceedings of the 21st ACM Symposium on the Theory of Computing*, pages 309–319, New York, NY, 1989. ACM Press.

[308] O. Berkman, C.S. Iliopoulos, and K. Park. The subtree max gap problem with application to parallel string covering. *Inf. Comput.*, 123(1):127–137, 1995.

[309] S. Y. Berkovich and A. E. F. A. Hegazy. Matching string patterns in large textual files. In *International Symposium on New Directions in Computing*, pages 122–127, Trondheim, Norway, 1985. IEEE Computer Society Press.

[310] P. Berman and S. Hannenhalli. Fast sorting by reversal. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 168–185, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[311] P. Berman, M. Karpinski, L. L. Larmore, W. Plandowski, and W. Rytter. On the complexity of pattern matching for highly compressed two-dimensional texts. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 40–51, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[312] P. Berman, Z. Zhang, Y. I. Wolf, E. V. Koonin, and W. Miller. Winnowing sequences from a database search. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 3rd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 50–58, Lyon, France, 1999. ACM Press.

[313] G. Berry and R. Sethi. From regular expression to deterministic automata. *Theor. Comput. Sci.*, 48(1):117–126, 1986.

[314] T. Berry and S. Ravindran. A fast string matching algorithm and experimental results. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '99*, pages 16–28, Czech Technical University, Prague, Czech Republic, 1999. Collaborative Report DC–99–05.

[315] T. Berry and S. Ravindran. String matching in a compressed DNA text. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 53–62, Perth, WA, Australia, 1999.

[316] J. Berstel. Mots sans carrés et morphismes itérés. Rapport LITP 78-42, Universités Paris 6 et 7, France, 1978.

[317] J. Berstel. Some recent results on squarefree words. Rapport LITP 84-2, Universités Paris 6 et 7, France, 1984.

[318] J. Berstel and L. Boasson. Partial words and a theorem of Fine and Wilf. *Theor. Comput. Sci.*, 218(1):135–141, 1999.

[319] J. Berstel and D. Perrin. *Theory of codes.* Academic Press, Orlando, FL, 1985.

[320] J. Berstel, D. Perrin, J.-F. Perrot, and A. Restivo. Sur le théorème du défaut. Rapport LITP 78-17, Universités Paris 6 et 7, France, 1978.

[321] J. Berstel, D. Perrin, J.-F. Perrot, and A. Restivo. Sur le théorème du défaut. *J. Algorithms*, 60(1):169–181, 1979.

[322] J. Berstel and M. Pocchiola. Average cost of Duval's algorithm for generating Lyndon words. Rapport 92-8, LIENS, 1992.

[323] J. Berstel and M. Pocchiola. A geometric proof of the enumeration formula for Sturmian words. *Int. J. Algebr. Comput.*, 3:349–355, 1993.

[324] J. Berstel and M. Pocchiola. Average cost of Duval's algorithm for generating Lyndon words. *Theor. Comput. Sci.*, 132(1-2):415–425, 1994.

[325] J. Berstel and P. Séébold. A characterization of Sturmian morphisms. In A. Borzyskovki and S. Sokolowski, editors, *Proceedings of the 18th Symposium on Mathematical Foundations of Computer Science*, number 711 in Lecture Notes in Computer Science, pages 281–290, Gdansk, Poland, 1993. Springer-Verlag, Berlin.

[326] J. Berstel and P. Séébold. A remark on morphic Sturmian word. *RAIRO Inform. Théor. Appl.*, 28(3-4):255–263, 1994.

[327] A. A. Bertossi. A VLSI system for string matching. *Integr., The VLSI J.*, 9:129–139, 1990.

[328] A. A. Bertossi, E. Lodi, F. Luccio, and L. Pagli. String matching with context dependent errors. Report, Dipartimento di Informatica, Universitá di Pisa, 1987.

[329] A. A. Bertossi, E. Lodi, F. Luccio, and L. Pagli. Context–dependent string matching. In R. M. Capocelli, editor, *Sequences, combinatorics, compression, security and transmission*, pages 25–40. Springer-Verlag, Berlin, 1990.

[330] A. A. Bertossi, F. Luccio, E. Lodi, and L. Pagli. String matching with weighted errors. *Theor. Comput. Sci.*, 73(3):319–328, 1990.

[331] A. A. Bertossi, F. Luccio, L. Pagli, and E. Lodi. A parallel solution to the approximate string matching algorithm problem. *Comput. J.*, 35(5):524–526, 1992.

[332] M. A. Bickel. Automatic correction to misspelled names: a fourth-generation language approach. *Commun. ACM*, 30(3):224–228, 1987.

[333] P. Bieganski. *Genetic sequence data retrieval and manipulation based on generalized suffix trees.* Ph. D. Thesis, University of Minnesota, 1995.

[334] A. Binstock and J. Rex. *Practical Algorithms for Programmers.* Addison-Wesley, Reading, MA, 1995.

[335] R. M. Bird, J. C. Tu, and R. M. Worthy. Associative/parallel processors for searching very large textual data bases. In M. J. McGill, editor, *Proceedings of the 3rd Workshop on Computer Architecture for Non-numeric Processing*, pages 8–16, ACM, New York (published as aspecial issue of SIGIR Forum 12(1)), 1977.

[336] R. S. Bird. Two-dimensional pattern matching. *Inf. Process. Lett.*, 6(5):168–170, 1977.

[337] R. S. Bird, J. Gibbons, and G. Jones. Formal derivation of a pattern matching algorithm. *Sci. Comp. Program.*, 12(2):93–104, 1989.

[338] M. J. Bishop, A. E. Friday, and E. A. Thompson. Inference of evolutionary relationships. In *Nucleic acid and protein sequence analysis: a practical approach*, pages 359–385. IRL Press Limited, Oxford, U. K., 1987.

[339] M. J. Bishop and C. J. Rawlings. *Nucleic acid and protein sequence analysis: a practical approach*. IRL Press Limited, Oxford, U. K., 1987.

[340] M. J. Bishop and E. A. Thompson. Fast computer search for similar DNA sequences. *Nucleic Acids Res.*, 12(13):5471–5474, 1984.

[341] M. J. Bishop and E. A. Thompson. Maximum likelihood alignment of DNA sequences. *J. Mol. Biol.*, 190:159–165, 1986.

[342] C. R. Blair. A program for correcting spelling errors. *Inf. Control*, 3:60–67, 1960.

[343] B. E. Blaisdell. A measure of the similarity of sets of sequences not requiring sequence alignment. *Proc. Natl. Acad. Sci. U.S.A.*, 83:5155–5159, 1986.

[344] M. Bloch. *Experimental evaluation and optimization of Aho-Corasick pattern matching machine*. Ph. D. Thesis, University of Prague, Czech Republic, 1989.

[345] M. Bloch. Efficiency of AC-machine and SFNA in practical string matching. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 15–18, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[346] A. Blum, T. Jiang, M. Li, J. Tromp, and M. Yannakakis. Linear approximation of shortest superstrings. In *Proceedings of the 23rd ACM Symposium on the Theory of Computing*, pages 328–336, New Orleans, LA, 1991. ACM Press.

[347] A. Blum, T. Jiang, M. Li, J. Tromp, and M. Yannakakis. Linear approximation of shortest superstrings. *J. Assoc. Comput. Mach.*, 41(4):630–647, 1994.

[348] N. Blum. On locally optimal alignments in genetic sequences. In A. Finkel and M. Jantzen, editors, *Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, number 577 in Lecture Notes in Computer Science, pages 426–436, Cachan, France, 1992. Springer-Verlag, Berlin.

[349] N. Blum. Speeding up dynamic programming without omitting any optimal solution and some applications in molecular biology. *J. Algorithms*, 35(2):129–168, 2000.

[350] A. Blumer, J. Blumer, A. Ehrenfeucht, D. Haussler, M. T. Chen, and J. Seiferas. The smallest automaton recognizing the subwords of a text. *Theor. Comput. Sci.*, 40(1):31–55, 1985.

[351] A. Blumer, J. Blumer, A. Ehrenfeucht, D. Haussler, and R. McConnel. Linear size finite automata for the set of all subwords of a word: an outline of results. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 21:12–20, 1983.

[352] A. Blumer, J. Blumer, A. Ehrenfeucht, D. Haussler, and R. McConnel. Building a complete inverted file for a set of text files in linear time. In *Proceedings of the 16th ACM Symposium on the Theory of Computing*, pages 349–351, Washington, D.C., 1984. ACM Press.

[353] A. Blumer, J. Blumer, A. Ehrenfeucht, D. Haussler, and R. McConnel. Building the minimal DFA for the set of all subwords of a word on-line in linear time. In J. Paredaens, editor, *Proceedings of the 11th International Colloquium on Automata, Languages and Programming*, number 172 in Lecture Notes in Computer Science, pages 109–118, Antwerp, Belgium, 1984. Springer-Verlag, Berlin.

[354] A. Blumer, J. Blumer, A. Ehrenfeucht, D. Haussler, and R. McConnel. Complete inverted files for efficient text retrieval and analysis. *J. Assoc. Comput. Mach.*, 34(3):578–595, 1987.

[355] A. Blumer, A. Ehrenfeucht, and D. Haussler. Average size of suffix trees and DAWGS. *Discret. Appl. Math.*, 24:37–45, 1989.

[356] J. Blumer and A. Blumer. On-line construction of a complete inverted file. Report UCSC-CRL-86-11, Department of Computer Science, University of Colorado, Boulder, 1986.

[357] J. A. Blumer. How much is that DAWG in the window ? a moving window algorithm for the directed acyclic word graph. *J. Algorithms*, 8(4):451–469, 1987.

[358] L. Boasson, P. Cegielski, I. Guessarian, and Y. Matiyasevich. Window-accumulated subsequence matching problem is linear. In *Proceedings of the 18th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 327–336, Philadelphia, Pennsylvania, 1999. ACM Press.

[359] H. Bodlaender, R. G. Downey, M. R. Fellows, M. T. Hallett, and H. T. Wareham. Parametrized complexity analysis in computational biology. *Comput. Appl. Biosci.*, 11(1):49–57, 1995.

[360] H. Bodlaender, R. G. Downey, M. R. Fellows, M. T. Hallett, and H. T. Wareham. Parametrized complexity analysis in computational biology. *Theor. Comput. Sci.*, 147(1–2):31–54, 1995.

[361] H. Bodlaender, R.G. Downey, M.R. Fellows, and H.T. Wareham. The parametrized complexity of sequence alignment and consensus. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 15–30, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[362] H. L. Bodlaender, M. R. Fellows, and P. A. Evans. Finite-state computability of annotations of strings and trees. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 384–391, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[363] P. Bonizzoni, M. Duella, and G. Mauri. Approximation complexity of longest common subsequence and shortest common supersequence over fixed alphabet. Report 117/94, Università degli Studi Milano, Italy, 1994.

[364] P. Bonizzoni, G. Della Vedova, and G. Mauri. Approximating the maximum isomorphic agreement subtree is hard. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 119–128, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[365] A. Bookstein. On Harrison's substring testing technique. *Commun. ACM*, 16(3):180–181, 1973.

[366] K. S. Booth. Lexicographically least circular substrings. *Inf. Process. Lett.*, 10(4):240–242, 1980.

[367] J. Boothroyd. Stringsort. *Commun. ACM*, 6(10):615, 1963.

[368] D. Botstein. Expression arrays: An open problem for mathematicians in post-genome sequence biology. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 56–56, Santa Fe, NM, 1997. ACM Press.

[369] J. Boyar and K. S. Larsen. Faster parallel computation of edit distances. Technical report, dept. Math. CS, Odense Univ., Denmark, 1993.

[370] R. S. Boyer and J. S. Moore. A fast string searching algorithm. *Commun. ACM*, 20(10):762–772, 1977.

[371] J. H. Bradford. Sequence matching with binary codes. *Inf. Process. Lett.*, 34(4):193–196, 1990.

[372] J. H. Bradford and T. A. Jenkyns. On the inadequacy of tournament algorithms for the N-SCS problem. *Inf. Process. Lett.*, 38(4):169–171, 1991.

[373] J. R. Bradley and S. S. Skiena. Fabricating arrays of strings. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 57–66, Santa Fe, NM, 1997. ACM Press.

[374] A. Brāzma. Learning of regular expressions by pattern matching. In Paul Vitányi, editor, *Proceedings of the 2nd European Conference on Computational Learning Theory*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 392–403. Springer-Verlag, Berlin, 1995.

[375] A. Brāzma. *Efficient learning of regular expressions from approximate examples*, volume IV: Making Learning Systems Practical, chapter 19, pages 337–352. MIT Press, 1997.

[376] A. Brāzma, I. Jonassen, I. Eidhammer, and D. Gilbert. Approaches to the automatic discovery of patterns in biosequences. *J. Comput. Bio.*, 5(2):277–304, 1998.

[377] A. Brāzma, I. Jonassen, E. Ukkonen, and J. Vilo. Discovering patterns and subfamilies in biosequences. In D. J. States, P. Agarwal, T. Gaasterland, L. Hunter, and R. Smith, editors, *Proceedings of the 4th International Conference on Intelligent Systems for Molecular Biology*, pages 34–43. AAAI Press, 1996.

[378] A. Brāzma, I. Jonassen, E. Ukkonen, and J. Vilo. Predicting gene regulatory elements in silico on a genomic scale. *Genome Research*, 8(11):1202–1215, 1998.

[379] A. Brāzma, I. Jonassen, J. Vilo, and E. Ukkonen. Pattern discovery in biosequences. In Vasant Honavar and Giora Slutzki, editors, *Proceedings of the 4th International Colloquium on Grammatical Inference (ICGI-98)*, volume 1433 of *Lecture Notes in Computer Science*, pages 257–270. Springer-Verlag, Berlin, 1998.

[380] A. Brāzma, J. Vilo, E. Ukkonen, and K. Valtonen. Data mining for regulatory elements in yeast genome. In T. Gaasterland, P. Karp, K. Karplus, C. Ouzounis, C. Sander, and A. Valencia, editors, *Proceedings of the 5th International Conference on Intelligent Systems for Molecular Biology*, pages 65–74. AAAI Press, 1997.

[381] E. Breimer, M. Goldberg, and D. Lim. A learning algorithm for the longest common subsequence problem. In *Proceedings of the 2nd Algorithm Engineering and Experiments*, pages 147–156, San Francisco, CA, 2000.

[382] D. Breslauer. An optimal $O(\log \log n)$ time parallel string matching algorithm. Report CUCS-492-89, Computer Science Department, Columbia University, NY, 1989.

[383] D. Breslauer. *Efficient string algorithmics.* Ph. D. Thesis, Computer Science Department, Columbia University, NY, 1992.

[384] D. Breslauer. Fast parallel string prefix-matching. Report CUCS-041-92, Computer Science Department, Columbia University, NY, 1992.

[385] D. Breslauer. An on-line string superprimitivity test. *Inf. Process. Lett.*, 44(6):345–347, 1992.

[386] D. Breslauer. Testing string superprimitivity in parallel. Report CUCS-053-92, Computer Science Department, Columbia University, NY, 1992.

[387] D. Breslauer. Saving comparisons in the Crochemore-Perrin string matching algorithm. In T. Lengauer, editor, *Proceedings of the 1st Annual European Symposium*, number 726 in Lecture Notes in Computer Science, pages 61–72, Bad Honnef, Germany, 1993. Springer-Verlag, Berlin.

[388] D. Breslauer. Dictionary matching on unbounded alphabets: uniform length dictionaries. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 184–197, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[389] D. Breslauer. Testing string superprimitivity in parallel. *Inf. Process. Lett.*, 49(5):235–241, 1994.

[390] D. Breslauer. Dictionary matching on unbounded alphabets: uniform length dictionaries. *J. Algorithms*, 18(2):278–295, 1995.

[391] D. Breslauer. Fast parallel string prefix-matching. *Theor. Comput. Sci.*, 132(2):269–278, 1995.

[392] D. Breslauer. Saving comparisons in the Crochemore–Perrin string-matching algorithm. *Theor. Comput. Sci.*, 158(1–2):177–192, 1996.

[393] D. Breslauer. The suffix tree of a tree and minimizing sequential transducers. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 116–129, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[394] D. Breslauer. The suffix tree of a tree and minimizing sequential transducers. *Theor. Comput. Sci.*, 191(1–2):131–144, 1998.

[395] D. Breslauer, L. Colussi, and L. Toniolo. Tight comparison bounds for the string prefix-matching problem. *Inf. Process. Lett.*, 47(1):51–57, 1993.

[396] D. Breslauer, L. Colussi, and L. Toniolo. Tight comparison bounds for the string prefix-matching problem. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 11–19, Padova, Italy, 1993. Springer-Verlag, Berlin.

[397] D. Breslauer, L. Colussi, and L. Toniolo. On the comparison complexity of the string prefix-matching problem. *J. Algorithms*, 29(1):18–67, 1998.

[398] D. Breslauer and Z. Galil. An optimal $O(\log \log n)$ time parallel string matching algorithm. *SIAM J. Comput.*, 19(6):1051–1058, 1990.

[399] D. Breslauer and Z. Galil. A lower bound for parallel string matching. In *Proceedings of the 23rd ACM Symposium on the Theory of Computing*, pages 439–443, New Orleans, LA, 1991. ACM Press.

[400] D. Breslauer and Z. Galil. Efficient comparison based string matching. Report CUCS-054-92, Computer Science Department, Columbia University, NY, 1992.

[401] D. Breslauer and Z. Galil. Finding all periods and initial palindromes of a string in parallel. Report CUCS-017-92, Computer Science Department, Columbia University, NY, 1992.

[402] D. Breslauer and Z. Galil. A lower bound for parallel string matching. *SIAM J. Comput.*, 21(5):856–862, 1992.

[403] D. Breslauer and Z. Galil. Parallel string matching algorithms. Report CUCS-002-92, Computer Science Department, Columbia University, NY, 1992.

[404] D. Breslauer and Z. Galil. Efficient comparison based string matching. *J. Complexity*, 9(3):339–365, 1993.

[405] D. Breslauer and Z. Galil. Finding all periods and initial palindromes of a string in parallel. *Algorithmica*, 14(4):355–366, 1995.

[406] D. Breslauer and L. Gąsieniec. Efficient string matching on coded texts. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 27–40, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[407] D. Breslauer and L. Gąsieniec. Efficient string matching on packed texts. *RAIRO Inform. Théor. Appl.*, 30(6):521–544, 1997.

[408] D. Breslauer and R. Hariharan. Optimal parallel construction of minimal suffix and factor automata. *Parallel Process. Lett.*, 6(1):35–44, 1996.

[409] D. Breslauer, T. Jiang, and Z. Jiang. Rotations of periodic strings and short superstrings. *J. Algorithms*, 24(2):340–353, 1997.

[410] S. Brlek. Enumeration of factors in the Thue-Morse word. *Discret. Appl. Math.*, 24:83–96, 1989.

[411] G. S. Brodal. Finger search trees with constant insertion time. In *Proceedings ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 540–549, San Francisco, CA, 1998.

[412] G. S. Brodal and L. Gąsieniec. Approximate dictionary queries. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 65–74, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[413] G. S. Brodal, R. B. Lyngsø, C. N. S. Pedersen, and J. Stoye. Finding maximal pairs with bounded gap. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 134–149, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[414] G. S. Brodal, R. B. Lyngsø, C. N. S. Pedersen, and J. Stoye. Finding maximal pairs with bounded gap. *J. Discret. Algorithms*, 1(1):77–103, 2000.

[415] G. S. Brodal and C. N. S. Pedersen. Finding maximal quasiperiodicities in strings. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 397–411, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[416] G. S. Brodal and S. Venkatesh. Improved bounds for dictionary look-up with one error. *Inf. Process. Lett.*, 75(1–2):57–59, 2000.

[417] A. Z. Broder. Identifying and filtering near-duplicate documents. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 1–10, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[418] C. Bron and E. J. Dijkstra. A better way to combine efficient string length encoding and zero-termination. *SIGPLAN Not.*, 24(6):11–19, 1989.

[419] P. Brown, J. Lai, and R. Mercer. Aligning sentences in parallel corpora. In *Proceedings of the 47th Annual Meeting of the ACL*, 1991.

[420] D. Brutlag and D. Naor. On suboptimal alignments of biological sequences. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber,

editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 179–196, Padova, Italy, 1993. Springer-Verlag, Berlin.

[421] D. L. Brutlag, J. P. Dautricourt, S. Maulik, and J. Relph. Improved sensitivity of biological sequence database searches. *Comput. Appl. Biosci.*, 6(3):237–245, 1990.

[422] V. Bruyère. *Automates de Boyer-Moore*. Thèse annexe, Université de Mons-Hainaut, Belgique, 1991.

[423] V. Bruyère, R. A. Baeza-Yates, O. Delgrange, and R. Scheihing. About the size of Boyer-Moore automata. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 31–46, Recife, Brazil, 1996. Carleton University Press.

[424] V. Bruyère and O. Delgrange. About the size of the Boyer-Moore automaton. Report, Université de Mons-Hainaut, Belgique, 1994.

[425] V. Bruyère and D. Perrin. Maximal bifix codes. *Theor. Comput. Sci.*, 218(1):107–121, 1999.

[426] D. Bryant. A lower bound for the breakpoint phylogeny problem. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 235–247, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[427] A. L. Buchsbaum and R. Giancarlo. Algorithmic aspects of speech recognition: A synopsis. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 27–32, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[428] R. Bundschuh and T. Hwa. An analytic study of the phase transition line in local sequence alignment with gaps. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 3rd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 70–76, Lyon, France, 1999. ACM Press.

[429] H. Bunke. Fast approximate matching of words against a dictionary. *Computing*, 55(1):75–89, 1995.

[430] H. Bunke and U. Buehler. Applications of approximate string matching to 2D shape recognition. *Pattern Recognit.*, 26(12):1797–1812, 1993.

[431] H. Bunke and J. Csirik. An algorithm for matching run-length coded strings. *Comput.*, 50:297–314, 1993.

[432] H. Bunke and J. Csirik. An improved algorithm for computing the edit distance of run-length coded strings. *Inf. Process. Lett.*, 54(2):93–96, 1995.

[433] J. Burghardt. A tree pattern matching algorithm with reasonable space requirements. In M. Dauchet and M. Nivat, editors, *Proceedings of the 13th Colloquium on Trees in Algebra and Programming*, number 299 in Lecture Notes in Computer Science, pages 1–15, Nancy, France, 1988. Springer-Verlag, Berlin.

[434] W. A. Burkhard. Hashing and trie algorithms for partial match retrieval. *ACM Trans. Database Syst.*, 1:175–187, 1976.

[435] S. Burkhardt, A. Crauser, P. Ferragina, H.-P. Lenhof, E. Rivals, and M. Vingron. q-gram based database searching using a suffix array. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 3rd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 77–83, Lyon, France, 1999. ACM Press.

[436] F. J. Burkowski. A hardware hashing schemme in the design of a multiterm string comparator. *IEEE Trans. Comput.*, C-31(9):825–834, 1982.

[437] CACM. Notes on Sunday's paper. *Commun. ACM*, 35(4):132–137, 1992.

[438] J. Cai, R. Paige, and R. Tarjan. More efficient bottom-up multi-pattern matching in trees. *Theor. Comput. Sci.*, 106(1):21–60, 1992.

[439] J. Cai, R. Paige, and R. Tarjan. More efficient bottom-up multi-pattern matching in trees. Report 604, Courant Institut, New York University, NY, 1992.

[440] J. Cai, R. Paige, and R. Tarjan. More efficient bottom-up multi-pattern matching in trees. *Theor. Comput. Sci.*, 106(1):21–60, 1992.

[441] A. Califano. SPLASH: Structural pattern localization analysis by sequential histograming. *Bioinformatics*, 15:341–357, 2000.

[442] E. Cambouropoulos, M. Crochemore, C. S. Iliopoulos, L. Mouchard, and Y. J. Pinzon. Algorithms for computing approximate repetitions in musical sequences. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 129–144, Perth, WA, Australia, 1999.

[443] Matteo Campanelli, Domenico Cantone, Simone Faro, and Emanuele Giaquinta. An efficient algorithm for approximate pattern matching with swaps. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 90–104, Czech Technical University in Prague, Czech Republic, 2009.

[444] Domenico Cantone, Salvatore Cristofaro, and Simone Faro. New efficient bit-parallel algorithms for the δ-matching problem with α-bounded gaps in musical sequences. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 170–184, Czech Technical University in Prague, Czech Republic, 2008.

[445] Domenico Cantone, Simone Faro, and Emanuele Giaquinta. Adapting boyer-moore-like algorithms for searching huffman encoded texts. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 29–39, Czech Technical University in Prague, Czech Republic, 2009.

[446] C. R. Cantor. Enhancements in sequence analysis with DNA arrays. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 50–50, New York, NY, 1998. ACM Press.

[447] R. Capocelli, A. De Santis, and U. Vaccaro, editors. *Sequences II*. Springer-Verlag, Berlin, 1993.

[448] A. Caprara. Sorting by reversals is difficult. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 75–83, Santa Fe, NM, 1997. ACM Press.

[449] A. Caprara. Formulations and complexity of multiple sorting by reversals. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 3rd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 84–93, Lyon, France, 1999. ACM Press.

[450] A. Caprara. Sorting permutations by reversals and Eulerian cycle decompositions. *SIAM J. Discret. Math.*, 12(1):91–110, 1999.

[451] A. Caprara, G. Lancia, and See Kiong Ng. Fast practical solution of sorting by reversals. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 12–21, San Francisco, CA, 2000.

[452] A. Cardon and C. Charras. *Introduction à l'algorithmique et à la programmation*. Ellipses, 1996.

[453] H. Carillo and D. Lipman. The multiple sequence alignment problem in biology. *SIAM J. Appl. Math.*, 48:1073–1082, 1988.

[454] A. Carpi. On abelian squares and substitutions. *Theor. Comput. Sci.*, 218(1):61–81, 1999.

[455] J. Cassaigne. Unavoidable binary patterns. *Acta Inform.*, 30(4):385–395, 1993.

[456] J. Cassaigne. *Motifs évitables et régularité dans les mots*. Thèse de doctorat, Université Paris 6, France, 1994.

[457] J. Cassaigne. Motifs évitables et régularité dans les mots. Rapport LITP 94-04, Universités Paris 6 et 7, France, 1994.

[458] J. Cassaigne. Limit values of the recurrence quotient of sturmian sequences. *Theor. Comput. Sci.*, 218(1):3–12, 1999.

[459] G. Castelli, F. Mignosi, and A. Restivo. Fine and Wilf's theorem for three periods and a generalization of sturmian words. *Theor. Comput. Sci.*, 218(1):83–94, 1999.

[460] K. Cattell, F. Ruskey, J. Sawada, M. Serra, and C. R. Miers. Fast algorithms to generate necklaces, unlabeled necklaces, and irreducible polynomials over GF(2). *J. Algorithms*, 37(2):267–282, 2000.

[461] S. Cenk Sahinalp and U. Vishkin. Symmetry breaking for suffix tree construction. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 300–309, Montréal, Canada, 1994. ACM Press.

[462] Y. Césari and M. Vincent. Une caractérisation des mots périodiques. *Comptes Rendus Acad. Sci. Paris Sér. I Math.*, pages 1175–1177, 1978.

[463] J.-M. Champarnaud, A. Khorsi, and T. Paranthoën. Split and join for minimizing: Brzozowski's algorithm. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 96–104, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[464] D. M. Champion and J. Rothstein. Immediate parallel solution of the longest common subsequence problem. *ICPP*, pages 70–77, 1987.

[465] S. C. Chan, A. K. C. Wong, and D. K. Y. Chiu. A survey of multiple sequence comparison methods. *Bull. Math. Biol.*, 54(4):563–598, 1992.

[466] C.-H. Chang and R. Paige. From regular expression to DFA's using NFA's. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 90–110, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[467] D. K. Chang. A string pattern-matching algorithm. *J. Syst. Softw.*, 22(3):207–?, 1993.

[468] W. I. Chang. *Approximate pattern matching and biological applications.* Ph. D. Thesis, University of California, Berkeley, 1991.

[469] W. I. Chang and J. Lampe. Theoretical and empirical comparisons of approximate string matching algorithms. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 175–184, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[470] W. I. Chang and E. L. Lawler. Approximated string matching in sublinear expected time. In *Proceedings of the 31st IEEE Annual Symposium on Foundations of Computer Science*, pages 116–124, Saint Louis, MO, 1990. IEEE Computer Society Press.

[471] W. I. Chang and E. L. Lawler. Sublinear approximate string matching and biological applications. *Algorithmica*, 12(4/5):327–344, 1994.

[472] W.I. Chang and T. Marr. Approximate string matching with local similarity. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 259–273, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[473] K-M. Chao. Computing all suboptimal alignments in linear space. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 31–42, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[474] K.-M. Chao, R. C. Hardison, and W. Miller. Locating well-conserved regions within a pairwise alignment. *Comput. Appl. Biosci.*, 4:387–396, 1993.

[475] K.-M. Chao and W. Miller. Linear-space algorithms that build local alignments from fragments. *Algorithmica*, 13(1/2):106–134, 1995.

[476] K.-M. Chao, J. Zhang, J. Ostell, and W. Miller. A local alignment tool for very long DNA sequences. *Comput. Appl. Biosci.*, 11(2):147–154, 1995.

[477] W. Charatonic and L. Pacholski. Solving word equations in two variables. In *Proceedings IWWERT*, Lecture Notes in Computer Science, pages 43–56. Springer-Verlag, Berlin, 1991.

[478] C. Charras and T. Lecroq. Exact string matching algorithms. URL:`http://www-igm.univ-mlv.fr/~lecroq/string/`, 1996.

[479] C. Charras and T. Lecroq. Exact string matching animation in java. Rapport LIR 97.10, Université de Rouen, France, 1997.

[480] C. Charras and T. Lecroq. Exact string matching animation in java. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 36–43, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[481] C. Charras and T. Lecroq. Sequence comparison. URL:`http://www-igm.univ-mlv.fr/~lecroq/seqcomp/`, 1998.

[482] C. Charras and T. Lecroq. Java tools to help understanding pattern matching techniques. In D.C. Dimitrov, editor, *Proceedings of the 5th International Conference on Compute Aided Engineering Education*, pages 105–111, Sofia, Bulgaria, 1999.

[483] C. Charras, T. Lecroq, and J.D. Pehoushek. A very fast string matching algorithm for small alphabets and long patterns. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 55–64, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[484] D. Chase. An improvement to bottom-up tree pattern matching. In *Proceedings of the 14th Annual ACM Symposium on Principles of Programming Languages*, pages 168–177, Munich, Germany, 1987. ACM Press.

[485] E. Chavez and J.L. Marroquín. Proximity queries in metric spaces. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 21–36, Valparaiso, Chile, 1997. Carleton University Press.

[486] G. H. Chen. An $O(1)$ time algorithm for string matching. *Int. J. Comput. Math.*, 42(3–4):185–191, 1992.

[487] M. T. Chen and J. Seiferas. Efficient and elegant subword tree construction. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 97–107. Springer-Verlag, Berlin, 1985.

[488] T. Chen and S. S. Skiena. Trie-based data structures for sequence assembly. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 206–223, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[489] W. Chen. Multi-subsequence searching. *Inf. Process. Lett.*, 74(5–6), 2000.

[490] H. Cheng and K. Fu. VLSI architectures for string matching and pattern matching. *Pattern Recognit.*, 20(1):125–141, 1987.

[491] F. Y. L. Chin and C. K. Poon. A fast algorithm for computing longest common subsequences of small alphabet size. *J. Inform. Process.*, 13(4):463–469, 1990.

[492] F. Y. L. Chin and C. K. Poon. Performance analysis of some simple heuristics for computing longuest common subsequences. *Algorithmica*, 12(4/5):293–311, 1994.

[493] B. S. Chlebus and L. Gąsieniec. Optimal pattern matching on meshes. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 213–224, Caen, France, 1994. Springer-Verlag, Berlin.

[494] M. Chodorow and J. L. Klarans. Locating syntactic patterns in text corpora. Report, IBM T.J.Watson Research Center, New York, NY, 1990.

[495] C. Choffrut. An optimal algorithm for building the Boyer-Moore automaton. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 40:217–225, 1990.

[496] C. Choffrut. Combinatoire et génome. Rapport LITP 91.72, Universités Paris 6 et 7, 1991.

[497] C. Choffrut. The word problem for single relator special monoids - the state of art. Rapport LITP 93.37, Universités Paris 6 et 7, 1993.

[498] C. Choffrut and S. Horwath. Equations in transfinite strings. In L. Brim, J. Gruska, and J. Zlatuska, editors, *Proceedings of the 23rdSymposium on Mathematical Foundations of Computer Science*, number 1450 in Lecture Notes in Computer Science, pages 656–664, Brno, Czech Republic, 1998. Springer-Verlag, Berlin.

[499] Y. Choi and T. W. Lam. Two-dimensional pattern matching on a dynamic library of texts. In D.-Z. Du and M. Li, editors, *Proceedings of the 1stAnnual International Conference on Computing and Combinatorics*, number 959 in Lecture Notes in Computer Science, pages 530–538, Xi'an, China, 1995. Springer-Verlag, Berlin.

[500] Y. Choi and T. W. Lam. Dynamic suffix tree and two-dimensional texts management. *Inf. Process. Lett.*, 61(4):213–220, 1997.

[501] N. Chomsky and M.-P. Schützenberger. The algebraic theory of context-free languages. In P. Braffort and D. Hirschberg, editors, *Computer Programming and Formal Systems*, pages 118–161. North-Holland, Amsterdam, The Netherlands, 1963.

[502] E. T. Chow, T. Hunkapiller, J. C. Peterson, B. A. Zimmerman, and M. S. Waterman. A systolic array processor for biological information signal processing. In *Proceedings of the International Conference on Supercomputing*, pages 216–223, Cologne, Germany, 1991.

[503] Manolis Christodoulakis and Gerhard Brey. Edit distance with single-symbol combinations and splits. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 208–217, Czech Technical University in Prague, Czech Republic, 2008.

[504] M. Chrobak and W. Rytter. Remarks on string-matching and one-way multihead automata. *Inf. Process. Lett.*, 24(5):325–329, 1987.

[505] M. Chrobak and W. Rytter. Unique decipherability for partially commutative alphabets. *Fundam. Inform.*, pages 323–336, 1987.

[506] W.-F. Chuan. Symmetric fibonacci words. *Fibonacci Quarterly*, 31:251–255, 1993.

[507] K.-L. Chung. A randomized parallel algorithm for string matching on hypercube. *Pattern Recognit.*, 25(10):1265–1268, 1992.

[508] K.-L. Chung. An improved algorithm for solving the banded cyclic string-to-string correction problem. *Theor. Comput. Sci.*, 201(1–2):275–279, 1998. Abstract.

[509] T.-C. Chung and J.-W. Cho. History sensitive string for multiple alphabets. *Inf. Process. Lett.*, 25(3):183–188, 1987.

[510] V. Chvátal, D.A. Klarner, and D.E. Knuth. Selected Combinatorial Research Problems. Technical Report STAN–CS–72–292, Computer Science Department, Stanford University, 1972.

[511] V. Chvátal and D. Sankoff. Longest common subsequences of random sequences. Technical Report STAN–CS–75–477, Computer Science Department, Stanford University, 1975.

[512] V. Chvátal and D. Sankoff. Longest common subsequences of two random sequences. *J. Appl. Probab*, 12:306–315, 1975.

[513] Ferdinando Cicalese, Gabriele Fici, and Zsuzsanna Lipták. Searching for jumbled patterns in strings. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 105–117, Czech Technical University in Prague, Czech Republic, 2009.

[514] Luigi Cinque and Sergio De Agostino. Lossless image compression by block matching on practical massively parallel architectures. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 26–34, Czech Technical University in Prague, Czech Republic, 2008.

[515] Luigi Cinque, Sergio De Agostino, and Luca Lombardi. Speeding up lossless image compression: Experimental results on a parallel machine. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 35–45, Czech Technical University in Prague, Czech Republic, 2008.

[516] D. R. Clark and J. I. Munro. Efficient suffix trees on secondary storage. In *Proceedings of the 7th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 383–391, Atlanta, Georgia, 1996.

[517] J. G. Cleary and I. H. Witten. Data compression using adaptive coding and partial string matching. *IEEE Trans. Comput.*, 32:396–402, 1984.

[518] Loek Cleophas and Kees Hemerik. Taxonomies of regular tree algorithms. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 146–159, Czech Technical University in Prague, Czech Republic, 2009.

[519] Loek Cleophas, Derrick G. Kourie, Tinus Strauss, and Bruce W. Watson. On minimizing deterministic tree automata. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 173–182, Czech Technical University in Prague, Czech Republic, 2009.

[520] B. Clift, D. Haussler, R. McConnell, T. D. Schneider, and G. D. Stormo. Sequences landscapes. *Nucleic Acids Res.*, 14:141–158, 1986.

[521] A. L. Cobbs. Fast identification of approximately matching substrings. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 64–74, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[522] A. L. Cobbs. Fast approximate matching using suffix trees. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 41–54, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[523] Wikus Coetser, Derrick G. Kourie, and Bruce W. Watson. On regular expression hashing to reduce FA size. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 227–241, Czech Technical University in Prague, Czech Republic, 2008.

[524] J. D. Cohen. Hardware-assisted algorithm for full-text large-dictionary string matching using n-gram hashing. *Inf. Process. Manage.*, 34(4), 1998.

[525] M. Cohn and R. Khazan. Parsing with suffix and prefix dictionaries. In J. A. Storer and M. Cohn, editors, *Proceedings Data Compression Conference*, pages 180–189, Snowbird, UTah, 1996. IEEE Computer Society Press.

[526] R. Cole. A new proof of the linearity of the Boyer-Moore string-matching algorithm. In *3rd Maryland Theoretical Science Day*, pages ?–?, 1990.

[527] R. Cole. Tight bounds on the complexity of the Boyer-Moore string matching algorithm. Report 512, Courant Institute, New York University, NY, 1990.

[528] R. Cole. Tight bounds on the complexity of the Boyer-Moore string matching algorithm. In *Proceedings of the 2nd ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 224–233, San Francisco, CA, 1991.

[529] R. Cole. Tight bounds on the complexity of the Boyer-Moore string matching algorithm. *SIAM J. Comput.*, 23(5):1075–1091, 1994.

[530] R. Cole, M. Crochemore, Z. Galil, L. Gąsieniec, R. Hariharan, S. Muthukrishnan, K. Park, and W. Rytter. Optimally fast parallel algorithms for preprocessing and pattern matching in one and two dimensions. In *Proceedings of the 34th IEEE Annual Symposium on Foundations of Computer Science*, pages 248–258, Palo Alto, CA, 1993. IEEE Computer Society Press.

[531] R. Cole and R. Hariharan. Tighter bounds on the exact complexity of string matching. In *Proceedings of the 33rd IEEE Annual Symposium on Foundations of Computer Science*, pages 600–609, Pittsburgh, PA, 1992. IEEE Computer Society Press.

[532] R. Cole and R. Hariharan. Tighter upper bounds on the exact complexity of string matching. *SIAM J. Comput.*, 26(3):803–856, 1997.

[533] R. Cole and R. Hariharan. Tree pattern matching and subset matching in randomized $o(n \log^3 m)$ time. In *Proceedings of the 29th ACM Symposium on the Theory of Computing*, pages 66–75, El Paso, TX, 1997.

[534] R. Cole and R. Hariharan. Approximate string matching: A faster simpler algorithm. In *Proceedings of the 9th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 463–472, San Francisco, CA, 1998.

[535] R. Cole and R. Hariharan. Dynamic LCA queries on trees. In *Proceedings ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 235–244, 1999.

[536] R. Cole and R. Hariharan. Faster suffix tree construction with missing suffix links. In *Proceedings ACM Symposium on the Theory of Computing*, pages 407–415, Portland, Oregon, 2000.

[537] R. Cole, R. Hariharan, M. Paterson, and U. Zwick. Which patterns are hard to find ? In *Proceedings of the 2nd Israeli Symposium on Theory of Computing and Systems*, pages 59–68. IEEE Computer Society Press, 1993.

[538] R. Cole, R. Hariharan, M. Paterson, and U. Zwick. Tighter lower bounds on the exact complexity of string matching. *SIAM J. Comput.*, 24(1):30–45, 1995.

[539] R. Cole, C. S. Iliopoulos, T. Lecroq, W. Plandowski, and W. Rytter. Relations between δ-matching and matching with don't care symbols: δ-distinguishing morphisms. In E. Tri Baskoro, editor, *Proceedings of the*

12th Australasian Workshop On Combinatorial Algorithms, pages 80–87, Bandung, Indonesia, 2001.

[540] R. Cole, C. S. Iliopoulos, T. Lecroq, W. Plandowski, and W. Rytter. On special families of morphisms related to δ-matching and don't care symbols. *Inf. Process. Lett.*, 85(5):227–233, 2003.

[541] J. F. Collins and A. F. W. Coulson. Molecular sequence comparison and alignments. In M. J. Bishop and C. J Rawlings, editors, *Nucleic acid and protein sequence analysis: A practical approach*, pages 323–358. IRL Press, 1987.

[542] L. Colussi. Correctness and efficiency of the pattern matching algorithms. *Inf. Comput.*, 95(2):225–251, 1991.

[543] L. Colussi. Fastest pattern matching in strings. *J. Algorithms*, 16(2):163–189, 1994.

[544] L. Colussi and A. De Col. A time and space efficient data structure for sring searching on large texts. *Inf. Process. Lett.*, 58(5):217–222, 1996.

[545] L. Colussi, Z. Galil, and R. Giancarlo. On the exact complexity of string matching. In *Proceedings of the 31st IEEE Annual Symposium on Foundations of Computer Science*, pages 135–144, Saint Louis, MO, 1990. IEEE Computer Society Press.

[546] L. Colussi and L. Toniolo. How the character comparison order shapes the shift function of on-line pattern matching algorithms. *Theor. Comput. Sci.*, 163(1–2):117–144, 1996.

[547] B. Commentz-Walter. A string matching algorithm fast on the average. Report 79.09.007, IBM Scientific Center in Heidelberg, 1979.

[548] B. Commentz-Walter. A string matching algorithm fast on the average. In H. A. Maurer, editor, *Proceedings of the 6th International Colloquium on Automata, Languages and Programming*, number 71 in Lecture Notes in Computer Science, pages 118–132, Graz, Austria, 1979. Springer-Verlag, Berlin.

[549] C. Consel and O. Danvy. Partial evaluation of pattern matching in strings. *Inf. Process. Lett.*, 30(2):79–86, 1989.

[550] Cook. Linear time simulation of deterministic two-way pushdown automata. In *Proceedings of the 5th IFIP World Computer Congress*, pages 172–179, Ljubljana, Yugoslavia, 1971.

[551] G. P. Copeland. String storage and searching database applications: implementation on the indy backend kernel. In *Proceedings of the 4th Workshop on Computer Architecture for Non-numeric Processing*, pages 8–17, New York, NY, 1978. ACM Press.

[552] N. G. Core, E. W. Edmiston, J. H. Saltz, and R. M. Smith. Supercomputers and biological sequence comparison algorithms. *Comp. Biomed. Res.*, 22:497–515, 1989.

[553] R. Cori and D. Perrin. Automates et commutations partielles. *RAIRO Inform. Théor.*, 19(1):21–32, 1985.

[554] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, 1990.

[555] F. Corpet. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.*, 16:10881–10890, 1988.

[556] E. Coward and F. Drabløs. Detecting periodic patterns in biological sequences. *Bioinformatics*, 14(6):498–507, 1998.

[557] T. C. Craven. *String Indexing*. Academic Press, London, U. K., 1986.

[558] T. C. Craven. Adapting of string indexing systems for retrieval using proximity operators. *Inf. Process. Manage.*, 24(2):133–140, 1988.

[559] M. Crochemore. Détermination optimale des répétitions dans un mot. Rapport LITP 80-53, Universités Paris 6 et 7, France, 1980.

[560] M. Crochemore. An optimal algorithm for computing the repetitions in a word. *Inf. Process. Lett.*, 12(5):244–250, 1981.

[561] M. Crochemore. Résultats de décidabilité pour les morphismes faiblement sans carré. Rapport LITP 82-38, Universités Paris 6 et 7, France, 1982.

[562] M. Crochemore. Sharp characterizations of squarefree morphisms. *Theor. Comput. Sci.*, 18(2):221–226, 1982.

[563] M. Crochemore. A solution to Berstel's problem P3. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 18:9–11, 1982.

[564] M. Crochemore. Sur les ensembles inévitables. Rapport LITP 82-56, Universités Paris 6 et 7, France, 1982.

[565] M. Crochemore. Mots et morphismes sans carré. *Ann. Discret. Math.*, 17:235–245, 1983.

[566] M. Crochemore. Recherche linéaire d'un carré dans un mot. *Comptes Rendus Acad. Sci. Paris Sér. I Math.*, 296:781–784, 1983.

[567] M. Crochemore. Recherche linéaire d'un carré dans un mot. *Comptes Rendus Acad. Sci. Paris Sér. I Math.*, T. 296:781–784, 1983.

[568] M. Crochemore. *Régularités évitables*. Thèse d'état, Rapport LITP 83-43, Université Paris 7, France, 1983.

[569] M. Crochemore. A solution to P12. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 19:15–16, 1983.

[570] M. Crochemore. Tests sur les morphismes faiblement sans carré. In L. Cummings, editor, *Combinatorics on Words*, pages 63–89. Academic Press, 1983.

[571] M. Crochemore. Linear searching for a square in a word. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 24:66–72, 1984.

[572] M. Crochemore. Optimal factor tranducers. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 31–44. Springer-Verlag, Berlin, 1985.

[573] M. Crochemore. Calcul de la f-distance de deux mots. In *Actes du séminaire d'informatique théorique du L.I.T.P.*, pages 25–36. Universités Paris 6 et 7, France, 1986.

[574] M. Crochemore. Computing LCF in linear time. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 30:57–61, 1986.

[575] M. Crochemore. Transducers and repetitions. *Theor. Comput. Sci.*, 45(1):63–86, 1986.

[576] M. Crochemore. Longest common factor of two words. In H. Ehrig, R. Kowalski, G. Levi, and U. Montanari, editors, *Proceedings of the 2nd Theory and Practice of Software Development*, number 249 in Lecture Notes in Computer Science, pages 26–36. Springer-Verlag, Berlin, Pisa, Italy, 1987.

[577] M. Crochemore. Compression de textes: les méthodes. Rapport LIPN 88–4, Université Paris-Nord, France, 1988.

[578] M. Crochemore. Constant-space string-matching. In K. V. Nori and S. Kumar, editors, *Proceedings of the 8th Conference on Foundations of Software Technology and Theoretical Computer Science*, number 338 in Lecture Notes in Computer Science, pages 80–87, Pune, India, 1988. Springer-Verlag, Berlin.

[579] M. Crochemore. String matching with constraints. In M. Chytil, L. Janiga, and V. Koubek, editors, *Proceedings 13th Symposium on Mathematical Foundations of Computer Science*, number 324 in Lecture Notes in Computer Science, pages 44–58, Carlsbad, Czechoslovakia, 1988. Springer-Verlag, Berlin.

[580] M. Crochemore. Automata and algorithms. In J.-E. Pin, editor, *Formal Properties of Finite Automata and Applications*, number 386 in Lecture Notes in Computer Science, pages 149–153. Springer-Verlag, Berlin, 1989.

[581] M. Crochemore. Data compression with substitution. In M. Gross and D. Perrin, editors, *Electronic, Dictionaries and Automata in Computational Linguistics*, number 377 in Lecture Notes in Computer Science, pages 1–16. Springer-Verlag, Berlin, 1989.

[582] M. Crochemore. String-matching and periods. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 39:149–153, 1989.

[583] M. Crochemore. Constant-space string-matching. In R. M. Capocelli, editor, *Sequences, combinatorics, compression, security and transmission*, pages 25–40. Springer-Verlag, Berlin, 1990.

[584] M. Crochemore. A la recherche d'un mot. In V. Bruyère, editor, *Actes des Journées Montoises*, pages 34–38, Université de Mons-Hainaut, Belgique, 1990.

[585] M. Crochemore. String-matching on ordered alphabets. *Theor. Comput. Sci.*, 92(1):33–47, 1992.

[586] M. Crochemore. Off-line serial exact string searching. Rapport 93.1, Institut Gaspard Monge, Université de Marne la Vallée, 1993.

[587] M. Crochemore. Off-line serial exact string searching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 1, pages 1–53. Oxford University Press, 1997.

[588] M. Crochemore. Reducing space for index implementation. *Theor. Comput. Sci.*, 292(1):185–197, 2003.

[589] M. Crochemore, J. Berstel, and J.-E. Pin. Thue-morse sequence and p-adic topology of the free monoid. *Ann. Discret. Math.*, 76:89–94, 1989.

[590] M. Crochemore, M. Chytil, B. Monien, and W. Rytter. On the parallel recognition of unambiguous context-free languages. *Theor. Comput. Sci.*, 81(2):311–316, 1991.

[591] M. Crochemore, A. Czumaj, L. Gąsieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter. Deux méthodes pour accélerer l'algorithme de Boyer-Moore. In D. Krob, editor, *Actes des 2e Journées franco-belges: Théories des Automates et Applications*, number 176 in Publications de l'Université de Rouen, pages 45–63, Rouen, France, 1991.

[592] M. Crochemore, A. Czumaj, L. Gąsieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter. Speeding up two string matching algorithms. In A. Finkel and M. Jantzen, editors, *Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, number 577 in Lecture Notes in Computer Science, pages 589–600, Cachan, France, 1992. Springer-Verlag, Berlin.

[593] M. Crochemore, A. Czumaj, L. Gąsieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter. Speeding up two string matching algorithms. Rapport LITP 92.21, Universités Paris 6 et 7, France, 1992.

[594] M. Crochemore, A. Czumaj, L. Gąsieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter. Fast practical multi-pattern matching. Rapport 93-3, Institut Gaspard Monge, Université de Marne la Vallée, 1993.

[595] M. Crochemore, A. Czumaj, L. Gąsieniec, S. Jarominek, T. Lecroq, W. Plandowski, and W. Rytter. Speeding up two string matching algorithms. *Algorithmica*, 12(4/5):247–267, 1994.

[596] M. Crochemore, A. Czumaj, L. Gąsieniec, T. Lecroq, W. Plandowski, and W. Rytter. Fast practical multi-pattern matching. *Inf. Process. Lett.*, 71((3–4)):107–113, 1999.

[597] M. Crochemore, Z. Galil, L. Gąsieniec, K. Park, and W. Rytter. Constant-time randomized parallel string matching. *SIAM J. Comput.*, 26(4):950–960, 1997.

[598] M. Crochemore, L. Gąsieniec, R. Hariharan, S. Muthukrishnan, and W. Rytter. A constant time optimal parallel algorithm for two-dimensional pattern matching. *SIAM J. Comput.*, 27(3):668–681, 1998.

[599] M. Crochemore, L. Gąsieniec, W. Plandowski, and W. Rytter. Two-dimensional pattern matching in linear time and small space. In E.W. Mayr and C. Puech, editors, *Proceedings of the 12th Annual Symposium on Theoretical Aspects of Computer Science*, number 900 in Lecture Notes in Computer Science, pages 181–192, Munich, Germany, 1995. Springer-Verlag, Berlin.

[600] M. Crochemore, L. Gąsieniec, and W. Rytter. Turbo-BM. Rapport LITP 92.61, Universités Paris 6 et 7, France, 1992.

[601] M. Crochemore, L. Gąsieniec, and W. Rytter. Two-dimensional pattern matching by sampling. *Inf. Process. Lett.*, 46(4):159–162, 1993.

[602] M. Crochemore, L. Gąsieniec, and W. Rytter. Constant-space string-matching in sublinear average time. *Theor. Comput. Sci.*, 218(1):197–203, 1999.

[603] M. Crochemore and P. Goralcik. Mutually avoiding ternary words of small exponent. Rapport LITP 91.26, Universités Paris 6 et 7, 1991.

[604] M. Crochemore and P. Goralcik. Mutually avoiding ternary words of small exponent. *Int. J. Algebr. Comput.*, 1(4):407–410, 1992.

[605] M. Crochemore and C. Hancart. Automata for matching patterns. Rapport LIR96.17 Informatique Fondamentale, Université de Rouen, France, 1996.

[606] M. Crochemore and C. Hancart. Pattern matching in strings. Rapport LIR96.18 Informatique Fondamentale, Université de Rouen, France, 1996.

[607] M. Crochemore and C. Hancart. Automata for matching patterns. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 2 Linear Modeling: Background and Application, chapter 9, pages 399–462. Springer-Verlag, Berlin, 1997.

[608] M. Crochemore and C. Hancart. Pattern matching in strings. In Mikhail J. Atallah, editor, *Algorithms and Theory of Computation Handbook*, pages 11.1–11.28. CRC Press Inc., Boca Raton, FL, 1998.

[609] M. Crochemore, C. Hancart, and T. Lecroq. A unifying look at the apostolico-giancarlo string matching algorithm. *J. Discret. Algorithms*, 2(1). To appear.

[610] M. Crochemore, C. Hancart, and T. Lecroq. *Algorithmique du texte.* Vuibert, 2001.

[611] M. Crochemore, C. Iliopoulos, and M. Korda. Two-dimensional prefix string matching and covering on square matrix. *Algorithmica*, 20(4):353–373, 1998.

[612] M. Crochemore, C. S. Iliopoulos, M. Korda, and J. F. Reid. Two-dimensional dictionary prefix-matching. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 147–158, Perth, WA, Australia, 1998.

[613] M. Crochemore, C. S. Iliopoulos, T. Lecroq, and Y. J. Pinzon. Approximate string matching in musical sequences. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference'01*, pages 26–36, Prague, Czech Republic, 2001. Annual Report DC–2001–06.

[614] M. Crochemore, C. S. Iliopoulos, T. Lecroq, W. Plandowski, and W. Rytter. Three heuristics for δ-matching: δ-BM algorithms. In A. Apostolico and M. Takeda, editors, *Proceedings of the 13th Annual Symposium on Combinatorial Pattern Matching*, number 2373 in Lecture Notes in Computer Science, pages 178–189, Fukuoka, Japan, 2002. Springer-Verlag, Berlin.

[615] M. Crochemore, C. S. Iliopoulos, Y. J. Pinzon, and J. F. Reid. A fast and practical bit-vector algorithm for the longest common subsequence problem. In L. Brankovic and J. Ryan, editors, *Proceedings of the 11th Australasian Workshop On Combinatorial Algorithms*, pages 75–86, Hunter Valley, Australia, 2000.

[616] M. Crochemore, C. S. Iliopoulos, Y. J. Pinzon, and J. F. Reid. A fast and practical bit-vector algorithm for the longest common subsequence problem. *Inf. Process. Lett.*, 80(6):279–285, 2001.

[617] M. Crochemore, C. S. Iliopoulos, and H. Yu. Algorithms for computing evolutionary chains in molecular and musical sequences. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 172–184, Perth, WA, Australia, 1998.

[618] M. Crochemore, C.S. Iliopoulos, T. Lecroq, Y.J. Pinzon, W. Plandowski, and W. Rytter. Occurrence and substring heuristics for δ-matching. *Fundamenta Informaticae*, 2003. Accepted.

[619] M. Crochemore and T. Lecroq. Pattern matching and text data compression algorithms. Rapport LIR95.11 Informatique Fondamentale, Université de Rouen, France, 1995.

[620] M. Crochemore and T. Lecroq. Pattern matching and text data compression algorithms. Rapport IGM 95–21, Université de Marne-la-Vallée, France, 1995.

[621] M. Crochemore and T. Lecroq. Pattern matching and text data compression algorithms. *ACM Comp. Surv.*, 28(1):39–41, 1996.

[622] M. Crochemore and T. Lecroq. Pattern matching and text data compression algorithms. In Allen B. Tucker Jr., editor, *The Computer Science and Engineering Handbook*, chapter 8, pages 162–202. CRC Press Inc., Boca Raton, FL, 1996.

[623] M. Crochemore and T. Lecroq. Text data compression algorithms. Rapport LIR96.15 Informatique Fondamentale, Université de Rouen, France, 1996.

[624] M. Crochemore and T. Lecroq. Tight bounds on the complexity of the Apostolico-Giancarlo algorithm. Rapport LIR96.12 Informatique Fondamentale, Université de Rouen, France, 1996.

[625] M. Crochemore and T. Lecroq. Tight bounds on the complexity of the Apostolico-Giancarlo algorithm. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 64–74, Recife, Brazil, 1996. Carleton University Press.

[626] M. Crochemore and T. Lecroq. Tight bounds on the complexity of the Apostolico-Giancarlo algorithm. *Inf. Process. Lett.*, 63(4):195–203, 1997.

[627] M. Crochemore and T. Lecroq. Text data compression algorithms. In Mikhail J. Atallah, editor, *Algorithms and Theory of Computation Handbook*, pages 12.1–12.23. CRC Press Inc., Boca Raton, FL, 1998.

[628] M. Crochemore, M. Lerest, and P. Wender. Sur les ensembles inévitables. Rapport LITP 82–56, Université Paris 6 et 7, France, 1982.

[629] M. Crochemore, M. Lerest, and P. Wender. An optimal test on finite unavoidable sets of words. *Inf. Process. Lett.*, 16(4):179–180, 1983.

[630] M. Crochemore, F. Mignosi, and A. Restivo. Automata and forbidden words. Technical Report IGM 98-5, Institut Gaspard-Monge, Université de Marne-la-Vallée, France, 1998.

[631] M. Crochemore, F. Mignosi, and A. Restivo. Automata and forbidden words. *Inf. Process. Lett.*, 67(3):111–117, 1998.

[632] M. Crochemore, F. Mignosi, and A. Restivo. Minimal forbidden words and factor automata. In L. Brim, J. Gruska, and J. Zlatuska, editors, *Proceedings of the 23rd Symposium on Mathematical Foundations of Computer Science*, number 1450 in Lecture Notes in Computer Science, pages 665–673, Brno, Czech Republic, 1998. Springer-Verlag, Berlin.

[633] M. Crochemore, F. Mignosi, A. Restivo, and S. Salemi. Data compression using antidictionaries. Technical Report IGM 98-10, Institut Gaspard-Monge, Université de Marne-la-Vallée, France, 1998.

[634] M. Crochemore, F. Mignosi, A. Restivo, and S. Salemi. Text compression using antidictionaries. In J. Wiedermann, P. van Emde Boas, and M. Nielsen, editors, *Proceedings of the 26th International Colloquium on Automata, Languages and Programming*, number 1644 in Lecture Notes in Computer Science, pages 261–270. Springer-Verlag, Berlin, 1999.

[635] M. Crochemore and J. Néraud. Unitary monoid with two generators: an algorithmic point of view. In A. Arnold, editor, *Proceedings of the 15th Colloquium on Trees in Algebra and Programming*, number 431 in Lecture Notes in Computer Science, pages 117–131, Copenhagen, Denmark, 1990. Springer-Verlag, Berlin.

[636] M. Crochemore and J.-P. Pécuchet. Compression de dictionnaires. Rapport LADL, Université Paris 7, France, 1987.

[637] M. Crochemore and J.-P. Pécuchet. Réduction d'un lexique pour micro-ordinateur. Rapport LADL, Université Paris 7, France, 1987.

[638] M. Crochemore and D. Perrin. Critical factorizations of words. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 34:47–52, 1988.

[639] M. Crochemore and D. Perrin. Pattern matching in strings. In V. Cantoni, V. Di Gesu, and S. Levialdi, editors, *Proceedings 4th Conference on Image Analysis and Processing*, number 249 in Lecture Notes in Computer Science, pages 67–79. Plenum Press, 1988.

[640] M. Crochemore and D. Perrin. Pattern matching in strings. Rapport LITP 88.5, Universités Paris 6 et 7, France, 1988.

[641] M. Crochemore and D. Perrin. Two-way string-matching. Rapport LITP 89-8, Universités Paris 6 et 7, France, 1989.

[642] M. Crochemore and D. Perrin. Two-way string-matching. *J. Assoc. Comput. Mach.*, 38(3):651–675, 1991.

[643] M. Crochemore and W. Rytter. Parallel computations on strings and arrays. In C. Choffrut and T. Lengauer, editors, *Proceedings of the 7th Annual Symposium on Theoretical Aspects of Computer Science*, number 415 in Lecture Notes in Computer Science, pages 109–125, Rouen, France, 1990. Springer-Verlag, Berlin.

[644] M. Crochemore and W. Rytter. Parallel construction of minimal suffix and factor automata. *Inf. Process. Lett.*, 35(3):121–128, 1990.

[645] M. Crochemore and W. Rytter. Parallel construction of minimal suffix and factor automata. In B. Rovan, editor, *Proceedings of the 15th Symposium on Mathematical Foundations of Computer Science*, number 452 in Lecture Notes in Computer Science, pages 217–223, Banská Bystrica, Czechoslovakia, 1990. Springer-Verlag, Berlin.

[646] M. Crochemore and W. Rytter. Testing square-freeness of strings by an optimal parallel algorithm in logarithmic time. Rapport LITP 90-63, Universités Paris 6 et 7, France, 1990.

[647] M. Crochemore and W. Rytter. Cubes, squares and time-space efficient string-searching. Rapport LITP 91-47, Universités Paris 6 et 7, France, 1991.

[648] M. Crochemore and W. Rytter. Efficient parallel algorithms to test square-freeness and factorize strings. *Inf. Process. Lett.*, 38(2):57–60, 1991.

[649] M. Crochemore and W. Rytter. Usefulness of the karp-miller-rosenberg algorithm in parallel computations on strings and arrays. *Theor. Comput. Sci.*, 88(1):59–82, 1991.

[650] M. Crochemore and W. Rytter. Note on two-dimensional pattern-matching by optimal parallel algorithms. In A. Nakamura, M. Nivat, A. Saoudi, P. S. P. Wang, and K. Inoue, editors, *Proceedings of the 2nd International Conference on Parallel Image Analysis*, pages 100–112, Ube, Japan, 1992. Springer-Verlag, Berlin.

[651] M. Crochemore and W. Rytter. Periodic prefix in strings. In R. Capocelli, A. De Santis, and U. Vaccaro, editors, *Proceedings Sequences II*, pages 153–165. Springer-Verlag, Berlin, 1993.

[652] M. Crochemore and W. Rytter. On two-dimensional pattern-matching by optimal parallel algorithms. *Theor. Comput. Sci.*, 132(2):403–414, 1994.

[653] M. Crochemore and W. Rytter. *Text algorithms.* Oxford University Press, 1994.

[654] M. Crochemore and W. Rytter. On linear-time alphabet-independent 2-dimensional pattern matching. In R.A. Baeza-Yates, E. Goles, and P.V. Poblete, editors, *Proceedings of the 2nd Latin American Symposium*, number 911 in Lecture Notes in Computer Science, pages 220–229, Valparaíso, Chile, 1995. Springer-Verlag, Berlin.

[655] M. Crochemore and W. Rytter. Squares, cubes and time-space efficient string-searching. *Algorithmica*, 13(5):405–425, 1995.

[656] M. Crochemore and R. Vérin. Direct construction of compact directed acyclic word graphs. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 116–129, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[657] M. Crochemore and R. Vérin. On compact directed acyclic word graphs. In J. Mycielski, G. Rozenberg, and A. Salomaa, editors, *Structures in Logic and Computer Science*, number 1261 in Lecture Notes in Computer Science, pages 192–211. Springer-Verlag, Berlin, 1997.

[658] M. Crochemore and M. Zipstein. Transducteurs arithmétiques. Rapport LITP 89.12, Universités Paris 6 et 7, 1989.

[659] Maxime Crochemore and Laura Giambruno. On-line construction of a small automaton for a finite set of words. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 15–28, Czech Technical University in Prague, Czech Republic, 2009.

[660] M. Cryan, L. A. Goldberg, and C. A. Phillips. Approximation algorithms for the fixed-topology phylogenetic number problem. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 130–149, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[661] K. Culik II and J. Karhumäki. Systems of equations over a free monoid and ehrenfeucht's conjecture. *Discret. Math.*, 43:139–153, 1983.

[662] P. Cull and J. Holloway. A divide and conquer approach to approximate string matching. Report TR-91-50-1, Department of Computer Science, Oregon State University, 1991.

[663] P. Cull, J. Holloway, and J. Cavener. Walking tree heuristics for string matching and gene location. Technical Report 97-20-02, Oregon State University, 1997.

[664] P. Cull and J. L. Holloway. Reconstructing sequences from shotgun data. In *Sequences: Combinatorics, Compression, Security, and Transmission.* Springer-Verlag, Berlin, 1991.

[665] L. J. Cummings. Overlapping substrings and Thue's problem. In *Proceedings of the 3rd Caribbean Conference on Combinatorics and Computing*, pages 99–109, University of the West Indies, Cave Hill, 1981.

[666] L. J. Cummings. Strongly qth power-free strings. *Ann. Discret. Math.*, 17:247–252, 1983.

[667] L. J. Cummings and W. F. Smyth. Weak repetitions in strings. *J. Comb. Math. Comb. Comp.*, 24:33–48, 1997.

[668] A. Czumaj, P. Ferragina, L. Gąsieniec, S. Muthukrishnan, and J. L. Träff. The architecture of a software library for string processing. In *Proceedings of the Workshop on Algorithm Engineering*, pages 166–176, Venice, Italy, 1997.

[669] A. Czumaj and L. Gąsieniec. On the complexity of determining the period of a string. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 412–422, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[670] A. Czumaj, L. Gąsieniec, M. Piotrow, and W. Rytter. Parallel and sequentiel approximation of shortest superstrings. In E. M. Schmidt and S. Skyum, editors, *Proceedings of the 4th Scandinavian Workshop on Algorithm Theory*, number 824 in Lecture Notes in Computer Science, pages 95–106, Aarhus, Denmark, 1994. Springer-Verlag, Berlin.

[671] A. Czumaj, L. Gąsieniec, M. Piotrow, and W. Rytter. Sequential and parallel approximation of shortest superstrings. *J. Algorithms*, 23(1):74–100, 1997.

[672] A. Daaboul. Local prediction for lossless image compression. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 44–50, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[673] B. Dahanayake. Generalized repetitions: A new approach. M. Sc. Thesis, Department of Computer Science and Systems, McMaster University, Canada, 1998.

[674] F. Damerau. A technique for computer detection and correction of spelling errors. *Commun. ACM*, 7(3):171–176, 1994.

[675] V. Dančík. *Expected length of longest common subsequences.* Ph. D. Thesis, Computer Science Departement, University of Warwick, UK, 1994.

[676] V. Dančík. Common subsequences and supersequences and their expected length. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 55–63, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[677] V. Dančík and M. Paterson. Upper bounds for the expected length of a longest common subsequence of two binary sequences. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 669–678, Caen, France, 1994. Springer-Verlag, Berlin.

[678] O. Danvy. Semantics-directed compilation of nonlinear patterns. *Inf. Process. Lett.*, 37(6):315–322, 1991.

[679] T. Darabont. Approximate string matching with k mismatches. Master's thesis, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic, 1992.

[680] G. Das, R. Fleischer, L. Gąsieniec, D. Gunopulos, and J. Kärkkäinen. Episode matching. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 12–27, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[681] B. J. D'Auriol and V. C. Bhavsar. Evaluation of the multi-transputer applications of the weighted Levenshtein distance computation. In *Proceedings of the Int. Conf. Parall. Comp. Transput. Appl.*, 1992.

[682] L. Davidson. Retrieval of misspelled names in an airlines passengerrecord system. *Commun. ACM*, 5(3):169–171, 1962.

[683] D.J.M. Davies. String searching in text editors. *Softw. Pract. Exp.*, 12(8):709–717, 1982.

[684] G. Davies and S. Bowsher. Algorithms for pattern matching. *Softw. Pract. Exp.*, 16(6):575–601, 1986.

[685] J. W. Daykin, C. S. Iliopoulos, and W. F. Smyth. Parallel RAM algorithms for factorizing words. *Theor. Comput. Sci.*, 127(1):53–67, 1994.

[686] Sergio De Agostino. Compressing bi-level images by block matching on a tree architecture. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 137–145, Czech Technical University in Prague, Czech Republic, 2009.

[687] N. F. de Almeida, Jr and V. C. Barbosa. A string-matching algorithm for the CREW PRAM. *Inf. Process. Lett.*, 47(5):257–259, 1993.

[688] B. De Iaco and F. Luccio. Finding all the palindromes of a string in linear time and space. Report TR 3/91, Department of Informatics, University of Pisa, Italy, 1991.

[689] B. De Iaco and F. Luccio. Finding all the palindromes in a binary tree in linear time and space. *Inf. Process. Lett.*, 44(2):73–77, 1992.

[690] R. de la Briandais. File searching using variable length keys. In *Proceedings of the Western Joint Computer Conference*, pages 295–298, 1959.

[691] V.L. Strube de Lima, P.C. Abrahao, and I. Paraboni. Approaching the dictionary in the implementation of a natural language processing system: toward a distributed structure. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 166–178, Valparaiso, Chile, 1997. Carleton University Press.

[692] A. de Luca. A combinatorial property of the Fibonacci word. *Inf. Process. Lett.*, 12(4):193–195, 1981.

[693] A. de Luca. A division property of the Fibonacci word. Rapport LITP 95/16, Universités Paris 6 et 7, France, 1995.

[694] A. de Luca. A division property of the Fibonacci word. *Inf. Process. Lett.*, 54(6):307–312, 1995.

[695] A. de Luca. On the combinatorics of finite words. *Theor. Comput. Sci.*, 218(1):13–39, 1999.

[696] A. de Luca and F. Mignosi. On some combinatorial properties of Sturmian words. Report 93/18, Dipartemento di Matematica, Università di Roma "La Sapienza", Italy, 1993.

[697] A. de Luca and F. Mignosi. On some combinatorial properties of Sturmian words. *Theor. Comput. Sci.*, 136(2):361–385, 1994.

[698] A. de Luca and L. Mione. On bispecial factors of the Thue-Morse word. *Inf. Process. Lett.*, 49(4):179–183, 1994.

[699] A. de Luca and S. Varrichio. Some combinatorial properties of the Thue-Morse sequences and a problem in semigroups. *Theor. Comput. Sci.*, 63(3):333–348, 1989.

[700] A. de Luca and S. Varrichio. Combinatorial properties of uniformly recurrent words and an application to semigroups. *Int. J. Algebr. Comput.*, 1:227–245, 1991.

[701] Satoshi Deguchi, Fumihito Higashijima, Hideo Bannai, Shunsuke Inenaga, and Masayuki Takeda. Parameterized suffix arrays for binary strings. In Jan Holub and Jan Žd̆árek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 84–94, Czech Technical University in Prague, Czech Republic, 2008.

[702] P. L. Deininger, D. J. Jolly, C. M. Rubin, T. Friedmann, and C. W. Schmid. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.*, 151:17–33, 1981.

[703] J. Deken. Some limit results for longest common subsequences. *J. Appl. Probab*, 12:306–315, 1975.

[704] A. Denise. Génération aléatoire et uniforme de mots. In A. Barlotti, M. Delest, and R. Pinzani, editors, *Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics*, pages 153–164, Firenze, Italy, 1993.

[705] A. Dermouche. Algorithms for pattern matching with and without errors. Report IR-1993-1, Computing Science Department, University of Glasgow, 1993.

[706] A. Dermouche. A fast algorithm for string matching with mismatches. *Inf. Process. Lett.*, 55(2):105–110, 1995.

[707] A. S. Deshpande, D. S. Richards, and W. R. Pearson. A platform for biological sequence comparison on parallel computers. *Comput. Appl. Biosci.*, 7(2):237–247, 1991.

[708] J. Devereux, P. Haeberli, and D. Smithies. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, 12:387–395, 1984.

[709] L. Devroye, W. Szpankoswki, and B. Rais. A note of the height of suffix trees. *SIAM J. Comput.*, 21(1):48–53, 1992.

[710] T. G. Dewey. A sequence alignment algorithm with an arbitrary gap penalty function. *J. Comput. Bio.*, 8(2):177–190, 2001.

[711] R. Dixon and T. Martin, editors. *Automatic speech and speaker recognition.* IEEE Press, 1979.

[712] O. C. Dogaru. Algorithm of straight string search. In *Proceedings of the 9th Romanian Symposium on Computer Science*, pages 172–177, Iasi, Romania, 1993.

[713] O. C. Dogaru. On the all occurrences of a word in a text. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 51–57, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[714] O. C. Dogaru. On the first occurrence of a pattern in a text. In *Proceedings of the 32nd International Conference on Modelling and Simulation of Systems*, volume 2, pages 45–50, Bystrice pod Hostinem, Czech Republic, 1998.

[715] P. Dömösi, S. Horváth, and M. Ito. Formal languages and primitive words. *Publ. Math. Debr.*, 42(3-4):315–321, 1993.

[716] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, and M. Katsura. Formal languages consisting of primitive words. In Z. Ésik, editor, *Proceedings Fundamental of Computation Theory*, number 710 in Lecture Notes in Computer Science, pages 194–203, Szeged, Hungary, 1993. Springer-Verlag, Berlin.

[717] X. Doubray. Palindromes in the Fibonacci word. Rapport LAMIFA 94-11, Université d'Amiens, France, 1994.

[718] X. Doubray. Palindromes in the Fibonacci word. *Inf. Process. Lett.*, 55(4):217–221, 1995.

[719] P. J. Downey, R. Sethi, and R. E. Tarjan. Variations on the common subexpression problem. *J. Assoc. Comput. Mach.*, 27(4):758–771, 1980.

[720] M. W. Du and S. C. Chang. An approach to designing very fast approximate string matching algorithms. *IEEE Trans. Knowl. Data Eng.*, 6(4):620–633, 1994.

[721] M. Dubiner, Z. Galil, and E. Magen. Faster tree pattern matching. In *Proceedings of the 31st IEEE Annual Symposium on Foundations of Computer Science*, pages 145–150, Saint Louis, MO, 1990. IEEE Computer Society Press.

[722] M. Dubiner, Z. Galil, and E. Magen. Faster tree pattern matching. *J. Assoc. Comput. Mach.*, 41(2):205–213, 1994.

[723] C. Duboc. On some equations in free partially commutative monoids. *Theor. Comput. Sci.*, 46(2–3):159–174, 1986.

[724] S. Dulucq and D. Gouyou-Beauchamps. Sur les facteurs des suites de Sturm. *Theor. Comput. Sci.*, 71(3):381–400, 1990.

[725] J.A. Dundan III. Implementing dynamic minimal-prefix tries. *Softw. Pract. Exp.*, 21(10):1027–1040, 1991.

[726] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. *Biological sequence analysis probabilistic models of proteins and nucleic acids.* Cambridge University Press, 1998.

[727] A. M. Duval and W. F. Smyth. Covering a circular string with substrings of fixed length. *J. Foudations Comp. Sc..*, 7(1):87–93, 1996.

[728] J.-P. Duval. Périodes et répétitions des mots du monoïde libre. *Theor. Comput. Sci.*, 9(1):17–26, 1979.

[729] J.-P. Duval. *Contribution à la combinatoire du monoïde libre.* Thèse, Université de Rouen, 1980.

[730] J.-P. Duval. Mots de Lyndon et périodicité. *RAIRO Inform. Théor.*, 14(2):181–191, 1980.

[731] J.-P. Duval. Lexicographically maximal and minimal segments of a word, linear string searching algorithms. Rapport LITP 81-23, Universités Paris 6 et 7, France, 1981.

[732] J.-P. Duval. Relationship between the period of a finite word and the length of its unbordered segments. Rapport LITP 81-22, Universités Paris 6 et 7, France, 1981.

[733] J.-P. Duval. A remark on the Knuth-Morris-Pratt's string searching algorithm. Rapport LITP 81-7, Universités Paris 6 et 7, France, 1981.

[734] J.-P. Duval. Relationship between the period of a finite word and the length of its unbordered segments. *Discret. Math.*, 40:31–44, 1982.

[735] J.-P. Duval. Factorizing words over an ordered alphabet. *J. Algorithms*, 4(4):363–381, 1983.

[736] J.-P. Duval. Génération d'une section des classes de conjugaison et arbre des mots de Lyndon de longueur bornée. Rapport LITP 88.20, Universités Paris 6 et 7, France, 1988.

[737] J.-P. Duval. Génération d'une section des classes de conjugaison et arbre des mots de Lyndon de longueur bornée. *Theor. Comput. Sci.*, 60(3):255–283, 1988.

[738] J.-P. Duval. Périodes locales et propagation de périodes dans un mot. Rapport LIR 97.05, Université de Rouen, France, 1997.

[739] J.-P. Duval. Périodes locales et propagation de périodes dans un mot. *Theor. Comput. Sci.*, 204(1–2):87–98, 1998.

[740] J.-P. Duval, R. Kolpakov, G. Kucherov, T. Lecroq, and A. Lefebvre. Linear-time computation of local periods. In *Proceedings of the 28th Symposium on Mathematical Foundations of Computer Science*, Bratislava, Slovak Republic, 2003. Accepted.

[741] J.-P. Duval, T. Lecroq, and A. Lefebvre. Border array on bounded alphabet. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference'02*, pages 28–35, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[742] J.-P. Duval and A. Lefebvre. Words over an ordered alphabet and suffix permutations. *RAIRO Inform. Théor. Appl.*, 36:249–259, 2002.

[743] J.-P. Duval, F. Mignosi, and A. Restivo. Recurrence and periodicity in infinite words from local periods. *Theor. Comput. Sci.*, 262(1–2):269–?, 2001.

[744] J.-P. Duval and L. Mouchard. Sequences obtained from a series of permutations of borders and ultimately periodic words. *J. Autom. Lang. Comb.*, 4(3):205–211, 1999.

[745] J. Dvorský and V. Snášel. Word-based compression method with direct access. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 24–31, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[746] J. Earley. A efficient context-free parsing algorithm. *Commun. ACM*, 13(2):94–102, 1970.

[747] S. Edelkamp. Dictionary automaton in optimal space. report00129, Institut f"ur Informatik, Universit"at Freiburg, 1999.

[748] E. W. Edmiston, N. G. Core, J. H. Saltz, and R. M. Smith. Parallel processing of biological sequence comparison algorithms. *Int. J. Parallel Program.*, 17(3):259–275, 1988.

[749] E. W. Edmiston and R. A. Wagner. Parallelization of the dynamic programming algorithm for comparison of sequences. In S. K. Sahni, editor, *Proceedings of the 1987 International Conference on Parallel Processing*, pages 79–80, Pennsylvania State University Press, PA, 1987.

[750] R. E. Eggen and C. Kurak. Pattern matching in fewer comparisons. In *Proceedings IEEE SOUTHEASTCON*, pages 410–412, 1992.

[751] A. Ehrenfeucht and D. Haussler. A new distance metric on strings computable in linear time. *Discret. Appl. Math.*, 20:191–203, 1988.

[752] T. Eilam-Tzoreff and U. Vishkin. Matching patterns in a string subject to multi-linear transformations. *Theor. Comput. Sci.*, 60(3):231–254, 1988.

[753] M. J. Eisler. A Research Report on the Design of an Approximate String Matching Device using the LCS Device Approach. Master's thesis, Univ. Central Florida, 1985.

[754] N. El-Mabrouk. Genome rearrangement by reversals and insertions/deletions of contiguous segments. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[755] N. El-Mabrouk. Sorting signed permutations by reversals and insertions/deletions of contiguous segments. *J. Discret. Algorithms*, 1(1):105–121, 2000.

[756] N. El-Mabrouk and M. Crochemore. Boyer-Moore strategy to efficient approximate string matching. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 24–38, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[757] N. El-Mabrouk, J. Nadeau, and D. Sankoff. Genome halving. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 235–250, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[758] N. El-Mabrouk and D. Sankoff. Hybridization and genome rearrangement. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 78–87, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[759] J. Ellis, F. Ruskey, and J. Sawada. Euclidean strings. In L. Brankovic and J. Ryan, editors, *Proceedings of the 11th Australasian Workshop On Combinatorial Algorithms*, pages 87–92, Hunter Valley, Australia, 2000.

[760] D. Eppstein. *Efficient Algorithms for Sequence Analysis with Concave and Convexe Gap Costs*. PhD thesis, Computer Science Department, Columbia University, New York, 1989.

[761] D. Eppstein. Sequence comparison with mixed convex and concave costs. *J. Algorithms*, 11(1):85–101, 1990.

[762] D. Eppstein, Z. Galil, and R. Giancarlo. Speeding up dynamic programming. In *Proceedings of the 29th IEEE Annual Symposium on Foundations of Computer Science*, pages 488–496, White Plains, NY, 1988. IEEE Computer Society Press.

[763] D. Eppstein, Z. Galil, R. Giancarlo, and G. F. Italiano. Sparse dynamic programming i: linear cost functions. *J. Assoc. Comput. Mach.*, 39(3):519–545, 1992.

[764] D. Eppstein, Z. Galil, R. Giancarlo, and G. F. Italiano. Sparse dynamic programming ii: convex and concave cost functions. *J. Assoc. Comput. Mach.*, 39(3):546–567, 1992.

[765] D. Eppstein, Z. Galil, R. Giancarlo, and G.F. Italiano. Sparse dynamic programming. In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 513–522, San Francisco, CA, 1990.

[766] B. W. Erickson and P. H. Sellers. *Recognition of patterns in genetic sequences*, pages 55–91. Addison-Wesley, Reading, MA, 1983.

[767] T. Etzold and P. Argos. SRS – an indexing and retrieval tool for flat file data libraries. *Comput. Appl. Biosci.*, 9:49–57, 1993.

[768] D. J. Evans and S. Ghanemi. Parallel string matching algorithm. *Kybern.*, 17(3):32–34, 1988.

[769] P. Evans. Finding common subsequences with arcs and pseudoknots. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 270–280, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[770] S. Even and M. Rodeh. Economical encoding of commas between strings. *Commun. ACM*, 21(4):315–317, 1978.

[771] S. D. Eyono Obono, P. Goralcik, and M. Maksimenko. Efficient solving of the word equations in one variable. In I. Prívara, B. Rovan, and P. Ružička, editors, *Proceedings 19th Symposium on Mathematical Foundations of Computer Science*, number 841 in Lecture Notes in Computer Science, pages 336–341, Koǒsice, Slovakia, 1994. Springer-Verlag, Berlin.

[772] C. Faloutsos. Acces methods for text. *ACM Comp. Surv.*, 17(1):49–74, 1985.

[773] C.-M. Fan, H. J. Shyr, and S. S. Yu. d-words and d-languages. *Acta Inform.*, 35(8):709–727, 1998.

[774] J. J. Fan and K. Y. Su. An efficient algorithm for matching multiple patterns. *IEEE Trans. Knowl. Data Eng.*, 5(2):339–351, 1993.

[775] J. J. Fan and K. Y. Su. An efficient algorithm for matching multiple patterns. In J. Aoe, editor, *Computer Algorithms: String Pattern Matching Strategies*, pages 91–103. IEEE Computer Society Press, 1994.

[776] T. I. Fan. Optimal matching of deformed patterns with positional influence. *Inf. Sci.*, 41:259–280, 1987.

[777] M. Farach. Optimal suffix tree construction with large alphabets. Technical Report 96–48, DIMACS, 1996.

[778] M. Farach. Optimal suffix tree construction with large alphabets. In *Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science*, pages 137–143, Miami Beach, FL, 1997.

[779] M. Farach, S. Kannan, and T. Warnow. A robust model for finding optimal evolutionary trees. *Algorithmica*, 13(1/2):155–179, 1995.

[780] M. Farach and S. Muthukrishnan. Optimal logarithmic time randomized suffix tree construction. In F. Meyer auf der Heide and B. Monien, editors, *Proceedings of the 23rd International Colloquium on Automata,*

Languages and Programming, volume 1099 of *Lecture Notes in Computer Science*, pages 550–561, Paderborn, Germany, 1996. Springer-Verlag, Berlin.

[781] M. Farach and S. Muthukrishnan. Perfect hashing for strings: Formalization and algorithms. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 130–140, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[782] M. Farach and M. Thorup. String matching in Lempel-Ziv compressed strings. In *Proceedings of the 27th ACM Symposium on the Theory of Computing*, pages 703–712, Las Vegas, Nevada, 1995.

[783] M. Farach and M. Thorup. String matching in Lempel-Ziv compressed strings. *Algorithmica*, 20(4):388–404, 1998.

[784] Simone Faro and Thierry Lecroq. Efficient variants of the backward-oracle-matching algorithm. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 146–160, Czech Technical University in Prague, Czech Republic, 2008.

[785] R. D. Faulk. An inductive approach to language translation. *Commun. ACM*, 7(11):647–653, 1964.

[786] D. V. Faulkner and J. Jurka. Multiple aligned sequence editor (MASE). *Trends Biochem. Sci.*, 13:321–322, 1988.

[787] M. Fellows, M. T. Hallett, C. Korostensky, and U. Stege. Analogs and duals of the MAST problem for sequences and trees. In G. Bilardi, G. F. Italiano, A. Pietracaprina, and G. Pu cci, editors, *Proceedings of the 6th Annual European Symposium*, number 1461 in Lecture Notes in Computer Science, pages 103–114, Venice, Italy, 1998. Springer-Verlag, Berlin.

[788] J. Felsenstein, S. Sawyer, and R. Kochin. An efficient method for matching nucleic acid sequences. *Nucleic Acids Res.*, 10:133–139, 1982.

[789] D.-F. Feng and R. F. Doolittle. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.*, 25:351–360, 1987.

[790] D.-F. Feng and R. F. Doolittle. Progressive alignment and phylogenetic tree construction of protein sequences. *Methods Enzymol.*, 188:375–402, 1990.

[791] D.-F. Feng, M. S. Johnson, and R. F. Doolittle. Aligning amino acid sequences: Comparison of commonly used methods. *J. Mol. Evol.*, 21:112–125, 1985.

[792] S. Ferenczi and Z. Kása. Complexity for finite factors of infinite sequences. *Theor. Comput. Sci.*, 218(1):177–195, 1999.

[793] D. Fernández-Baca, T. Seppäläinen, and G. Slutzki. Parametric multiple sequence alignment and phylogeny construction. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 69–83, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[794] P. Ferragina. Incremental text editing: A new data structure. In J. van Leeuwen, editor, *Proceedings of the 2nd Annual European Symposium*, number 855 in Lecture Notes in Computer Science, pages 495–507, Utrecht, The Netherlands, 1994. Springer-Verlag, Berlin.

[795] P. Ferragina. *Dynamic data structures for string matching problems*. PhD thesis, Computer Science Dept., Università di Pisa, 1997.

[796] P. Ferragina. Dynamic text indexing under string updates. *J. Algorithms*, 22(2):296–328, 1997.

[797] P. Ferragina and R. Grossi. Fast incremental text editing. In *Proceedings of the 6th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 531–540, San Francisco, CA, 1995.

[798] P. Ferragina and R. Grossi. A fully-dynamic data structure for external substring search. In *Proceedings of the 27th ACM Symposium on the Theory of Computing*, Las Vegas, Nevada, 1995. ACM Press.

[799] P. Ferragina and R. Grossi. Optimal on-line search and sublinear time update in string matching. In *Proceedings of the 36th IEEE Annual Symposium on Foundations of Computer Science*, pages 604–612, Milwaukee, Wisconsin, 1995. IEEE Computer Society Press.

[800] P. Ferragina and R. Grossi. Fast string searching in secondary storage: Theoretical developments and experimental results. In *Proceedings of the 7th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 373–382, Atlanta, Georgia, 1996.

[801] P. Ferragina and R. Grossi. Optimal on-line search and sublinear time update in string matching. *SIAM J. Comput.*, 27(3):713–736, 1998.

[802] P. Ferragina and R. Grossi. Improved dynamic text indexing. *J. Algorithms*, 31(2):291–319, 1999.

[803] P. Ferragina and R. Grossi. The string B-tree: A new data structure for string search in external memory and its applications. *J. Assoc. Comput. Mach.*, 46:236–280, 1999.

[804] P. Ferragina, R. Grossi, and M. Montangero. A note on updating suffix tree labels. In *CIAC: Italian Conference on Algorithms and Complexity*, 1997.

[805] P. Ferragina, R. Grossi, and M. Montangero. A note on updating suffix tree labels. *Theor. Comput. Sci.*, 201(1/2):249–262, 1998.

[806] P. Ferragina and F. Luccio. Dynamic dictionary matching in external memory. *Inf. Comput.*, 146(2):85–99, 1998.

[807] P. Ferragina and F. Luccio. String search in coarse-grained parallel computers. *Algorithmica*, 24(3/4):177–194, 1999.

[808] P. Ferragina, S. Muthukrishnan, and de Berg M. Multi-method dispatching: A geometric approach with applications to string matching problems. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 483–491, 1999.

[809] V. Ferretti, J. H. Nadeau, and D. Sankoff. Original synteny. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 159–167, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[810] J. W. Fickett. Fast optimal alignment. *Nucleic Acids Res.*, 12(1):175–179, 1984.

[811] W. A. Finamore, R. D. de Azevedo, and M. da Silva Pinho. On procedures for multiple-string match with respect to two sets. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '99*, pages 29–39, Czech Technical University, Prague, Czech Republic, 1999. Collaborative Report DC–99–05.

[812] N. J. Fine and H. S. Wilf. Uniqueness theorems for periodic functions. *Proc. Amer. Math. Soc.*, 16:109–114, 1965.

[813] D. Fischer, R. Norel, R. Nussinov, and H. J. Wolfson. 3-D docking of protein molecules. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 20–34, Padova, Italy, 1993. Springer-Verlag, Berlin.

[814] D. Fischer, R. Nussinov, and H. J. Wolfson. 3-d structure matching in protein molecules. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 136–150, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[815] M. J. Fischer and M. Paterson. String matching and other products. In R. M. Karp, editor, *Proceedings SIAM-AMS Complexity of Computation*, pages 113–125, Providence, RI, 1974.

[816] V. A. Fischetti, G. M. Landau, J. P. Schmidt, and P. H. Sellers. Identifying periodic occurrences of a template with applications to protein struture. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 111–120, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[817] V. A. Fischetti, G. M. Landau, P. H. Sellers, and J. P. Schmidt. Identifying periodic occurrences of a template with applications to protein struture. *Inf. Process. Lett.*, 45(1):11–18, 1993.

[818] W. M. Fitch and T. F. Smith. Optimal sequence alignments. *Proc. Natl. Acad. Sci. U.S.A.*, 80:1382–1386, 1983.

[819] P. Flajolet and C. Puech. Partial match retrieval of multidimensional data. *J. Assoc. Comput. Mach.*, 33(2):371–407, 1986.

[820] A. Floratos, I. Rigoutsos, L. Parida, G. Stolovitzky, and Y. Gao. Sequence homology detection through large-scale pattern discovery. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 164–169, 1999.

[821] D. E. Foulser. *On random strings and sequence comparisons*. PhD thesis, Computer Science Department, Stanford University, 1986.

[822] D. E. Foulser. A linear time algorithm for DNA sequencing. Report YALEU/DCS/RR-812, Department of Computer Science, Yale University, 1990.

[823] A. Fraenkel. Recent results and questions in combinatorial game complexities. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 124–146, Perth, WA, Australia, 1998.

[824] A. S. Fraenkel and S. T. Klein. Complexity aspects of guessing prefix codes. *Algorithmica*, 12(4/5):409–419, 1994.

[825] A. S. Fraenkel and J. Simpson. How many squares must a binary sequence contain? *Electronic Journal of Combinatorics*, 2:R2, 1995.

[826] A. S. Fraenkel and J. Simpson. How many squares can a string contain? *J. Combin. Theory Ser. A*, 82:112–120, 1998.

[827] A. S. Fraenkel and J. Simpson. The exact number of squares in Fibonacci words. *Theor. Comput. Sci.*, 218(1):95–106, 1999.

[828] A. S. Fraenkel, J. Simpson, and M. Paterson. On weak circular squares in binary words. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264

in Lecture Notes in Computer Science, pages 76–82, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[829] F. Franěk, A. Karaman, and W. F. Smyth. Repetitions in sturmian words. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 107–116, Perth, WA, Australia, 1998.

[830] F. Franěk, A. Karaman, and W. F. Smyth. Repetitions in sturmian words. *Theor. Comput. Sci.*, 249(2):289–303, 2000.

[831] F. Franěk, W. Lu, P. J. Ryan, W. F. Smyth, Y. Sun, and L. Yang. Verifying a border array in linear time. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 26–33, Perth, WA, Australia, 1999.

[832] F. Franěk, W. Lu, and W. F. Smyth. Repetitions in two-pattern strings. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 32–48, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[833] F. Franěk, W. F. Smyth, and X. Xiao. A note on Crochemore's repetitions algorithm a fast space-efficient approach. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 36–43, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[834] Frantisek Franek and Mei Jiang. Crochemore's repetitions algorithm revisited – computing runs. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 214–224, Czech Technical University in Prague, Czech Republic, 2009.

[835] C. B. Fraser. *Subsequences and supersequences of strings.* Ph. D. Thesis, University of Glasgow, UK, 1995.

[836] C. B. Fraser. Consistent subsequences and supersequences. *Theor. Comput. Sci.*, 165(2):233–246, 1996.

[837] C. B. Fraser and R. W. Irving. Maximal common subsequences and minimal common supersequences. *Inf. Comput.*, 124(2):145–153, 1996.

[838] E. Fredkin. *Trie memory.* Information Memorandum, Bolt Beranek and NewMan Inc., Cambridge, MA, 1959.

[839] E. Fredkin. Trie memory. *Commun. ACM*, 3(9):490–499, 1960.

[840] M. L. Fredman. On computing the length of longest increasing subsequences. *Discret. Math.*, 11:29–35, 1975.

[841] K. Fredriksson and E. Ukkonen. Rotation invariant filters for multidimensional string matching and orientation search. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 118–125, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[842] J. French, A. Powell, and E. Schulman. Applications of approximate word matching in information retrieval. In *Proceedings ACM CIKM'97*, pages 9–15, 1997.

[843] H. Freund and R. Wolter. Evolution of bit strings: Some preliminary results. *Complex Syst.*, 5(3):279–298, 1991.

[844] T. Friedemann. Alignment of multiple DNA and protein sequence data. *Comput. Appl. Biosci.*, 4:213–214, 1988.

[845] A. M. Frieze and W. Szpankowski. Greedy algorithms for the shortest common superstring that are asymptotically optimal. *Algorithmica*, 21(1):21–36, 1998.

[846] J. Fu. Pattern matching in directed graphs. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 64–77, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[847] K. J. Fu. Approximate pattern matching in directed graphs. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 373–383, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[848] I. Fudos, E. Pitoura, and W. Szpankowski. On pattern occurrences in a random text. *Inf. Process. Lett.*, 57:307–312, 1996.

[849] Z. Galil. Real time algorithms for string matching and palindrome recognition. In *Proceedings of the 8th ACM Symposium on the Theory of Computing*, pages 161–173, Hershey, PA, 1976. ACM Press.

[850] Z. Galil. Two fast simulations which imply some fast string matching and palindrome recognition algorithms. *Inf. Process. Lett.*, 4(4):85–87, 1976.

[851] Z. Galil. Some open problems in the theory of computations as questions about two-way deterministic pushdown automaton languages. *Math. Syst. Theory*, 10:211–228, 1977.

[852] Z. Galil. Palindrome recognition in real time by a multitape turing machine. *J. Comput. Syst. Sci.*, 16(2):140–157, 1978.

[853] Z. Galil. Palindrome recognition in real time by a multitape turing machine. In G. Ausiello and C. Böhm, editors, *Proceedings of the 5th International Colloquium on Automata, Languages and Programming*, number 62 in Lecture Notes in Computer Science, pages 241–250, Udine, Italy, 1978. Springer-Verlag, Berlin.

[854] Z. Galil. On improving the worst case running time of the Boyer-Moore string searching algorithm. *Commun. ACM*, 22(9):505–508, 1979.

[855] Z. Galil. String matching in real time. *J. Assoc. Comput. Mach.*, 28(1):134–149, 1981.

[856] Z. Galil. Optimal parallel algorithms for string matching. Report CUCS-74-83, Computer Science Department, Columbia University, NY, 1983.

[857] Z. Galil. Optimal parallel algorithms for string matching. In *Proceedings of the 16th ACM Symposium on the Theory of Computing*, pages 240–248, Washington, D.C., 1984. ACM Press.

[858] Z. Galil. Open problems in stringology. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, pages 1–8. Springer-Verlag, Berlin, 1985.

[859] Z. Galil. Optimal parallel algorithms for string matching. *Inf. Control*, 67(1–3):144–157, 1985.

[860] Z. Galil. Hunting lions in the desert optimally or a constant time optimal parallel string matching algorithms. In *Proceedings of the 24th ACM Symposium on the Theory of Computing*, pages 69–76, Victoria, Canada, 1992. ACM Press.

[861] Z. Galil and R. Giancarlo. Improved string matching with k mismatches. *SIGACT News*, 17(4):52–54, 1986.

[862] Z. Galil and R. Giancarlo. Data structures and algorithms for approximate string matching. Report CUCS-299-87, Computer Science Department, Columbia University, NY, 1987.

[863] Z. Galil and R. Giancarlo. Parallel string matching with k mismatches. *Theor. Comput. Sci.*, 51(1):341–348, 1987.

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

[865] Z. Galil and R. Giancarlo. Speeding up dynamic programming with applications to molecular biology. *Theor. Comput. Sci.*, 64(1):107–118, 1989.

[866] Z. Galil and R. Giancarlo. On the exact complexity of string matching: lower bounds. *SIAM J. Comput.*, 20(6):1008–1020, 1991.

[867] Z. Galil and R. Giancarlo. On the exact complexity of string matching: upper bounds. *SIAM J. Comput.*, 21(3):407–437, 1992.

[868] Z. Galil, R. Giancarlo, and K. Park. Fully dynamic dictionary matching. Report CUCS-043-91, Computer Science Department, Columbia University, NY, 1991.

[869] Z. Galil and K. Park. An improved algorithm for approximate string matching. In G. Ausiello, M. Dezani-Ciancaglini, and S. Ronchi Della Rocca, editors, *Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, number 372 in Lecture Notes in Computer Science, pages 394–404, Stresa, Italy, 1989. Springer-Verlag, Berlin.

[870] Z. Galil and K. Park. An improved algorithm for approximate string matching. Report CUCS-468-89, Computer Science Department, Columbia University, NY, 1989.

[871] Z. Galil and K. Park. An improved algorithm for approximate string matching. *SIAM J. Comput.*, 19(6):989–999, 1990.

[872] Z. Galil and K. Park. A linear-time algorithm for concave one-dimensional dynamic programming. *Inf. Process. Lett.*, 33(6):309–311, 1990.

[873] Z. Galil and K. Park. Dynamic programming with convexity, concavity, and sparsity. *Theor. Comput. Sci.*, 92(1):49–76, 1992.

[874] Z. Galil and K. Park. Truly alphabet-independent two-dimensional pattern matching. In *Proceedings of the 33th IEEE Annual Symposium on Foundations of Computer Science*, pages 247–256, Pittsburgh, PA, 1992. IEEE Computer Society Press.

[875] Z. Galil and K. Park. Alphabet-independent two-dimensional witness computation. *SIAM J. Comput.*, 25(5):907–935, 1996.

[876] Z. Galil and Y. Rabani. On the space complexity of some algorithms for sequence comparison. *Theor. Comput. Sci.*, 95(2):231–244, 1992.

[877] Z. Galil and J. Seiferas. Recognizing certain repetitions and reversals within strings. In *Proceedings of the 17th IEEE Annual Symposium on Foundations of Computer Science*, pages 236–252, Houston, TX, 1976. IEEE Computer Society Press.

[878] Z. Galil and J. Seiferas. Saving space in fast string-matching. In *Proceedings of the 18th IEEE Annual Symposium on Foundations of Computer Science*, pages 179–188, Providence, Rhode Island, 1977. IEEE Computer Society Press.

[879] Z. Galil and J. Seiferas. A linear-time on-line recognition algorithm for "palstars". *J. Assoc. Comput. Mach.*, 25(1):102–111, 1978.

[880] Z. Galil and J. Seiferas. Saving space in fast string-matching. *SIAM J. Comput.*, 9(2):417–438, 1980.

[881] Z. Galil and J. Seiferas. Linear-time string matching using only a fixed number of local storage locations. *Theor. Comput. Sci.*, 13(3):331–336, 1981.

[882] Z. Galil and J. Seiferas. Time-space optimal string matching. In *Proceedings of the 13th ACM Symposium on the Theory of Computing*, pages 106–113, Milwaukee, Wisconsin, 1981.

[883] Z. Galil and J. Seiferas. Time-space optimal string matching. *J. Comput. Syst. Sci.*, 26(3):280–294, 1983.

[884] Z. Galil and I. Yudkiewicz. On-line parallel exact string searching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 2, pages 55–88. Oxford University Press, 1997.

[885] J. Gallant, D. Maier, and J. Storer. On finding minimal length superstring. *J. Comput. Syst. Sci.*, 20(1):50–58, 1980.

[886] Matthias Gallé, Pierre Peterlongo, and François Coste. In-place update of suffix array while recoding words. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 54–67, Czech Technical University in Prague, Czech Republic, 2008.

[887] E. Garel. Minimal separators of two words. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 35–53, Padova, Italy, 1993. Springer-Verlag, Berlin.

[888] E. Garel. Séparateurs dans les mots infinis engendrés par morphismes. *Theor. Comput. Sci.*, 180(1–2):81–113, 1997.

[889] M. R. Garey and D. S. Johnson. *Computers and intractability: a guide to the theory of NP-completeness.* Freeman, New York, NY, 1979.

[890] L. Gąsieniec, P. Indyk, and P. Krysta. External inverse pattern matching. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 90–101, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[891] L. Gąsieniec, J. Jansson, and A. Lingas. Approximation algorithms for hamming clustering problems. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern*

Matching, number 1848 in Lecture Notes in Computer Science, pages 108–118, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[892] L. Gąsieniec, M. Karpinski, W. Plandowski, and W. Rytter. Randomized efficient algorithms for compressed strings: The finger-print approach (extended abstract). In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 39–49, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[893] L. Gąsieniec and K. Park. Work-time optimal parallel prefix matching. In J. van Leeuwen, editor, *Proceedings of the 2nd Annual European Symposium*, number 855 in Lecture Notes in Computer Science, pages 471–482, Utrecht, The Netherlands, 1994. Springer-Verlag, Berlin.

[894] L. Gąsieniec, W. Plandowski, and W. Rytter. Constant-space string matching with smaller number of comparisons: sequential sampling. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 78–89, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[895] L. Gąsieniec, W. Plandowski, and W. Rytter. The zooming method: a recursive approach to time-space efficient string-matching. *Theor. Comput. Sci.*, 147(1–2):19–30, 1995.

[896] L. Gąsieniec and W. Rytter. Searching patterns with short variables. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 145–151, Perth, WA, Australia, 1999.

[897] M. S. Gelfand, E. V. Koonin, and A. A. Mironov. Prediction of transcription regulatory sites in archaea by a comparative genomic approach. *Nucleic Acids Res.*, 28(3):695–705, 2000.

[898] M. S. Gelfand and E.V. Koonin. Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes. *Nucleic Acids Res.*, 25:2430–2439, 1997.

[899] M. S. Gelfand, C. G. Kozhukhin, and P. A. Pevzner. Extendable words in nucleotide sequences. *Comput. Appl. Biosci.*, 8:129–135, 1992.

[900] M. S. Gelfand, A. A. Mironov, and P. A. Pevzner. Spliced alignment: A new approach to gene recognition. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 141–158, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[901] M. Geréb-Graus and M. Li. Three one-way heads cannot do string matching. *J. Comput. Syst. Sci.*, 48(1):1–8, 1994.

[902] D. Giammarresi. Periodicities on trees. *Theor. Comput. Sci.*, 205(1/2):145–181, 1998.

[903] D. Giammarresi and A. Restivo. Recognizable picture languages. In *Proceedings of the 1st International Colloquium on Parallel Image Processing*, pages ?–?, 1991.

[904] D. Giammarresi and A. Restivo. Recognizable picture languages. *Internat. J. Pattern Recognition and Artificial Intelligence*, 6(2–3):241–256, 1992.

[905] D. Giammarresi, A. Restivo, S. Seibert, and W. Thomas. Monadic second-order logic over pictures and recognizability. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 365–375, Caen, France, 1994. Springer-Verlag, Berlin.

[906] R. Giancarlo. *Efficient algorithms on strings*. Ph. D. Thesis, Departement of Computer Science, Columbia University, New-York, 1990.

[907] R. Giancarlo. An index data structure for matrices, with applications to fast two-dimensional pattern matching. In F. Dehne, J.-R. Sack, N. Santoro, and S. Whitesides, editors, *Proceedings of the 3rd Workshop on Algorithms and Data Structures*, number 709 in Lecture Notes in Computer Science, pages 337–348, Montréal, Canada, 1993. Springer-Verlag, Berlin.

[908] R. Giancarlo. The suffix tree of a square matrix with applications. In *Proceedings of the 4th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 402–411, Austin, TX, 1993.

[909] R. Giancarlo. A generalization of the suffix tree to square matrices, with applications. *SIAM J. Comput.*, 24(3):520–562, 1995.

[910] R. Giancarlo. Dynamic programming: special cases. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 7, pages 201–236. Oxford University Press, 1997.

[911] R. Giancarlo and R. Grossi. Parallel construction and query of suffix trees for two-dimensional matrices. In *Proceedings of the 5th ACM Symposium on Parallel Algorithms and Architectures*, pages 86–97, Velen, Germany, 1993. ACM Press.

[912] R. Giancarlo and R. Grossi. Multi-dimensional pattern matching with dimensional wildcards. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 90–101, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[913] R. Giancarlo and R. Grossi. On the construction of classes of suffix trees for square matrices: Algorithms and applications. *Inf. Comput.*, 130(2):151–182, 1996.

[914] R. Giancarlo and R. Grossi. Multi-dimensional pattern matching with dimensional wildcards: Data structures and optimal on-line search algorithms. *J. Algorithms*, 24(2):223–265, 1997.

[915] R. Giancarlo and R. Grossi. Suffix tree data structures for matrices. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 10, pages 293–340. Oxford University Press, 1997.

[916] A. Gibbons and W. Rytter. On the decidability of some problems about rational subsets of free partially commutative monoids. *Theor. Comput. Sci.*, 48(2–3):329–337, 1986.

[917] A. Gibbons and W. Rytter. *Efficient parallel algorithms.* Cambridge University Press, Cambridge, 1988.

[918] A. Gibbons and W. Rytter. Optimal parallel algorithms for dynamic expression evaluation and context-free recognition. *Inf. Comput.*, 81(1):32–45, 1989.

[919] R. Giegerich. Explaining and controlling ambiguity in dynamic programming. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 46–59, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[920] R. Giegerich and S. Kurtz. Suffix trees in the functional programming paradigm. In *Programming Languages and Systems*, pages 225–240, 1994.

[921] R. Giegerich and S. Kurtz. Suffix trees in the functional programming paradigm. In D. Sannella, editor, *Proceedings of the 5th European Symposium on Programming*, number 788 in Lecture Notes in Computer Science, pages 225–240, Edinburgh, U.K., 1994. Springer-Verlag, Berlin.

[922] R. Giegerich and S. Kurtz. A comparison of imperative and purely functional suffix tree constructions. *Sci. Comp. Program.*, 25(2–3):187–218, 1995.

[923] R. Giegerich and S. Kurtz. From ukkonen to mccreight and weiner: A unifying view of linear-time suffix tree construction. *Algorithmica*, 19(3):331–353, 1997.

[924] R. Giegerich, S. Kurtz, F. Hischke, and E. Ohlebusch. A general technique to improve filter algorithms for approximate string matching. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 38–52, Valparaiso, Chile, 1997. Carleton University Press.

[925] R. Giegerich, S. Kurtz, and J. Stoye. Efficient implementation of lazy suffix trees. In J.S. Vitter and C.D. Zaroliagis, editors, *Proceedings of the 3rd Workshop on Algorithm Engineering*, number 1668 in Lecture Notes in Computer Science, pages 30–42, London, UK, 1999. Springer-Verlag, Berlin.

[926] W. Gilbert. Genome sequencing: Creating a new biology for the twenty–first century. *Issues Sci. Tech.*, 3:26–35, 1987.

[927] W. B. Goad and M. I. Kanehisa. Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. *Nucleic Acids Res.*, 10:247–263, 1982.

[928] H. Goeman and M. Clausen. A new practical linear space algorithm for the longest common subsequence problem. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '99*, pages 40–60, Czech Technical University, Prague, Czech Republic, 1999. Collaborative Report DC–99–05.

[929] L. A. Goldberg, P. W. Goldberg, C. A. Philipps, E. Sweedyk, and T. Warnow. Minimizing phylogenetic number to find good evolutionary trees. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 102–127, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[930] L. A. Goldberg, P. W. Goldberg, C. A. Phillip, and G. B. Sorkin. Constructing computer virus phylogenies. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 253–270, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[931] R. N. Goldberg. Minimal string difference encodings. *J. Algorithms*, 3(2):147–156, 1982.

[932] T. Goldberg and U. Zwick. Faster parallel string matching via larger deterministic samples. *J. Algorithms*, 16(2):295–308, 1994.

[933] M. Goldwurm. Random generation of words in an algebraic language in linear binary space. *Inf. Process. Lett.*, 54(4):229–233, 1995.

[934] J. Golic and M. Mihaljevic. A noisy clock-controlled shift register cryptanalysis concept based on sequence comparison approach. In I. Damgard, editor, *Proceedings of the 7th EUROCRYPT*, number 473 in Lecture Notes in Computer Science, pages 487–491, Aarhus, Denmark, 1990. Springer-Verlag, Berlin.

[935] G. Gonnet, R. Baeza-Yates, and T. Snider. New indices for text: PAT trees and PAT arrays. In B. Frakes and R. Baeza-Yates, editors, *Information Retrieval Data Structures & Algorithms*. Prentice-Hall, 1992.

[936] G. H. Gonnet. Unstructured data bases or very efficient text searching. In *Proceedings of the 2nd ACM SIGACT-SIGMOD Symposium on Principles of Database Systems*, pages 117–124, Atlanta, GA, 1983. ACM Press.

[937] G. H. Gonnet. The PAT text searching system. Report, Department of Computer Science, University of Waterloo, Ontario, 1987.

[938] G. H. Gonnet. Efficient searching of text and pictures. Report OED-88-02, University of Waterloo, 1988.

[939] G. H. Gonnet. Efficient two-dimensional searching. In O. Nurmi and E. Ukkonen, editors, *Proceedings of the 3rd Scandinavian Workshop on Algorithm Theory*, number 621 in Lecture Notes in Computer Science, page 317, Helsinki, Finland, 1992. Springer-Verlag, Berlin.

[940] G. H. Gonnet. A tutorial introduction to computation biochemistry using darwin. Technical report, Informatik E.T.H., Zürich, Switzerland, 1992.

[941] G. H. Gonnet and R. A. Baeza-Yates. An analysis of the Karp-Rabin string matching algorithm. *Inf. Process. Lett.*, 34(5):271–274, 1990.

[942] G. H. Gonnet and R. A. Baeza-Yates. *Handbook of Algorithms and Data Structures*. Addison-Wesley, Reading, MA, 1991.

[943] G. H. Gonnet and R. A. Baeza-Yates. *Text algorithms*, chapter 7, pages 251–288. Addison-Wesley, Wokingham, U. K., second edition, 1991.

[944] G. H. Gonnet, M. Cohen, and S. Benner. The all-against-all matching of a major protein sequence data base: reliable parameters for aligning sequences and scoring gaps. Report, Swiss Federal Institute of Technology, 1991.

[945] G. H. Gonnet and F. W. Tompa. Mind your grammar: a new approach to modelling text. In P. M. Stocker and W. Kent, editors, *Proceedings of the 13th International Conference on Very Large Data Bases*, pages 339–346, Brighton, U. K., 1987.

[946] M. T. Goodrich and R. Tamassia. *Data Structures and Algorithms in Java*. John Wiley & Sons, 1998.

[947] O. Gotoh. An improved algorithm for matching biological sequences. *J. Mol. Biol.*, 162:705–708, 1982.

[948] O. Gotoh. Alignments of three biological sequences with an efficient trace-back procedure. *J. Theor. Biol.*, 121:327–337, 1986.

[949] O. Gotoh. Consistency of optimal sequence alignments. *Bull. Math. Biol.*, 52:509–525, 1990.

[950] O. Gotoh. Optimal alignment between groups of sequences and its application to multiple sequence alignment. *Comput. Appl. Biosci.*, 9(3):361–370, 1993.

[951] O. Gotoh. Homology-based gene structure prediction: Simplified matching algorithm using a translated codon (tron) and improved accuracy by allowing for long gaps. *Bioinformatics*, 16(?):?–?, 2000.

[952] A. Gräf. Left-to-right tree pattern matching. In R.V. Book, editor, *RTA*, number 488 in Lecture Notes in Computer Science, pages 323–334, Como, Italy, 1991. Springer-Verlag, Berlin.

[953] D. Gries and B. Burkhardt. Presenting an algorithm to find the minimum edit distance. Technical Report TR88-903, Cornell University, Computer Science Department, 1988.

[954] J. Griggs, P. Halton, A. Odlyzko, and M. S. Waterman. On the number of alignments of k sequences. *Graphs Comb.*, 6:133–146, 1990.

[955] J. Griggs, P. Halton, and M. S. Waterman. Sequence alignments with matched sections. *SIAM J. Algebr. Discret. Methods*, 7:604–608, 1986.

[956] R. E. Griswold. String scanning in the Icon programming language. *Comput. J.*, 33(2), 1990.

[957] M. Gross. Constructing lexicon-grammars. In B. T. S. Atkins and A. Zampolli, editors, *Computational Approaches to the Lexicon*, pages 259–316. Oxford University Press, 1993.

[958] M. Gross and D. Perrin, editors. *Electronic dictionaries and automata in computational linguistics*, volume 377 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1989.

[959] R. Grossi. A fast VLSI solution for the approximate string matching. *Integr., The VLSI J.*, 13(2):195–206, 1992.

[960] R. Grossi. On finding common subtrees. *Theor. Comput. Sci.*, 108(2):345–356, 1993.

[961] R. Grossi and F. Luccio. Simple and efficient string matching with k mismatches. *Inf. Process. Lett.*, 33(3):113–120, 1989.

[962] R. Grossi and J. S. Vitter. Compressed suffix arrays and suffix trees with applications to text indexing and string matching. In *Proceedings ACM Symposium on the Theory of Computing*, pages 397–406, Portland, Oregon, 2000. ACM Press.

[963] M. Gu, M. Farach, and R. Beigel. An efficient algorithm for dynamic text indexing. In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 697–704, Arlington, VA, 1994.

[964] Q.-P. Gua, S. Penga, and H. Sudboroughb. A 2-approximation algorithm for genome rearrangements by reversals and transpositions. *Theor. Comput. Sci.*, 210(2):327–339, 1999.

[965] A. Guénoche and P. Vitte. Distances entre chaînes de caractères par un algorithme sous-linéaire en place mémoire. In *Proceedings Distancia*, Rennes, France, 1992.

[966] L. Guibas. Periodicities in strings. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, pages 257–269. Springer-Verlag, Berlin, 1985.

[967] L. J. Guibas and A. M. Odlyzko. A new proof of the linearity of the Boyer-Moore string searching algorithm. In *Proceedings of the 18th IEEE Annual Symposium on Foundations of Computer Science*, pages 189–195, Providence, Rhode Island, 1977. IEEE Computer Society Press.

[968] L. J. Guibas and A. M. Odlyzko. A new proof of the linearity of the Boyer-Moore string searching algorithm. *SIAM J. Comput.*, 9(4):672–682, 1980.

[969] L. J. Guibas and A. M. Odlyzko. Periods in strings. *J. Comb. Theory Ser. A*, 30:19–43, 1981.

[970] L. J. Guibas and A. M. Odlyzko. String overlaps, pattern matching and nontransitive games. *J. Comb. Theory Ser. A*, 30:183–208, 1981.

[971] K. Guimarães and J. Melo. String alignment and approximate matching. Report TR-005/96, Departamento de Informática, Universidade Federal de Pernambuco, Recife, Brazil, 1996.

[972] K. Guimarães and J. Melo. String alignment and approximate matching. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 85–100, Recife, Brazil, 1996. Carleton University Press.

[973] S. K. Gupta, J. D. Kececioglu, and A. A. Schäffer. Making the shortest-paths approach to sum-of-pairs multiple sequence alignment more space efficient in practice. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 128–143, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[974] D. Gusfield. Efficient method for multiple sequence alignment with guaranteed error bounds. Report CSE-91-4, Computer Science Division, University of California, Davis, 1991.

[975] D. Gusfield. Efficient methods for multiple sequence alignment with guaranteed error bounds. *Bull. Math. Biol.*, 55(1):141–154, 1993.

[976] D. Gusfield. Faster implementation of a shortest superstring approximation. *Inf. Process. Lett.*, 51(5):271–274, 1994.

[977] D. Gusfield. *Algorithms on strings, trees and sequences: computer science and computational biology.* Cambridge University Press, Cambridge, 1997.

[978] D. Gusfield, K. Balasubramanian, and D. Naor. Parametric optimization of sequence alignment. In *Proceedings of the 3rd ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 432–439, Orlando, FL, 1992.

[979] D. Gusfield, K. Balasubramanian, and D. Naor. Parametric optimization of sequence alignment. *Algorithmica*, 12(4/5):312–326, 1994.

[980] D. Gusfield, R. M. Karp, L. Wang, and P. Stelling. Graph traversals, genes, and matroids: An efficient case of the travelling salesman problem. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 304–319, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[981] D. Gusfield, G. Landau, and B. Schieber. An efficient algorithm for all the pairs suffix-prefix problem. *Inf. Process. Lett.*, 41(4):181–185, 1992.

[982] C. Gutiérrez. Solving equations in strings: On Makanin's algorithm. In C. L. Lucchesi and A. V. Moura, editors, *Proceedings of the 3rd Latin American Symposium*, number 1380 in Lecture Notes in Computer Science, pages 358–373, Campinas, Brazil, 1998. Springer-Verlag, Berlin.

[983] R. Gwangsoo. An approximation algorithm for the reversible shortest common superstring problem. *Journal of the Korea Information Science Society*, 21(7):1261–1268, 1994. In Korean.

[984] E. h. Yang and Z. Zhang. The shortest common superstring problem: Average case analysis for both exact and approximate matching. *IEEE Trans. Inf. Theory*, 45(6):1867–1886, 1999.

[985] T. Hagerup. Optimal parallel string algorithms: Sorting, merging and computing the minimum. Report MPI-I-93-152, Max-Planck-Institut für Mathematik, Universität Saarbrücken, 1993.

[986] T. Hagerup. Optimal parallel string algorithms: Merging, sorting and computing the minimum. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 382–391, Montréal, Canada, 1994. ACM Press.

[987] L. C. K. Hai. Color set size problem with applications to string matching. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 230–243, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[988] H. Hakonen, T. Kaukoranta, and T. Raita. A comparison of longest common subsequence algorithms. University of Turku, Finland, 1996.

[989] H. Hakonen and T. Raita. A family of fast constant-space substring search algorithms. *Acta Cybernetica*, 14(2):239–250, 1999.

[990] K. Hakuta and H. Imai. The longest common subsequence problem for small alphabet size between many strings. In T. Ibaraki, Y. Inagaki, K. Iwama, T. Nishizeki, and M. Yamashita, editors, *Proceedings of the 3rd International Symposium on Algorithms and Computation*, number 650 in Lecture Notes in Computer Science, pages 469–478, Nagoya, Japan, 1992. Springer-Verlag, Berlin.

[991] A. Halaas. A systolic VLSI matrix for a family of fundamental search problem. *Integr., The VLSI J.*, 1(4):269–282, 1983.

[992] P. A. V. Hall and G. R. Dowling. Approximate string matching. *ACM Comp. Surv.*, 12(4):381–402, 1980.

[993] E. Hamilton. Expanded generalized regular expression. *Commun. ACM*, 31(12):1376–1385, 1988.

[994] R. W. Hamming. Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2):147–160, 1950.

[995] R. W. Hamming. *Coding and information theory.* Prentice Hall, Englewood, Cliff, NJ, 1982.

[996] R. W. Hamming. Hamming distance and its discovery. *UMAP: The Journal of Undergraduate Mathematics and its Applications*, 14, 1993.

[997] C. Hancart. Sur l'analyse en moyenne des algorithmes de recherche d'un mot dans un texte. Mémoire de dea d'informatique fondamentale, CERIL, Université Paris 7, 1990.

[998] C. Hancart. Algorithme de Morris et Pratt et ses raffinements : une analyse en moyenne. Rapport LITP 91.56, Universités Paris 6 et 7, France, 1991.

[999] C. Hancart. Une analyse en moyenne de l'algorithme de Morris et Pratt et de ses raffinements. In D. Krob, editor, *Actes des 2e Journées franco-belges: Théories des Automates et Applications*, number 176 in Publicatins de l'Université de Rouen, pages 99–110, Rouen, France, 1991.

[1000] C. Hancart. *Analyse exacte et en moyenne d'algorithmes de recherche d'un motif dans un texte.* Thèse de doctorat, Université Paris 7, 1993.

[1001] C. Hancart. Analyse exacte et en moyenne d'algorithmes de recherche d'un motif dans un texte. Rapport 93-11, IGM, Université de Marne-la-Vallée, France, 1993.

[1002] C. Hancart. On Simon's string searching algorithm. *Inf. Process. Lett.*, 47(2):95–99, 1993.

[1003] C. Hancart. Des bornes exactes de la complexité des algorithmes séquentiels de recherche d'un motif. *Bull. Belg. Math. Soc.*, 1(2):239–252, 1994.

[1004] L. Hanks, R. K. Cytron, and W. Gillett. An efficient algorithm for developing topologically valid matchings. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 144–161, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1005] S. Hannenhalli. Polynomial-time algorithm for computing translocation distance between genomes. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 162–176, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1006] S. Hannenhalli and P. A. Pevzner. Transforming cabbage into turnip: Polynomial algorithm for sorting signed permutations by reversals. *J. Assoc. Comput. Mach.*, 46(1):1–27, 1999.

[1007] D. Hanson. Literate programming – printing common words. *Commun. ACM*, 30(7):594–599, 1987.

[1008] D. Harel and R. E. Tarjan. Fast algorithms for finding nearest common ancestors. *SIAM J. Comput.*, 13(2):338–355, 1984.

[1009] R. Hariharan. Optimal parallel suffix tree construction. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 290–299, Montréal, Canada, 1994. ACM Press.

[1010] R. Hariharan. Optimal parallel suffix tree construction. *J. Comput. Syst. Sci.*, 55(1):44–69, 1997.

[1011] M. C. Harrison. Implementation of the substring test by hashing. *Commun. ACM*, 14(12):777–779, 1971.

[1012] W. E. Hart and S. Istrail. Invariant patterns in crystal lattices: Implications for protein folding algorithms (extended abstract). In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 288–303, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1013] A. Hartman and M. Rodeh. Optimal sparsing of strings. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 155–167. Springer-Verlag, Berlin, 1985.

[1014] K. Hashiguchi and K. Yamada. Two recognizable string-matching problems over free partially commutative monoids. *Theor. Comput. Sci.*, 92(1):77–86, 1992.

[1015] K. Hashigushi and K. Yamada. String matching problems over free partially commutative monoid. *Inf. Comput.*, 101(2):131–149, 1992.

[1016] R. L. Haskins. Hardware for searching very large text data-bases. In *Proceedings of the 5th Workshop on Computer Architecture for Non-numeric Processing*, pages ?–?, ACM, New York (published as a special issue of SIGIR Forum 15(2)), 1980.

[1017] R. L. Haskins. Special purpose processors for text retrieval. *Database Eng.*, 4(1):16–29, 1981.

[1018] R. L. Haskins and L. A. Hollaar. Operational characteristics of a hardware-based pattern matcher. *ACM Trans. Database Syst.*, 8(1):15–40, 1983.

[1019] P. J. Hatcher and T. W. Cristopher. High quality code generation via bottom up trie pattern matching. In *Proceedings of the 13th Annual ACM Symposium on Principles of Programming Languages*, pages 119–130, St. Petersburg Beach, FA, 1986.

[1020] J. P. Haton. *Contribution à l'analyse, paramétrisation et la reconnaissance automatique de la parole.* Thèse de doctorat d'état, Université de Nancy, France, 1973.

[1021] G. Havas and J.-X. Lian. A new problem in string searching. In Ding-Zhu Du and Xiang sun Zhang, editors, *Proceedings of the 5th International Symposium on Algorithms and Computation*, number 834 in Lecture Notes in Computer Science, pages 660–668, Beijing, China, 1994. Springer-Verlag, Berlin.

[1022] H. S. Heaps. *Information retrieval: theoretical and computational aspects.* Academic Press, New York, NY, 1978.

[1023] J.-J. Hébrard. *Distances sur les mots. Application à la recherche de motifs.* Thèse de doctorat, Université de Rouen, 1984.

[1024] J.-J. Hébrard. An algorithm for distinguishing efficiently bit-strings by their subsequences. *Theor. Comput. Sci.*, 82(1):35–49, 1991.

[1025] J.-J. Hébrard and M. Crochemore. Calcul de la distance par les sous-mots. *RAIRO Inform. Théor. Appl.*, 20(4):441–456, 1986.

[1026] P. Heckel. A technique for isolating differences between files. *Commun. ACM*, 21(4):264–268, 1978.

[1027] J. Hein. A new method that simultaneously aligns and reconstruct ancestral sequences for any number of homologous sequences, when the phylogeny is given. *Mol. Biol. Evol.*, 6:649–668, 1989.

[1028] J. Hein. Unified approach to alignment and phylogenies. *Methods Enzymol.*, 183:626–645, 1990.

[1029] J. Hein, T. Jiang, L. Wang, and K. Zhang. On the complexity of comparing evolutionary trees. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 177–190, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1030] G. Held. *Data Compression - Techniques and Applications, Hardware and Software Considerations*. John Wiley & Sons, New York, NY, 1987.

[1031] J. P. Held and J. V. Carlis. MATCH – A new high-level relational operator for pattern matching. *Commun. ACM*, 30(1):62–75, 1987.

[1032] M. Held and R. M. Karp. A dynamic programming approach to sequencing problems. *SIAM J. Appl. Math.*, 10:196–210, 1962.

[1033] C. M. Henneke. A multiple sequence alignment algorithm for homologous proteins using secondary structure information and optionally keying alignments to functionally important sites. *Comput. Appl. Biosci.*, 5:141–150, 1989.

[1034] G. Herrmannsfeldt. A highly parallel finite state automaton processor for biological pattern matching. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 58–72, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1035] G. Z. Hertz and G. D. Stormo. Identifying DNA and protein patterns with statistically sign ificant alignments of multiple sequences. *Bioinformatics*, 15:563–577, 1999.

[1036] J. Herz and R. D. Hersch. Analysing character shapes by string matching techniques. *Electronic Publishing*, 6(3):261–272, 1993.

[1037] J. H. Hester and D. S. Hirschberg. Self-organizing linear search. *ACM Comp. Surv.*, 17:295–311, 1985.

[1038] K. Heumann and H.-W. Mewes. The Hashes Position Tree (HPT): a suffix tree variant for large data sets stored on slow mass storage devices. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 101–114, Recife, Brazil, 1996. Carleton University Press.

[1039] T. Hickey and J. Cohen. Uniform random generation of strings in acontext-free language. *SIAM J. Comput.*, 12(4):645–655, 1983.

[1040] D. G. Higgins, A. J. Bleasby, and R. Fuchs. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.*, 8(2):189–191, 1992.

[1041] D. G. Higgins and P. M. Sharp. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, 73:237–244, 1988.

[1042] D. G. Higgins and P. M. Sharp. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.*, 5:151–153, 1989.

[1043] Kazunori Hirashima, Hideo Bannai, Wataru Matsubara, Akira Ishino, and Ayumi Shinohara. Bit-parallel algorithms for computing all the runs in a string. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 203–213, Czech Technical University in Prague, Czech Republic, 2009.

[1044] M. Hirata, H. Yamada, H. Nagai, and K. Takahashi. A versatile data string-search VLSI. In *Proceedings of the IEEE Custom Integrated Circuits Conference*, pages 563–566, 1987.

[1045] M. Hirata, H. Yamada, H. Nagai, and K. Takahashi. A versatile data string-search VLSI. *IEEE J. Solid-State Circuits*, 23(2):329–335, 1988.

[1046] D. S. Hirschberg. A linear space algorithm for computing maximal common subsequences. *Commun. ACM*, 18(6):341–343, 1975.

[1047] D. S. Hirschberg. *The longest common subsequence problem*. PhD thesis, Princeton University, 1975.

[1048] D. S. Hirschberg. Algorithms for the longest common subsequence problem. *J. Assoc. Comput. Mach.*, 24(4):664–675, 1977.

[1049] D. S. Hirschberg. An information-theoretic lower bound for the longest common subsequence problem. *Inf. Process. Lett.*, 7(1):40–41, 1978.

[1050] D. S. Hirschberg. Recent results on the complexity of common subsequence problems. In D. Sankoff and J. B. Kruskall, editors, *Time Warps, String Edits, and Macromolecules: the Theory and Practice of Sequence Comparison*, chapter 12, pages 325–330. Addison-Wesley, Reading, MA, 1983.

[1051] D. S. Hirschberg. Serial computations of Levenshtein distances. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 4, pages 123–141. Oxford University Press, 1997.

[1052] D. S. Hirschberg. Bounds on the number of string subsequences. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 115–122, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1053] D. S. Hirschberg and L. L. Larmore. The least weight subsequence problem. *SIAM J. Comput.*, 16(4):628–638, 1987.

[1054] D. S. Hirschberg and L. L. Larmore. New applications of failure functions. *J. Assoc. Comput. Mach.*, 24(3):616–625, 1987.

[1055] D. S. Hirschberg and L. L. Larmore. The set LCS problem. *Algorithmica*, 2:91–95, 1987.

[1056] D. S. Hirschberg and L. L. Larmore. The set-set LCS problem. *Algorithmica*, 4:503–510, 1989.

[1057] D. S. Hirschberg and M. Régnier. Tight bounds on the number of string subsequences. *J. Discret. Algorithms*, 1(1):123–132, 2000.

[1058] D. T. Hoang. A systolic array for the sequence alignment problem. Report CS-92-22, Brown University, 1992.

[1059] C. M. Hoffman and M. J. O'Donnell. Pattern matching in trees. *J. Assoc. Comput. Mach.*, 29(1):68–95, 1982.

[1060] L. A. Hollaar. Text retrievals computers. *IEEE Trans. Comput.*, 12:40–50, 1979.

[1061] L. A. Hollaar and D. C. Roberts. Current research into specialized processors for text information retrieval. In S. Bing Yao, editor, *Proceedings of the 4th International Conference on Very Large Data Bases*, pages 270–279, West Berlin, Germany, 1978.

[1062] I. Holmes and R. Durbin. Dynamic programmig alignment accuracy. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 102–108, New York, NY, 1998. ACM Press.

[1063] N. Holsti and E. Sutinen. Approximate string matching using q-gram places. In *Proceedings of the 7th Finnish Symposium on Computer Science*, pages 23–32, 1994.

[1064] J. Holub. Approximate string matching in text. Master's thesis, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic, 1996.

[1065] J. Holub. Reduced nondeterministic finite automata for approximate string matching. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 19–27, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[1066] J. Holub. Simulation of NFA in approximate string and sequence matching. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 39–46, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[1067] J. Holub. Dynamic programming for reduced NFAs for approximate string and sequence matching. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 73–82, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1068] J. Holub. Simulation of nondeterministic finite automata in approximate string and sequence matching. Technical Report DC–98–04, Department of Computer Science and Engineering, Czech Technical University, Prague, Czech Republic, 1998.

[1069] J. Holub, C. S. Iliopoulos, B. Melichar, and L. Mouchard. Distributed string matching using finite automata. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 114–128, Perth, WA, Australia, 1999.

[1070] J. Holub and B. Melichar. Approximate string matching using factor automata. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 28–39, Perth, WA, Australia, 1998.

[1071] J. Holub and B. Melichar. Implementation of nondeterministic finite automata for approximate pattern matching. In J.-M. Champarnaud, D. Maurel, and D. Ziadi, editors, *Proceedings of the 3rd International Workshop on Implementing Automata'98*, number 1660 in Lecture Notes in Computer Science, pages 92–99, Rouen, France, 1999. Springer-Verlag, Berlin.

[1072] J. Holub and B. Melichar. Approximate string matching using factor automata. *Theor. Comput. Sci.*, 249(2):305–311, 2000.

[1073] Jan Holub and Jan Žďárek, editors. *Proceedings of the Prague Stringology Conference 2008, Prague, Czech Republic, September 1–3, 2008*. Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, 2008.

[1074] Jan Holub and Jan Žďárek, editors. *Proceedings of the Prague Stringology Conference 2009, Prague, Czech Republic, August 31–September 2, 2009*. Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, 2009.

[1075] J. E. Hopcroft and J. D. Ullman. *Introduction to automata, languages and computations.* Addison-Wesley, Reading, MA, 1979.

[1076] R. N. Horspool. Practical fast searching in strings. *Softw. Pract. Exp.*, 10(6):501–506, 1980.

[1077] P. Horton and T. Sukuba. BB: A branch and bound algorithm for local multiple sequence alignment. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 84–98, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1078] P. G. Howard. Lossless and lossy compression of text images by soft pattern matching. In J. A. Storer and M. Cohn, editors, *Proceedings of the 6th Data Compression Conference*, pages 210–219, Snowbird, Utah, 1996. IEEE Computer Society Press.

[1079] W. J. Hsu and M. W. Du. Computing a longest common subsequence for a set of strings. *BIT*, 24:45–59, 1984.

[1080] W. J. Hsu and M. W. Du. New algorithms for the LCS problem. *J. Comput. Syst. Sci.*, 29(2):133–152, 1984.

[1081] Y.-J. Hu, S. Sandmeyer, C. McLaughlin, and D. Kibler. Combinatorial motif analysis and hypothesis generation on a genomic scale. *Bioinformatics*, 16(3):222–232, 2000.

[1082] Y. Hua, T. Jiang, and B. Wu. Aligning DNA sequences to minimized the change in protein. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 221–234, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1083] S. S. Huang and H. S. Asuri. Algorithms for the set LCS and the set-set LCS problems. In R. Sanicki and W. W. Koczkodaj, editors, *Computing and Information*, pages 81–87. Elsevier, Amsterdam, The Netherlands, 1989.

[1084] X. Huang. A lower bound for the edit-distance problem under an arbitrary cost function. *Inf. Process. Lett.*, 27(6):319–321, 1988.

[1085] X. Huang. A space-efficient parallel sequence comparison algorithm for a message-passing multiprocessor. *Int. J. Parallel Program.*, 18(3):223–239, 1989.

[1086] X. Huang. A contig assembly program based on sensitive detection of fragment overlaps. *Gene*, 14:18–25, 1992.

[1087] X. Huang. A context dependent method for comparing sequences. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 54–63, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1088] X. Huang. On global sequence alignement. *Comput. Appl. Biosci.*, 10:227–235, 1994.

[1089] X. Huang, R. C. Hardison, and W. Miller. A space-efficient algorithm for local similarities. *Comput. Appl. Biosci.*, 6(4):373–381, 1990.

[1090] X. Huang and W. Miller. A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.*, 12:337–357, 1991.

[1091] X. Huang, P. A. Pevzner, and W. Miller. Parametric recomputing in alignment graphs. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 87–101, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1092] D. A. Huffman. A method for the construction of minimum redundancy codes. *Proc. I.R.E.*, 40:1098–1101, 1951.

[1093] L. C. K. Hui. Color set size problem with applications to string matching. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 644 in Lecture Notes in Computer Science, pages 230–243, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1094] A. Hume. A tale of two greps. *Softw. Pract. Exp.*, 18(11):1063–1072, 1988.

[1095] A. Hume. Keyword matching. Report, Bell Laboratories, 1990.

[1096] A. Hume and D. M. Sunday. Fast string searching. *Softw. Pract. Exp.*, 21(11):1221–1248, 1991.

[1097] A. Hume and D. M. Sunday. Fast string searching. In *Proceedings Summer USENIX Conference*, pages 221–233, Nashville, TN, 1991.

[1098] J. W. Hunt and M. D. McIlroy. An algorithm for differential file comparison. Computing Science Technical Report 41, Bell Laboratories, 1975.

[1099] J. W. Hunt and T. G. Szymanski. A fast algorithm for computing longest common subsequences. *Commun. ACM*, 20(5):350–353, 1977.

[1100] T. Hwa and M. Lässig. Optimal detection of sequence similarity by local alignment. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 109–116, New York, NY, 1998. ACM Press.

[1101] H. Hyyrö. A bit-vector algorithm for computing levenshtein and Damerau edit distances. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 44–54, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[1102] Heikki Hyyrö. An input sensitive online algorithm for lcs computation. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 192–202, Czech Technical University in Prague, Czech Republic, 2009.

[1103] O. H. Ibarra, T. Jiang, and H. Wang. String editing on a one–way linear array of finite state machines. *IEEE Trans. Comput.*, 41(1):112–118, 1992.

[1104] O. H. Ibarra, T. Pong, and S. M. Sohn. Hypercube algorithms for some string comparison problems. In *Proceedings of the International Conference on Parallel Processing*, volume 3, pages 190–193, 1988.

[1105] O. H. Ibarra, T. C. Pong, and S. M. Sohn. String processing on the hypercube. *IEEE Trans. Acoust. Speech Signal Process.*, 38(1):160–164, 1990.

[1106] R. Idury. Dynamic multiple pattern matching. Technical Report TR93-207, Rice University, 1998.

[1107] R. M. Idury and A. A. Schäffer. Dynamic dictionary matching with failure functions. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 276–287, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1108] R. M. Idury and A. A. Schäffer. Multiple matching of rectangular patterns. In *Proceedings of the 25th ACM Symposium on the Theory of Computing*, pages 81–90, San Diego, CA, 1993. ACM Press.

[1109] R. M. Idury and A. A. Schäffer. Dynamic dictionary matching with failure functions. *Theor. Comput. Sci.*, 131(2):295–310, 1994.

[1110] R. M. Idury and A. A. Schäffer. Multiple matching of parameterized patterns. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 226–239, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1111] R. M. Idury and A. A. Schäffer. Multiple matching of rectangular patterns. *Inf. Comput.*, 117(1):78–90, 1995.

[1112] R. M. Idury and A. A. Schäffer. Multiple matching of parameterized patterns. *Theor. Comput. Sci.*, 154(2):203–224, 1996.

[1113] C. S. Iliopoulos, M. Kumar, L. Mouchard, and S. Venkatesh. Computing all motif occurrences in polyphonic musical sequences. Technical Report 9-98, School of Computing, Curtin University, Australia, 1998.

[1114] C. S. Iliopoulos, M. Kumar, L. Mouchard, and S. Venkatesh. Motif evolution in polyphonic musical sequences. In L. Brankovic and J. Ryan, editors, *Proceedings of the 11th Australasian Workshop On Combinatorial Algorithms*, pages 53–66, Hunter Valley, Australia, 2000.

[1115] C. S. Iliopoulos and M. Kurokawa. String matching with gaps for musical melodic recognition. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 55–64, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[1116] C. S. Iliopoulos, T. Lecroq, L. Mouchard, and Y. J. Pinzon. Computing approximate repetitions in musical sequences. *Int. J. Comput. Math.*, 83(1/2). To appear.

[1117] C. S. Iliopoulos, T. Lecroq, L. Mouchard, and Y. J. Pinzon. Computing repetitions in musical sequences. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 49–59, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[1118] C. S. Iliopoulos, M. Mohamed, L. Mouchard, K. G. Perdikuri, W. F. Smyth, and A. K. Tsakalidis. String regularities with don't cares. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 65–74, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[1119] C. S. Iliopoulos, D. Moore, and W. F. Smyth. A characterization of the squares in a Fibonacci string. *Theor. Comput. Sci.*, 172(1–2):281–291, 1997.

[1120] C. S. Iliopoulos, D. W. G. Moore, and K. Park. Covering a string. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 54–62, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1121] C. S. Iliopoulos, D. W. G. Moore, and W. F. Smyth. A linear algorithm for computing the squares of a fibonacci string. In P. Eades and M. Moule, editors, *Proceedings of the Computing: Australian Theory Symposium*, pages 55–63, Melbourne, Australia, 1996.

[1122] C. S. Iliopoulos and L. Mouchard. An $o(n \log n)$ algorithm for computing all maximal quasiperiodicities in strings. In C. S. Calude and M. J. Dinneen, editors, *Combinatorics, Computation and Logic. Proceedings of DMTCS'99 and CATS'99*, Lecture Notes in Computer Science, pages 262–272, Auckland, New-Zealand, 1999. Springer-Verlag, Singapore.

[1123] C. S. Iliopoulos and L. Mouchard. Quasiperiodicity and string covering. *Theor. Comput. Sci.*, 218(1):205–216, 1999.

[1124] C. S. Iliopoulos and L. Mouchard. Quasiperiodicity: From detection to normal forms. *J. Autom. Lang. Comb.*, 4(3):213–228, 1999.

[1125] C. S. Iliopoulos, L. Mouchard, and S. Venkatesh. Optimal computation of all the repetitions in a string. Technical Report 8-98, School of Computing, Curtin University, Australia, 1998.

[1126] C. S. Iliopoulos and K. Park. An optimal $o(\log\log n)$-time algorithm for parallel superprimitivity testing. *J. Korea Information Science Society*, 21(8):1400–1404, 1994.

[1127] C. S. Iliopoulos and K. Park. A work-time optimal algorithm for computing all string covers. *Theor. Comput. Sci.*, 164(1–2):299–310, 1996.

[1128] C. S. Iliopoulos and Y. J. Pinzon. Recovering an LCS in $o(n^2/w)$ time and space. In E. Tri Baskoro, editor, *Proceedings of the 12th Australasian Workshop On Combinatorial Algorithms*, pages 80–87, Bandung, Indonesia, 2001.

[1129] C. S. Iliopoulos and J. F. Reid. An optimal parallel algorithm for analysing occluded images. In N. Sharda and A. Tam, editors, *Proceedings of the 4th Annual Australian Conference on Parallel and Real-Time Systems*, pages 104–113, Newcastle, Australia, 1997. Springer-Verlag, Berlin.

[1130] C. S. Iliopoulos and J. F. Reid. Validating and decomposing partially occluded two-dimensional images. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 83–94, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1131] C. S. Iliopoulos and J. Simpson. On-line validation and analysis of occluded images. In V. Estivill-Castro, editor, *Proceedings of the 8th Australasian Workshop On Combinatorial Algorithms*, pages 25–36, Queensland University of Technology, Australia, 1997.

[1132] C. S. Iliopoulos and W. F. Smyth. Optimal algorithms for computing the canonical form of a circular string. *Theor. Comput. Sci.*, 92(1):87–105, 1992.

[1133] C. S. Iliopoulos and W. F. Smyth. A characterization of the squares in a Fibonacci string. *Theor. Comput. Sci.*, 172(1–2):281–291, 1997.

[1134] C. S. Iliopoulos and W. F. Smyth. On-line algorithms for k-covering. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 97–106, Perth, WA, Australia, 1998.

[1135] Costas S. Iliopoulos, Mirka Miller, and Solon P. Pissis. Parallel algorithms for degenerate and weighted sequences derived from high throughput sequencing technologies. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 249–262, Czech Technical University in Prague, Czech Republic, 2009.

[1136] H. Imai and T. Ikeda. $k-$Group Multiple Alignment Based on A* Search. In *Proceedings of the 6th Genome Inform. Workshop*, pages 9–18, 1995.

[1137] P. Indyk. Deterministic superimposed coding with applications to pattern matching. In *Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science*, pages 127–136. IEEE Computer Society Press, 1997.

[1138] P. Indyk. Faster algorithms for string matching problems: matching the convolution bound. In *Proceedings of the 39th IEEE Annual Symposium on Foundations of Computer Science*, pages 166–173. IEEE Computer Society Press, 1998.

[1139] S. Inenaga. Bidirectional construction of suffix trees. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 75–87, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[1140] Shunsuke Inenaga and Hideo Bannai. Finding characteristic substrings from compressed texts. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 40–54, Czech Technical University in Prague, Czech Republic, 2009.

[1141] K. Inoue and I. Takanami. A characterization of recognizable picture languages. In A. Nakamura, M. Nivat, A. Saoudi, P.S.P. Wang, and K. Inoue, editors, *Proceedings of the 2nd International Colloquium on Parallel Image Analysis*, number 654 in Lecture Notes in Computer Science, pages 133–143, Ube, Japan, 1992. Springer-Verlag, Berlin.

[1142] R. W. Irving. Suffix binary search trees. Technical report, University of Glasgow, Computing Science Department, 1996.

[1143] R. W. Irving and C. B. Fraser. Two algorithms for the longest common subsequence of three (or more) strings. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 214–229, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1144] R. W. Irving and C. B. Fraser. On the worst-case behaviour of some approximation algorithms for the shortest common supersequence of k

strings. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 63–73, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1145] R. W. Irving and C. B. Fraser. Maximal common subsequences and minimal common supersequences. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 173–183, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1146] M. E. Isenman and D. E. Shasha. Performance and architectural issues for string matching. *IEEE Trans. Comput.*, 39(2):238–250, 1990.

[1147] A. Itai. Optimal alphabetic trees. *SIAM J. Comput.*, 5(1):9–18, 1976.

[1148] K. Itano, Y. Sato, H. Hirai, and T. Yamagata. An incremental pattern matching algorithm for the pipelined lexical scanner. *Inf. Process. Lett.*, 27(5):253–258, 1988.

[1149] Ito and Shizuya. On the difficulty of searching for a string without decryption. *TIEICE: IEICE Transactions on Communications/Electronics/Information and Systems*, pages ?–?, 1999.

[1150] M. Ito, M. Katsura, H. J. Shyr, and S. S. Yu. Automata accepting primitive words. *Semigroup Forum*, 37:45–50, 1988.

[1151] M. Ito, K. Shimizu, M. Nakanishi, and A. Hashimoto. Polynomial-time algorithms for computing characteristic strings. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 274–288, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1152] T. Ito and M. Kizawa. Hierarchical file organization and its application to similar-string matching. *ACM Trans. Database Syst.*, 8(3):410–433, 1983.

[1153] S. Y. Itoga. The string merging problem. *BIT*, 21:20–30, 1981.

[1154] A. G. Ivanov. Distinguishing an approximate word's inclusion on turing machine in real time. *Izv. Akad. Nauk SSSR Ser. Mat.*, 48:520–568, 1984.

[1155] S. Iyenger and V. Alia. A string search algorithm. *Appl. Math. Comput.*, 6:123–131, 1980.

[1156] G. Jacobson and K.-P. Vo. Heaviest increasing/common subsequence. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 52–66, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1157] P. Jacquet, B. Rais, and W. Szpankowski. Compact suffix trees ressemble patricia trees: limiting distribution of depth. Report CSD-TR-92-048, Purdue University, 1992.

[1158] P. Jacquet and W. Szpankowski. Autocorrelation on words and its applications. analysis of suffix tree by string-ruler approach. Rapport TR-1106, INRIA, 1989.

[1159] P. Jacquet and W. Szpankowski. What can we learn about suffix trees from independent tries ? In F. Dehne, J.-R. Sack, and N. Santoro, editors, *Proceedings of the 2nd Workshop on Algorithms and Data Structures*, number 519 in Lecture Notes in Computer Science, pages 228–239, Ottawa, Canada, 1991. Springer-Verlag, Berlin.

[1160] J. Jájá. *An introduction to parallel algorithms*. Addison-Wesley, Reading, MA, 1992.

[1161] J. Jájá, K.W. Ryu, and U. Vishkin. Sorting strings and constructing digital search trees in parallel. *Theor. Comput. Sci.*, 154(2):225–245, 1996.

[1162] Jan Janoušek. String suffix automata and subtree pushdown automata. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 160–172, Czech Technical University in Prague, Czech Republic, 2009.

[1163] K. Jantke. Polynomial time inference of general pattern languages. In M. Fontet and K. Mehlhorn, editors, *Proceedings of the 1st Annual Symposium on Theoretical Aspects of Computer Science*, number 166 in Lecture Notes in Computer Science, pages 314–325, Paris, France, 1984. Springer-Verlag, Berlin.

[1164] T. Jiang, Z. Jiang, and D. Breslauer. Rotation of periodic strings and short superstrings. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 115–125, Recife, Brazil, 1996. Carleton University Press.

[1165] T. Jiang, E. L. Lawler, and L. Wang. Aligning sequences via an evolutionary tree: complexity and approximation. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 760–769, Montréal, Canada, 1994. ACM Press.

[1166] T. Jiang and M. Li. Towards a DNA sequencing theory (learning a string). Report 91-08, Computer Science Department, McMaster University Hamilton, Ontario, 1991.

[1167] T. Jiang and M. Li. Approximating shortest superstrings with constraints. In F. Dehne, J.-R. Sack, N. Santoro, and S. Whitesides, editors, *Proceedings of the 3rd Workshop on Algorithms and Data Structures*, number 709 in Lecture Notes in Computer Science, pages 385–396, Montréal, Canada, 1993. Springer-Verlag, Berlin.

[1168] T. Jiang and M. Li. k one-way heads cannot do string matching. In *Proceedings of the 25th ACM Symposium on the Theory of Computing*, pages 62–70, San Diego, CA, 1993. ACM Press.

[1169] T. Jiang and M. Li. On the complexity of learning strings and sequences. *Theor. Comput. Sci.*, 119(2):363–373, 1993.

[1170] T. Jiang and M. Li. On the approximation of shortest common supersequences and longest common subsequences. In S. Abiteboul and E. Shamir, editors, *Proceedings of the 21st International Colloquium on Automata, Languages and Programming*, number 820 in Lecture Notes in Computer Science, pages 191–202, Jerusalem, Israel, 1994. Springer-Verlag, Berlin.

[1171] T. Jiang and M. Li. Approximating shortest superstrings with constraints. *Theor. Comput. Sci.*, 134(2):473–491, 1995.

[1172] T. Jiang and M. Li. On the approximation of shortest common supersequences and longest common subsequences. *SIAM J. Comput.*, 24(5):1122–1139, 1995.

[1173] T. Jiang, M. Li, and D. Du. A note on shortest superstrings with flipping. *Inf. Process. Lett.*, 44(4):195–199, 1992.

[1174] T. Jiang, G.-H. Lin, B. Ma, and K. Zhang. The longest common subsequence problem for arc-annotated sequences. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 154–165, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1175] T. Jiang, A. Salomaa, K. Salomaa, and Y. Sheng. Inclusion is undecidable for pattern languages. In A. Lingas, R. G. Karlsson, and S. Carlsson, editors, *Proceedings of the 20th International Colloquium on Automata, Languages and Programming*, number 700 in Lecture Notes in Computer Science, pages 301–312, Lund, Sweden, 1993. Springer-Verlag, Berlin.

[1176] T. Jiang, L. Wang, and K. Zhang. Alignment of trees: an alternative to tree edit. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 75–86, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1177] P. Johansen. Inductive inference of ultimately periodic sequences. *BIT*, 1:573–580, 1988.

[1178] P. Johansen. Combinatorial pattern matching, the method and the program package. Report 10, Department of Computer Science, University of Copenhagen, Denmark, 1994.

[1179] P. Johansen. On-line string matching with feedback. *Theor. Comput. Sci.*, 141(1–2):53–67, 1995.

[1180] M. S. Johnson and R. F. Doolittle. A method for the simultaneous alignment of three or more amino acid sequences. *J. Mol. Evol.*, 23:267–278, 1986.

[1181] T. Johtela, J. Smed, H. Hakonen, and T. Raita. An efficient heuristic for the lcs problem. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 126–140, Recife, Brazil, 1996. Carleton University Press.

[1182] P. Jokinen, J. Tarhio, and E. Ukkonen. A comparison of approximate string matching algorithms. Report A-1991-7, Department of Computer Science, University of Helsinky, 1991.

[1183] P. Jokinen, J. Tarhio, and E. Ukkonen. A comparison of approximate string matching algorithms. *Softw. Pract. Exp.*, 26(12):1439–1458, 1996.

[1184] P. Jokinen and E. Ukkonen. Two algorithms for approximate string matching in static texts. In A. Tarlecki, editor, *Proceedings of the 16th Symposium on Mathematical Foundations of Computer Science*, number 520 in Lecture Notes in Computer Science, pages 240–248, Kazimierz Dolny, Poland, 1991. Springer-Verlag, Berlin.

[1185] L. P. Jones, E. W. Gassie, and S. Radhakrishnan. PORTREP: a portable repeated string finder. *Softw. Pract. Exp.*, 19(1):63–77, 1989.

[1186] N. D. Jones. A note on linear time simulation of deterministic two-way pushdown automata. *Inf. Process. Lett.*, 6(4):110–112, 1977.

[1187] D. Joseph, J. Meidanis, and P. Tiwari. Determining DNA sequences similarity using maximum independant set algorithms for interval graphs. In O. Nurmi and E. Ukkonen, editors, *Proceedings of the 3rd Scandinavian Workshop on Algorithm Theory*, number 621 in Lecture Notes in Computer Science, pages 326–337, Helsinki, Finland, 1992. Springer-Verlag, Berlin.

[1188] D. Julian and M. Davies. String matching in text editors. *Softw. Pract. Exp.*, 12(8):709–717, 1982.

[1189] A. Julich. Implementations of BLAST for parallel computers. *Comput. Appl. Biosci.*, 11(1):3–6, 1995.

[1190] W. Just and G. Della Vedova. Multiple sequence alignment as a facility location problem. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 60–70, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[1191] J. Justin and G. Pirillo. On a combinatorial property of Sturmian words. *Theor. Comput. Sci.*, 154(2):387–394, 1996.

[1192] D. Kandel, Y. Matias, R. Unger, and P. Winkler. Shuffling biological sequences. *Discret. Appl. Math.*, 71(1–3):171–185, 1996.

[1193] S. K. Kannan and E. W. Myers. An algorithm for locating non-overlapping regions of maximum alignment score. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 74–86, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1194] M.-Y. Kao. Tree contractions and evolutionary trees. *SIAM J. Comput.*, 27(6):1592–1616, 1998.

[1195] M.-Y. Kao, T. W. Lam, W.-K. Sung, and H.-F. Ting. A faster and unifying algorithm for comparing trees. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 129–142, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1196] M.-Y. Kao and S. R. Tate. Online matching with blocked input. *Inf. Process. Lett.*, 38(3):113–116, 1991.

[1197] J. Karhumäki. On cube-free ω-words generated by binary morphisms. *Discret. Appl. Math.*, 5:279–297, 1983.

[1198] J. Karhumäki, J. Manuch, and W. Plandowski. On defect effect of bi-infinite words. In L. Brim, J. Gruska, and J. Zlatuska, editors, *Proceedings of the 23rd Symposium on Mathematical Foundations of Computer Science*, number 1450 in Lecture Notes in Computer Science, pages 674–682, Brno, Czech Republic, 1998. Springer-Verlag, Berlin.

[1199] J. Karhumäki, W. Plandowski, and W. Rytter. On the complexity of computing the order of repetition of a string. Technical Report TUCS-TR-226, Turku Centre for Computer Science, Finland, 1998.

[1200] J. Karhumäki, W. Plandowski, and W. Rytter. The compression of subsegments of images described by finite automata. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 186–195, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1201] J. Karhumäki, W. Plandowski, and W. Rytter. Generalized factorizations of words and their algorithmic properties. *Theor. Comput. Sci.*, 218(1):123–133, 1999.

[1202] J. Karhumäki, W. Plandowski, and W. Rytter. Pattern-matching problems for two-dimensional images described by finite automata. *Nordic J. Comput.*, 7(1):1–13, 2000.

[1203] L. Kari, A. Mateescu, G. Paun, and A. Salomaa. On parallel deletions applied to aword. *RAIRO Inform. Théor. Appl.*, 29(2):129–144, 1995.

[1204] J. Kärkkäinen. Suffix cactus: a cross between suffix tree and suffix array. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 191–204, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1205] J. Kärkkäinen, G. Navarro, and E. Ukkonen. Approximate string matching over ziv-lempel compressed text. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 195–209, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1206] J. Kärkkäinen and E. Ukkonen. Two and higher dimensional pattern matching in optimal expected time. In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 715–723, Arlington, VA, 1994.

[1207] J. Kärkkäinen and E. Ukkonen. Lempel-ziv parsing and sublinear-size index structures for string matching. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 141–155, Recife, Brazil, 1996. Carleton University Press.

[1208] J. Kärkkäinen and E. Ukkonen. Sparse suffix trees. In J.-K. Cai and C. K. Wong, editors, *Proceedings of the 2ndAnnual International Conference on Computing and Combinatorics*, volume 1090, pages 219–230, 1996.

[1209] J. Kärkkäinen and E. Ukkonen. Two- and higher-dimensional pattern matching in optimal expected time. *SIAM J. Comput.*, 29(2):571–589, 2000.

[1210] S. Karlin, G. Ghandour, F. Ost, S. Tavare, and J. L. Korn. New approaches for computer analysis of nucleic acid sequences. *Proc. Natl. Acad. Sci. U.S.A.*, 80:5660–5664, 1983.

[1211] S. Karlin, M. Morris, G. Ghandour, and M.-Y. Leung. Efficient algorithms for molecular sequence analysis. *Proc. Natl. Acad. Sci. U.S.A.*, 85:841–845, 1988.

[1212] S. Karlin and F. Ost. Counts of long aligned matches among random letter sequences. *Ann. Appl. Probab.*, 19:239–351, 1987.

[1213] S. Karlin and F. Ost. Maximum length of common words among random letter sequences. *Ann. Appl. Probab.*, 16:535–563, 1988.

[1214] S. Karlin, H. W. Trickey, and J. D. Ullman. Algorithms for the compilation of regular expressions into PLAs. *Algorithmica*, 2:283–314, 1987.

[1215] H. Karloff. Fast algorithms for approximately counting mismatches. *Inf. Process. Lett.*, 48(2):53–60, 1993.

[1216] R. M. Karp. On-line algorithms versus off-line algorithms: How much is it worth to know the future. In J. van Leeuwen, editor, *Algorithms, Software, Architecture*, volume Information Processing 92, Volume I of *IFIP Transactions A-12*, pages 416–429. Elsevier Science Publishers B.V., 1992.

[1217] R. M. Karp, R. E. Miller, and A. L. Rosenberg. Rapid identification of repeated patterns in strings, trees and arrays. In *Proceedings of the 4th ACM Symposium on the Theory of Computing*, pages 125–136, Denver, CO, 1972. ACM Press.

[1218] R. M. Karp and M. O. Rabin. Efficient randomized pattern-matching algorithms. *IBM J. Res. Dev.*, 31(2):249–260, 1987.

[1219] M. Karpinski. Alphabet-independent optimal parallel search for three-dimensional patterns. *Theor. Comput. Sci.*, 205(1/2):243–260, 1998.

[1220] M. Karpinski and W. Rytter. An alphabet-independent optimal parallel search for three dimensional patterns. Report 070, International Computer Science Institute, Berkeley, CA, 1993.

[1221] M. Karpinski and W. Rytter. An alphabet-independent optimal parallel search for three dimensional patterns. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 125–135, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1222] M. Karpinski and W. Rytter. Alphabet-independent optimal parallel search for three-dimensional patterns. *Theor. Comput. Sci.*, 205(1–2):243–260, 1998.

[1223] M. Karpinski, W. Rytter, and A. Shinohara. Pattern-matching for strings with short descriptions. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 205–214, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1224] M. Karpinski, W. Rytter, and A. Shinohara. An efficient pattern-matching algorithm for strings with short descriptions. *Nordic J. Comput.*, 4(2):172–186, 1997.

[1225] T. Kasami. An efficient recognition and syntax analysis algorithm for context-free languages. Report AFCRL-65-758, Air Force Cambridge Research Laboratory, Bedford, MA, 1963.

[1226] S. R. Kasaraju. Efficient string matching. Extended abstract, Department of Computer Science, John Hopkins University, 1988.

[1227] S. R. Kasaraju. Efficient tree pattern matching. In *Proceedings of the 30th IEEE Annual Symposium on Foundations of Computer Science*, pages 178–183, Research Triangle Park, North Carolina, 1989. IEEE Computer Society Press.

[1228] R. L. Kashyap and B. J. Oommen. An effective algorithm for string correction using generalized edit distances I: description of the algorithm and its optimality. *Inf. Sci.*, 23(2):123–142, 1981.

[1229] R. L. Kashyap and B. J. Oommen. A common basis for similarity and dissimilarity measures involving two strings. *Int. J. Comput. Math.*, 13:17–40, 1983.

[1230] R. L. Kashyap and B. J. Oommen. The noisy substring matching problem. *IEEE Trans. Softw. Eng.*, 9(3):365–370, 1983.

[1231] R. L. Kashyap and B. J. Oommen. Similarity measures for set of strings. *Int. J. Comput. Math.*, 13:95–104, 1983.

[1232] R. L. Kashyap and B. J. Oommen. String correction using probabilistic methods. *Pattern Recognit. Lett.*, pages 147–154, 1984.

[1233] J. Katajainen and T. Raita. An approximation algorithm for space-optimal encoding. *Comput. J.*, 32:228–237, 1989.

[1234] J. Katajainen and T. Raita. An analysis of the longest match and the greedy heuristics in text encoding. *J. Assoc. Comput. Mach.*, 39(2):281–294, 1992.

[1235] J. Kececioglu. *Exact and approximation algorithms for DNA sequence reconstruction.* Ph. D. Thesis, Report 91-26, Department of Computer Science, University of Arizona, Tucson, 1991.

[1236] J. Kececioglu. The maximum weight trace problem in multiple sequence alignment. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 106–119, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1237] J. Kececioglu, M. Li, and J. Tromp. Inferring a DNA sequence from erroneous copies. *Theor. Comput. Sci.*, 185(1):3–113, 1997.

[1238] J. Kececioglu and E. W. Myers. Combinatorial algorithms for DNA sequence assembly. Report TR92-37, Department of Computer Science, University of Arizona, Tucson, AZ (revised 15 January 1993), 1992.

[1239] J. Kececioglu and E. W. Myers. Combinatorial algorithms for DNA sequence assembly. *Algorithmica*, 13(1/2):7–51, 1995.

[1240] J. Kececioglu and R. Ravi. Of mice and men: Algorithms for evolutionary distances between genomes with translocation. In *Proceedings of the 6th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 604–613, San Francisco, CA, 1995.

[1241] J. Kececioglu and D. Sankoff. Exact and approximation algorithms for the inversion distance between two chromosomes. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 87–105, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1242] J. Kececioglu and D. Sankoff. Efficient bounds for oriented chromosome inversion distance. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 307–325, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1243] J. Kececioglu and D. Sankoff. Exact and approximation algorithms for sorting by reversals, with application to genome rearrangement. *Algorithmica*, 13(1/2):180–210, 1995.

[1244] J. Kececioglu and W. Zhang. Aligning alignments. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 189–208, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1245] Z. M. Kedem, G. M. Landau, and K. V. Palem. Optimal parallel suffix-prefix matching algorithms and applications. In *Proceedings of the 1st ACM Symposium on Parallel Algorithms and Architectures*, pages 388–398, Santa Fe, NM, 1989.

[1246] Z. M. Kedem, G. M. Landau, and K. V. Palem. Parallel suffix-prefix-matching algorithms and applications. *SIAM J. Comput.*, 25(5):998–1023, 1996.

[1247] Z. M. Kedem and K. V. Palem. Optimal parallel algorithms for forest and term matching. *Theor. Comput. Sci.*, 93(2):245–264, 1989.

[1248] R. Kemp. Generating words lexicographically: An average-case analysis. *Acta Inform.*, 35(1):17–89, 1998.

[1249] M. Kempf, R. Bayer, and U. Güntzer. Time optimal left to right construction of position trees. *Acta Inform.*, 24(4):461–474, 1987.

[1250] D. Kesner. Pattern matching in order-sorted languages. In A. Tarlecki, editor, *Proceedings 16th Symposium on Mathematical Foundations of Computer Science*, number 520 in Lecture Notes in Computer Science, pages 267–276, Kazimierz Dolny,Poland, 1991. Springer-Verlag, Berlin.

[1251] A. J. Kfoury. A linear-time algorithm to decide whether a binary word contains an overlap. *Bulletin of the European Association for Theoretical Computer Science*, 30:74–80, 1986.

[1252] A. J. Kfoury. A linear-time algorithm to decide whether a binary word contains an overlap. *RAIRO Inform. Théor. Appl.*, 22:135–145, 1988.

[1253] T. Kida, M. Takeda, A. Shinohara, and S. Arikawa. Shift-And approach to pattern matching in LZW compressed text. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 1–13, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1254] T. Kida, M. Takeda, A. Shinohara, M. Niyazaki, and S. Arikawa. Multiple pattern matching in LZW compressed text. *J. Discret. Algorithms*, 1(1):133–158, 2000.

[1255] P. Kilpeläinen. *Tree Matching Problems with Applications to Structured Text Databases*. Ph. d. thesis, University of Helsinki, Finland, 1992.

[1256] P. Kilpeläinen and H. Mannila. The tree inclusion problem. In S. Abramsky and T. S. E. Maibaum, editors, *Proceedings of the International Joint Conference on Theory and Practice of Software Development*, number 493 in Lecture Notes in Computer Science, pages 202–214, Brighton, U.K., 1991. Springer-Verlag, Berlin.

[1257] P. Kilpeläinen and H. Mannila. Grammatical tree matching. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 644 in Lecture Notes in Computer Science, pages 162–174, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1258] P. Kilpeläinen and H. Mannila. Query primitive for tree-structured data. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 213–225, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1259] D. K. Kim, Y. A. Kim, and K. Park. Constructing suffix arrays for multi-dimensional matrices. In M. Farach-Colton, editor, *Proceedings of* *the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 126–139, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1260] D. K. Kim, J. S. Lee, K. Park, and Y. Cho. Efficient algorithms for approximate string matching with swaps. *J. Complexity*, 15:128–147, 1999.

[1261] J.-Y. Kim and J. Shawe-Taylor. Fast expected string matching using an n-gram algorithm. Report CSD-TR-91-16, University of London, 1991.

[1262] J.-Y. Kim and J. Shawe-Taylor. An approximate string-matching algorithm. *Theor. Comput. Sci.*, 92(1):107–117, 1992.

[1263] J.-Y. Kim and J. Shawe-Taylor. Fast multiple keyword searching. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 41–51, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1264] J.-Y. Kim and J. Shawe-Taylor. Fast expected two dimensional pattern matching. In R. Baeza-Yates and N. Ziviani, editors, *Proceedings of the 1st South American Workshop on String Processing*, pages 77–92, 1993.

[1265] J.-Y. Kim and J. Shawe-Taylor. Fast string matching using n-gram. *Softw. Pract. Exp.*, 24(1):79–88, 1994.

[1266] S.-R. Kim and K. Park. A dynamic edit distance table. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 60–68, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1267] R. E. Kimbrell. Searching for text? send an N-gram! *Byte*, 13(5):297–312, 1988.

[1268] J.P. Kitajima, G. Navarro, B. Ribeiro-Neto, and N. Ziviani. Distributed generation of suffix arrays: A quicksort based approach. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 53–69, Valparaiso, Chile, 1997. Carleton University Press.

[1269] Y. Kitazume, E. Ohira, and T. Endo. LSI implementation of a pattern matching algorithm for speech recognition. *IEEE Trans. Acoust. Speech Signal Process.*, 33(1):1, 1985.

[1270] M. M. Klawe and D. J. Kleitman. An almost linear time algorithm for generalized matrix searching. Report, IBM Almaden Research Center, 1988.

[1271] P. Klein. Computing the edit-distance between unrooted ordered trees. In G. Bilardi, G. F. Italiano, A. Pietracaprina, and G. Pu cci, editors, *Proceedings of the 6th Annual European Symposium*, number 1461 in Lecture Notes in Computer Science, pages 91–102, Venice, Italy, 1998. Springer-Verlag, Berlin.

[1272] S. T. Klein. Space- and time-efficient decoding with canonical Huffman trees. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 65–75, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1273] S. T. Klein. Improving static compression schemes by alphabet extension. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 210–221, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1274] Shmuel T. Klein and Moti Meir. Delta encoding in a compressed domain. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 55–64, Czech Technical University in Prague, Czech Republic, 2009.

[1275] Shmuel T. Klein and Dana Shapira. Huffman coding with non-sorted frequencies. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 46–53, Czech Technical University in Prague, Czech Republic, 2008.

[1276] Shmuel T. Klein and Dana Shapira. On the usefulness of backspace. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 80–89, Czech Technical University in Prague, Czech Republic, 2009.

[1277] L. Knecht. Pairwise alignment with scoring on tuples. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 215–229, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1278] J. R. Knight and E. W. Myers. Approximate regular expression pattern matching with concave gap penalties. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 67–78, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1279] J. R. Knight and E. W. Myers. Approximate regular expression pattern matching with concave gap penalties. Report 92-12, Department of Computer Science, University of Arizona, Tucson, AZ, 1992.

[1280] J. R. Knight and E. W. Myers. Super-pattern matching. Report 29, Department of Computer Science, University of Arizona, Tucson, AZ, 1992.

[1281] J. R. Knight and E. W. Myers. Approximate regular expression pattern matching with concave gap penalties. *Algorithmica*, 14(1):85–121, 1995.

[1282] J. R. Knight and E. W. Myers. Super-pattern matching. *Algorithmica*, 13(1/2):211–243, 1995.

[1283] D. E. Knuth. *The art of computer programming: fundamental algorithms*, volume 1. Addison-Wesley, Reading, MA, 1973.

[1284] D. E. Knuth. *The art of computer programming: seminumerical algorithms*, volume 2. Addison-Wesley, Reading, MA, 1973.

[1285] D. E. Knuth. *The art of computer programming: Sorting and searching*, volume 3. Addison-Wesley, Reading, MA, 1973.

[1286] D. E. Knuth. Dynamic huffman coding. *J. Algorithms*, 6(2):163–180, 1985.

[1287] D. E. Knuth, J. H. Morris, Jr, and V. R. Pratt. Fast pattern matching in strings. *SIAM J. Comput.*, 6(1):323–350, 1977.

[1288] D. E. Knuth and M. F. Plass. Breaking paragraphs into lines. *Softw. Pract. Exp.*, 11(11):1119–1184, 1981.

[1289] D. E. Knuth and V. R. Pratt. Automata theory can be useful. Report, Stanford University, 1971.

[1290] K. I. Ko and C. M. Hua. A note on the two variable pattern finding problem. *J. Comput. Syst. Sci.*, 34(1):75–86, 1987.

[1291] T. Kohonen and E. Reuhkala. A very fast associative method for the recognition and correction of misspelt words, based on redundant hash-addressing. In *Proceedings of the 4th Joint Conference on Pattern Recognition*, pages 807–809, Kyoto, Japan, 1978.

[1292] R. Kolpakov and G. Kucherov. Maximal repetitions in words or how to find all squares in linear time. Report LORIA 98-R-227, Laboratoire Lorrain de Recherche en Informatique et ses Applications, 1998.

[1293] R. Kolpakov and G. Kucherov. Finding maximal repetitions in a word in linear time. In *Proceedings of the 40th IEEE Annual Symposium on Foundations of Computer Science*, pages 596–604, New York, 1999. IEEE Computer Society Press.

[1294] R. Kolpakov and G. Kucherov. On maximal repetitions in words. In G. Ciobanu and Gh. Păun, editors, *Proceedings of the 12th International Symposium on Fundamentals of Computation Theory*, volume 1684 of *Lecture Notes in Computer Science*, pages 374–385, Iasi, Romania, 1999. Springer-Verlag, Berlin.

[1295] R. Kolpakov and G. Kucherov. On the sum of exponents of maximal repetitions in a word. Report LORIA 99-R-034, Laboratoire Lorrain de Recherche en Informatique et ses Applications, 1999.

[1296] R. Kolpakov and G. Kucherov. On maximal repetitions in words. *J. Discret. Algorithms*, 1(1):159–186, 2000.

[1297] R. Kolpakov, G. Kucherov, and Y. Tarannikov. On repetition-free binary words of minimal density. In L. Brim, J. Gruska, and J. Zlatuska, editors, *Proceedings of the 23rdSymposium on Mathematical Foundations of Computer Science*, number 1450 in Lecture Notes in Computer Science, pages 683–692, Brno, Czech Republic, 1998. Springer-Verlag, Berlin.

[1298] R. Kolpakov, G. Kucherov, and Y. Tarannikov. On repetition-free binary words of minimal density. *Theor. Comput. Sci.*, 218(1):161–175, 1999.

[1299] G. A. Komatsoulis and M. S. Waterman. Chimeric alignment by dynamic programming: Algorithm and biological uses. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 174–180, Santa Fe, NM, 1997. ACM Press.

[1300] I. Kontoyiannis and Y. M. Suhov. Stationary entrophy estimation via string matching. In J. A. Storer and M. Cohn, editors, *Proceedings of the 6th Data Compression Conference*, page 444, Snowbird, UT, 1996. IEEE Computer Society Press.

[1301] D. G. Korn and K.-P. Vo. SFIO: Safe/fast string/file IO. In *Proceedings Summer USENIX Conference*, pages 235–256, Nashville, TN, 1991.

[1302] S. R. Kosaraju. Localized search in sorted lists. In *Proceedings of the 13th ACM Symposium on the Theory of Computing*, pages 62–69, Milwaukee, Wisconsin, 1981.

[1303] S. R. Kosaraju. Computation of squares in string. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 146–150, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1304] S. R. Kosaraju. Real-time pattern matching and quasi-real-time construction of suffix tree. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 310–316, Montréal, Canada, 1994. ACM Press.

[1305] S. R. Kosaraju and A. L. Delcher. Large-scale assembly of DNA strings and space-efficient construction of suffix trees. In *Proceedings of the 27th ACM Symposium on the Theory of Computing*, pages 169–177, Las Vegas, Nevada, 1995.

[1306] S. R. Kosaraju and A. L. Delcher. Correction: Large-scale assembly of DNA strings and space-efficient construction of suffix trees. In *Proceedings of the 28th ACM Symposium on the Theory of Computing*, page 659, Philadelphia, Pennsylvania, 1996.

[1307] S. R. Kosaraju, J.K. Park, and C. Stein. Long tours and short superstrings. In *Proceedings of the 35th IEEE Annual Symposium on Foundations of Computer Science*, pages 166–177, Santa Fe, NM, 1994. ACM Press.

[1308] G. Kowalski and A. Meltzer. New multi-term high speed text search algorithms. In *Proceedings of the 1st IEEE Conference on Computer Applications*, pages 514–522, 1984.

[1309] R. E. Krichevsky. Occam's razor, partially specified Boolean functions, string matching, and independent sets. *Inf. Comput.*, 108(1):158–174, 1994.

[1310] K. Krithivasan and R. Sitalakshmi. Efficient two-dimensional pattern matching in the presence of errors. *Inf. Sci.*, 43:169–184, 1987.

[1311] F.T. Krogh. Notes on a very fast substring search algorithm. *Commun. ACM*, 35(4):132, 1992.

[1312] J. B. Kruskal. An overview of sequence comparison: time warps, string edits, and macromolecules. *SIAM Rev.*, 25:201–237, 1983.

[1313] G. Kucherov and M. Rusinowitch. Matching a set of strings with variable length don't cares. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 230–247, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1314] Manfred Kufleitner. On bijective variants of the burrows-wheeler transform. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 65–79, Czech Technical University in Prague, Czech Republic, 2009.

[1315] K. Kukich. Techniques for automatically correcting words in text. *ACM Comp. Surv.*, 24(4):377–439, 1992.

[1316] M. Oğuzhan Külekci. Filter based fast matching of long patterns by using simd instructions. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 118–128, Czech Technical University in Prague, Czech Republic, 2009.

[1317] R. Kulkarni, S. Date, and B. Kulkarni. PRAS: Parallel Alignment of Sequences Algorithm. *Parallel Process. Lett.*, 3(3):243–252, 1993.

[1318] D. C. Kulp. Very fast pattern matching for highly repetitive text. Report, Department of Computer Science, University of Canterbury, 1992.

[1319] S. Kumar and E. Spafford. A pattern-matching model for intrusion detection. In *Proceedings National Computer Security Conference*, pages 11–21, 1994.

[1320] S. K. Kumar and C. P. Rangan. A linear-space algorithm for the LCS problem. *Acta Inform.*, 24(3):353–362, 1987.

[1321] S. Kuo and G. R. Cross. An improved algorithm to find the length of the longest common subsequence of two strings. *SIGIR Forum*, 23, 1989.

[1322] S. Kuo and G. R. Cross. A two–step string–matching procedure. *Pattern Recognit.*, 24:711–716, 1991.

[1323] S. Kurtz. Approximate string searching under weighted edit distance. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 156–170, Recife, Brazil, 1996. Carleton University Press.

[1324] S. Kurtz. Reducing the space requirements of suffix trees. *Softw. Pract. Exp.*, 29(13):1149–1171, 1999.

[1325] S. Kurtz and G. Myers. Estimating the probability of approximate matches. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 52–64, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1326] Kazuhiko Kusano, Wataru Matsubara, Akira Ishino, and Ayumi Shinohara. Average value of sum of exponents of runs in strings. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 185–192, Czech Technical University in Prague, Czech Republic, 2008.

[1327] R. E. Ladner and M. J. Fisher. Parallel prefix computation. *J. Assoc. Comput. Mach.*, 27(4):831–838, 1980.

[1328] K. B. Lakshmanan. On universal codeword sets. *IEEE Trans. Inf. Theory*, 27:659–662, 1981.

[1329] G. Lancia and R. Ravi. GESTALT: Genomic steiner alignments. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 101–114, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1330] J. K. Lanctot, M. Li, and E. Yang. Estimating DNA sequence entropy. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, San Francisco, CA, 2000.

[1331] G. M. Landau. *String matching in erroneus input.* Ph. D. Thesis, Department of Computer Science, Tel-Aviv University, 1986.

[1332] G. M. Landau, E. W. Myers, and J. P. Schmidt. Incremental string comparison. *SIAM J. Comput.*, 27(2):557–582, 1998.

[1333] G. M. Landau, B. Schieber, and U. Vishkin. Parallel construction of a suffix tree. In T. Ottmann, editor, *Proceedings of the 14th International Colloquium on Automata, Languages and Programming*, number 267 in Lecture Notes in Computer Science, pages 314–325, Karlsruhe, Germany, 1987. Springer-Verlag, Berlin.

[1334] G. M. Landau and J. P. Schmidt. An algorithm for approximate tandem repeats. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 120–133, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1335] G. M. Landau and U. Vishkin. Efficient string matching in the presence of errors. In *Proceedings of the 26th IEEE Annual Symposium on Foundations of Computer Science*, pages 126–136, Portland, Oregon, 1985. IEEE Computer Society Press.

[1336] G. M. Landau and U. Vishkin. Efficient parallel and serial approximate string matching. Report TR-221-U101, Courant Institut of Mathematical Sciences, New York University, NY, 1986.

[1337] G. M. Landau and U. Vishkin. Efficient string matching with k mismatches. *Theor. Comput. Sci.*, 43(2–3):239–249, 1986.

[1338] G. M. Landau and U. Vishkin. Introducing efficient parallelism into approximate string matching and a new serial algorithm. In *Proceedings of the 18th ACM Symposium on the Theory of Computing*, pages 220–230, Berkeley, CA, 1986. ACM Press.

[1339] G. M. Landau and U. Vishkin. Fast string matching with k differences. *J. Comput. Syst. Sci.*, 37(1):63–78, 1988.

[1340] G. M. Landau and U. Vishkin. Fast parallel and serial approximate string matching. *J. Algorithms*, 10(2):157–169, 1989.

[1341] G. M. Landau and U. Vishkin. Two dimensional pattern matching in digitized image. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 134–151, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1342] G. M. Landau and U. Vishkin. Pattern matching in digitalized image. *Algorithmica*, 12(4/5):375–408, 1994.

[1343] G. M. Landau and U. Vishkin. Approximate string searching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 6, pages 185–199. Oxford University Press, 1997.

[1344] G. M. Landau, U. Vishkin, and R. Nussinov. An efficient string matching algorithm with k differences for nucleotide and amino acid sequences. Report TR-37/85, Department of Computer Science, Tel Aviv University, 1985.

[1345] G. M. Landau, U. Vishkin, and R. Nussinov. Locating alignments with k differences for nucleotide and amino acid sequences. *Comput. Appl. Biosci.*, 4(1):19–24, 1988.

[1346] G. M. Landau and M. Ziv-Ukelson. On the shared substring alignment problem. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 804–814, San Francisco, CA, 2000.

[1347] E. Lander, J. P. Mesirov, and W. Taylor. Protein sequence comparison on a data parallel computer. In D. H. Bayley, editor, *Proceedings of the 1988 International Conference on Parallel Processing, (Algorithms and Applications)*, volume 3, pages 257–263, Pennsylvania State University Press, PA, 1988.

[1348] E. Lander, J. P. Mesirov, and W. Taylor, IV. Study of protein sequence comparison metrics on the Connection Machine CM-2. In J. L. Martin and S. F. Lundstrom, editors, *Proceedings Supercomputing '88*, volume 2, pages 2–12, Orlando, FL, 1988. IEEE Computer Society Press.

[1349] E. Lander, J. P. Mesirov, and W.Taylor, IV. Study of protein sequence comparison metrics on the Connection Machine CM-2. *J. Supercomputing*, 3(4):255–269, 1989.

[1350] S. Lange and R. Wiehagen. Polynomial-time inference of arbitrary pattern languages. *New Gener. Comput.*, 8(4):361–370, 1991.

[1351] L. L. Larmore. An optimal algorithm with unknown time complexity for convex matrix searching. *Inf. Process. Lett.*, 6(3):147–151, 1990.

[1352] L. L. Larmore and W. Rytter. Efficient sublinear time parallel algorithms for dynamic programming and context-free recognition. In A. Finkel and M. Jantzen, editors, *Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, number 577 in Lecture Notes in Computer Science, pages 121–132, Cachan, France, 1992. Springer-Verlag, Berlin.

[1353] L. L. Larmore and W. Rytter. Almost optimal sublinear time parallel recpgnition algorithms for three subclasses of context free languages. *Theor. Comput. Sci.*, 197(1–2):189–201, 1998.

[1354] L. L. Larmore and B. Schieber. On-line dynamic programming with applications to the prediction of RNA secondary structure. In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 503–512, San Francisco, CA, 1990.

[1355] N. J. Larsson. Extended application of suffix trees to data compression. In J. A. Storer and M. Cohn, editors, *Proceedings Data Compression Conference*, pages 190–199, Snowbird, UT, 1996. IEEE Computer Society Press.

[1356] R. H. Lathrop, T. A. Webster, and T. F. Smith. ARIADNE: pattern-directed inference and hierarchical abstraction in protein structure recognition. *Commun. ACM*, 30(11):909–921, 1987.

[1357] D. Lavenier. *MicMacs : un Réseau Systolique pour le Traitement de Chaînes de Caractères*. PhD thesis, Univ. Rennes 1, France, 1989.

[1358] D. Lavenier. An Integrated 2D Systolic Array for Spelling Correction. *Integr., The VLSI J.*, 15, 1993.

[1359] D. Lavenier. Dedicated Hardware for Biological Sequence Comparison. *J. Univ. Comp. Sci.*, 2(2), 1996.

[1360] D. Lavenier. SAMBA: Systolic Accelerator for Molecular Biological Applications. Rapport IRISA 2845, 1996.

[1361] C. E. Lawrence, S. F. Altschul, M. S. Boguski, J. S. Liu, A. F. Neuwald, and J. C. Wootton. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science*, 262:208–214, 1993.

[1362] T. Lecroq. *Recherches de mot*. Thèse de doctorat, Université d'Orléans, France, 1992.

[1363] T. Lecroq. Recherches de mot. Rapport LITP 92.15, Universités Paris 6 et 7, France, 1992.

[1364] T. Lecroq. A variation on the Boyer-Moore algorithm. *Theor. Comput. Sci.*, 92(1):119–144, 1992.

[1365] T. Lecroq. Experimental results on string matching algorithms. Rapport LITP 94.12, Universités Paris 6 et 7, France, 1994.

[1366] T. Lecroq. Experimental results on string matching algorithms. *Softw. Pract. Exp.*, 25(7):727–765, 1995.

[1367] T. Lecroq. Experimental results on string matching over infinite alphabets. Rapport LIR 97.01, Université de Rouen, France, 1997.

[1368] T. Lecroq. Experiments on string matching in memory structures. *Softw. Pract. Exp.*, 28(5):561–568, 1998.

[1369] T. Lecroq. Theoretical and practical aspects of string matching. In *Abstracts of invited lectures and short communcations deliverered at the 7th International Colloquium on Numerical Analysis and Computer Science with Applications*, page 75, Plovdiv, Bulgaria, 1998.

[1370] T. Lecroq. Theoretical and practical string matching. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, page 2, Perth, WA, Australia, 1999.

[1371] T. Lecroq. New experimental results on exact string-matching. Rapport LIFAR 2000.03, Université de Rouen, France, 2000.

[1372] T. Lecroq. *Quelques aspects de l'algorithmique du texte*. Habilitation à diriger des recherches, Université de Rouen, France, 2000.

[1373] T. Lecroq. Une introduction à la recherche de mot. Actes de l'École Jeunes Chercheurs Algorithmique et Calcul Formel du GDR ALP, Caen, France, 2000.

[1374] T. Lecroq, G. Luce, and J.-F. Myoupo. A faster linear systolic algorithm for recovering a longest common subsequence. Rapport LaRIA 96–07, Université de Picardie Jules Verne, France, 1996.

[1375] T. Lecroq, G. Luce, and J.-F. Myoupo. A faster linear systolic algorithm for recovering a longest common subsequence. *Inf. Process. Lett.*, 61(3):129–136, 1997.

[1376] T. Lecroq and J.-F. Myoupo. Détermination d'un plus long sous-mot commun à deux mots sur un réseau linéaire. In L. Bougé, M. Cosnard, and P. Fraigniaud, editors, *Proceedings of the 6th Rencontres Francophones du Parallélisme*, page 301, Lyon, France, 1994.

[1377] T. Lecroq and J.-F. Myoupo. Détermination d'un plus long sous-mot commun à deux mots sur un réseau linéaire. Rapport LIR95.08 Informatique Fondamentale, Université de Rouen, France, 1995.

[1378] T. Lecroq, J.-F. Myoupo, and D. Semé. A one-phase parallel algorithm for the sequence alignment problem. Rapport LaRIA 97–16, Université de Picardie Jules Verne, France, 1997.

[1379] T. Lecroq, J.-F. Myoupo, and D. Semé. A one-phase parallel algorithm for the sequence alignment problem. *Parallel Process. Lett.*, 8(4):515–526, 1998.

[1380] D. Lee and F. Lochovsky. *Text retrieval machine*. Springer-Verlag, Berlin, 1985.

[1381] J.-S. Lee, D. K. Kim, K. Park, and Y. Cho. Efficient algorithms for approximate string matching with swaps. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 28–39, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1382] K. C. Lee and V. W. Mak. Design and analysis of a parallel VLSI string search algorithm. In H. Boral and P. Faudemay, editors, *Proceedings of the 6th International Workshop on Database Machines*, number 368 in Lecture Notes in Computer Science, pages 215–229, Deauville, France, 1989. Springer-Verlag, Berlin.

[1383] K.-F. Lee. *Automatic Speech Recognition.* Kluwer, 1989.

[1384] A. Lefebvre and T. Lecroq. Computing repeated factors with a factor oracle. In L. Brankovic and J. Ryan, editors, *Proceedings of the 11th Australasian Workshop On Combinatorial Algorithms*, pages 145–158, Hunter Valley, Australia, 2000.

[1385] A. Lefebvre and T. Lecroq. Estimating topological entropy of biological sequences using a factor oracle. In *Proccedings of the 5th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI 2001)*, Orlando, Florida, 2001.

[1386] A. Lefebvre and T. Lecroq. Compror: on-line lossless data compression with a factor oracle. *Inf. Process. Lett.*, 83(1):1–6, 2002.

[1387] A. Lefebvre and T. Lecroq. A heuristic for computing repeats with a factor oracle: Applications to biological sequences. *Int. J. Comput. Math.*, 79(12):1303–1315, 2002.

[1388] A. Lefebvre, T. Lecroq, and J. Alexandre. Utilisations de l'oracle des facteurs. In L. Duret, C. Gaspin, and T. Schiex, editors, *Actes des Journées Ouvertes Biologie Informatique Mathématiques*, pages 153–157, Toulouse, France, 2001.

[1389] A. Lefebvre, T. Lecroq, and J. Alexandre. Drastic improvements over repeats found with a factor oracle. In E. Billington, D. Donovan, and A. Khodkar, editors, *Proceedings of the 13th Australasian Workshop On Combinatorial Algorithms*, pages 253–265, Fraser Island, Australia, 2002.

[1390] A. Lefebvre, T. Lecroq, H. Dauchel, and J. Alexandre. FORRepeats: detects repeats on entire chromosomes and between genomes. *Bioinformatics*, 19(3):319–326, 2003.

[1391] A. Lempel, S. Even, and M. Cohn. An algorithm for optimal prefix parsing of a noiseless and memoryless channel. *IEEE Trans. Inf. Theory*, 19:208–214, 1973.

[1392] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inf. Theory*, 22:75–81, 1976.

[1393] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inf. Theory*, 22(1):75–81, 1976.

[1394] A. Lesk. *Computational molecular biology, sources and methods for sequence analysis.* Oxford University Press, 1988.

[1395] A. M. Lesk. Three-dimensional pattern matching in protein structure analysis. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 248–260, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1396] A. M. Lesk. Application of sequence alignment methods to multiple structural alignment. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 95–100, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1397] A. M. Lesk, M. Levitt, and C. Chothia. Alignment of amino acid sequences of distantly related proteins using variable gap penalties. *Protein Engineering*, 1:77–78, 1986.

[1398] L. Lestrée. *Recherche de mots et d'arbres en parallèle.* Thèse de doctorat, Université Paris 7, 1993.

[1399] L. Lestrée. Unit route upper bound for string-matching on hypercube. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 136–145, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1400] Peter Leupold. Reducing repetitions. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 225–236, Czech Technical University in Prague, Czech Republic, 2009.

[1401] V. I. Levenshtein. Binary codes capable of correcting spurious insertions and deletions of ones. *Problems of Information Transmission*, 1:8–17, 1965.

[1402] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Sov. Phys. Dokl.*, 6:707–710, 1966.

[1403] H. Li, M. Lu, and J. Fang. An optimal algorithm for the longest common subsequence problem. In *Proceedings of the 3rd IEEE Symposium on Parallel and Distributed Processing*, pages 630–639, Los Alamitos, 1991. IEEE Computer Society Press.

[1404] M. Li. Lower bounds on string matching. Report TR-84-636, Department of Computer Science, Cornell University, Ithaca, 1984.

[1405] M. Li. Towards a DNA sequencing theory. In *Proceedings of the 31st IEEE Annual Symposium on Foundations of Computer Science*, pages 125–134, Saint Louis, MO, 1990. IEEE Computer Society Press.

[1406] M. Li and T. Jiang. Shortest common superstrings. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 8, pages 238–265. Oxford University Press, 1997.

[1407] M. Li, B. Ma, and L. Wang. Finding similar regions in many strings. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 473–482, 1999.

[1408] M. Li and Y. Yesha. String matching cannot be done by a two-head one way deterministic finite automaton. *Inf. Process. Lett.*, 22(5):231–235, 1986.

[1409] F. M. Liang. *Word hy-phen-a-tion by com-pu-ter*. Ph. D. Thesis, Computer Science Department, Stanford University, 1983.

[1410] D. Liben-Nowell. On the structure of syntenic distance. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 50–65, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1411] D. Liben-Nowell and J. M. Kleinberg. Structural properties and tractability results for linear synteny. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 248–263, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1412] Jie Lin, Yue Jiang, and Don Adjeroh. The virtual suffix tree: An efficient data structure for suffix trees and suffix arrays. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 68–83, Czech Technical University in Prague, Czech Republic, 2008.

[1413] Y. C. Lin. New systolic arrays for the longest subsequence problem. *Parallel Comput.*, 20:1323–1334, 1994.

[1414] D. C. Lindsay. A file difference program. *Commun. ACM*, 32(6):740–755, 1989.

[1415] D. J. Lipman. Computing discoveries in molecular biology (abstract). In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, page 64, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1416] D. J. Lipman, S. F. Altschul, and J. D. Kececioglu. A tool for multiple sequence alignment. *Proc. Natl. Acad. Sci. U.S.A.*, 86:4412–4415, 1989.

[1417] D. J. Lipman and W. R. Pearson. Rapid and sensitive protein similarity searches. *Sci.*, 227:1435–1441, 1985.

[1418] R. S. Lipton and D. P. Lopresti. A systolic array for rapid string comparison. In H. Fuchs, editor, *Proceedings of the 1985 Chapel Hill Conference on Very Large Scale Integration*, pages 363–376, Rockville, MD, 1985. Computer Science Press.

[1419] R. S. Lipton and D. P. Lopresti. Comparing long strings on a short systolic array. Report CS-TR-026-86, Princeton University, Princeton, NJ, 1986.

[1420] R. S. Lipton and D. P. Lopresti. Comparing long strings on a short systolic array. In W. Moore, A. McCabe, R. Urquhart, and A. Hilger, editors, *Systolic Arrays*. Computer Science Press, 1986.

[1421] R. S. Lipton and D. P. Lopresti. Comparing long strings on a short systolic array. In W. Moore, A. McCabe, R. Urquart, and A. Hilger, editors, *Proceedings of the 1st International Workshop on Systolic Arrays*, pages 181–190, Bristol, 1987.

[1422] H. H. Liu and K. S. Fu. VLSI arrays for minimum-distance classifications. In K. S. Fu, editor, *VLSI for Pattern Recognition and Image Processing*, number 13 in Springer Series in Information Sciences, pages ?–? Springer-Verlag, Berlin, 1984.

[1423] J. S. Liu, A. F. Neuwald, and C. E. Lawrence. Bayesian models for multiple local sequence alignment and Gibbs sampling strategies. *J. American Statistical Association*, 90(432):1156–??, 1995.

[1424] J. S. Liu, A. F. Neuwald, and C. E. Lawrence. Markovian structures in biological sequence alignment. *J. American Statistical Association*, 94(445):1–??, 1999.

[1425] K.-C. Liu. On string pattern matching: A quantitative analysis and a proposal. *Computer Languages*, 13(1):23–29, 1988.

[1426] L. Liu. On string pattern matching: a new model with a polynomial time algorithm. *SIAM J. Comput.*, 10(1):118–139, 1981.

[1427] Z. Liu, X. Du, and N. Ishii. An improved adaptive string searching algorithm. *Softw. Pract. Exp.*, 28(2):191–198, 1998.

[1428] E. Lodi. A parallel solution to the approximate string matching problem. *Comput. J.*, 35(5):524–526, 1992.

[1429] D. M. Loewenstern and P. N. Yianilos. Significantly lower entropy estimates for natural DNA sequences. Technical Report 96-51, DIMACS, 1996.

[1430] R. K. Loke and B. J. Oommen. Noisy subsequence recognition using constrained string editing involving substitutions, insertions, deletions and generalized transpositions. In *Proceedings of the 3rd Int. Comp. Sci. Conf.*, number 1024 in Lecture Notes in Computer Science, pages 116–123. Springer-Verlag, Berlin, 1995.

[1431] S. Lonardi. *Off-Line Data Compression by Textual Substitution.* PhD thesis, Università di Padova, Dipartimento di Ingegneria Elettronica e Informatica, 1999.

[1432] D. P. Lopresti. *Discounts for Dynamic Programming with Applications in VLSI Process or Arrays.* PhD thesis, Princeton University, 1987.

[1433] D. P. Lopresti. P-NAC: a systolic array for comparing nucleic acid sequences. *Comput.*, 20(7):98–99, 1987.

[1434] D. P. Lopresti. A systolic array for comparing nucleic acid sequences. *IEEE Comp. Mag.*, 20:98–99, 1987.

[1435] D. P. Lopresti. Rapid implementation of a genetic sequence comparator using FPGAs. *Adv. Res. VLSI*, pages 139–152, 1991.

[1436] D. P. Lopresti and A. Tomkins. Block edit models for approximate string matching. *Theor. Comput. Sci.*, 181(1):159–179, 1997.

[1437] M. Lothaire, editor. *Combinatorics on Words.* Addison-Wesley, Reading, MA, 1983.

[1438] M. Lothaire, editor. *Combinatorics on Words.* Cambridge University Press, second edition, 1997.

[1439] G. Louchard and W. Szpankowski. A probabilistic analysis of a string edit problem. Rapport 1814, INRIA, 1992.

[1440] G. Louchard and W. Szpankowski. Analysis of a string edit problem in a probabilistic framework. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 152–163, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1441] R. Lowrance and R. A. Wagner. An extension of the string-to-string correction problem. *J. Assoc. Comput. Mach.*, 22(2):177–183, 1975.

[1442] M. Lu. Parallel computation of longest-common-subsequence. In S.G. Akl, F. Fiala, and W.W. Koczkodaj, editors, *Proceedings Advances in Computing and Information*, number 468 in Lecture Notes in Computer Science, pages 383–394. Springer-Verlag, Berlin, 1990.

[1443] M. Lu and H. Lin. Parallel algorithms for the longest common subsequence problem. *IEEE Trans. Parallel Distr. Syst.*, 5(8):835–848, 1994.

[1444] S.-Y. Lu. A tree-to-tree distance and its application to cluster analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1(2):219–224, 1979.

[1445] S.-Y. Lu. A tree-matching algorithm based on node splitting and merging. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(2):249–256, 1984.

[1446] S. Lucci. *String matching: A Comparative Study of Algorithms, and its Relation to Problems of Parallel and Distributed Computing*. PhD thesis, City University of New york, 1987.

[1447] F. Luccio and L. Pagli. An efficient algorithm for some tree matching problems. *Inf. Process. Lett.*, 39(1):51–57, 1991.

[1448] F. Luccio and L. Pagli. Simple solutions for approximate tree matching problems. In S. Abramsky and T. S. E. Maibaum, editors, *Proceedings of the International Joint Conference on Theory and Practice of Software Development*, number 493 in Lecture Notes in Computer Science, pages 193–201, Brighton, U.K., 1991. Vol I: Colloqium on Trees in Algebra and Programming, Springer-Verlag, Berlin.

[1449] F. Luccio and L. Pagli. Approximate matching for two families of trees. *Inf. Comput.*, 123(1):111–120, 1995.

[1450] G. Luce. Réduction de la Complexité d'un Assouplissement des Règles de Mutations en Alignement de Séquences Biologiques, 1994. Rapport de DEA, Université de Rouen, France.

[1451] G. Luce. *Algorithmique systolique pour le problème du plus long sous-mot commun et implémentation sur un réseau de transputers*. PhD thesis, Université d'Amiens, France, 1997.

[1452] G. Luce and J.-F. Myoupo. An efficient linear systolic algorithm for recovering longest common subsequences. Rapport LAMIFA 95-06, Université d'Amiens, France, 1995.

[1453] G. Luce and J.-F. Myoupo. An efficient linear systolic array for recovering longest common subsequences. In *Proceedings of the 1st IEEE Int. Conf. Algo. Archi. Parall. Proc.*, volume 1, pages 20–29, 1995.

[1454] G. Luce and J.-F. Myoupo. Systolic-based parallel architecture for the longest common subsequence problem. Rapport LAMIFA 95-15, Université d'Amiens, France, 1995.

[1455] G. Luce and J.-F. Myoupo. Efficient Parallel Computing of a Longest Common Subsequence of Three Strings. In *Proceedings of the 8th SIAM Conf. Parall. Proc. Sci. Comp.*, 1997.

[1456] T. Luczak and W. Szpankowski. A lossy data compression based on string matching: Preliminary analysis and suboptimal algorithms. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 102–112, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1457] T. Luczak and W. Szpankowski. A suboptimal lossy data compression based on approximate pattern matching. *IEEE Trans. Inf. Theory*, 43:1439–1451, 1997.

[1458] R. C. Lyndon and M. P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *Mich. Math. J.*, 9:177–183, 1962.

[1459] G. Lyon. Syntax-directed least-errors analysis for context-free languages: a practical approach. *Commun. ACM*, 17(1):3–14, 1974.

[1460] R. M. and W. I. Grosky. Shape matching utilizing indexed hypotheses generation and testing. *IEEE J. Robot. Automat.*, 5(1):70–77, 1989.

[1461] B. Ma. A polynominal time approximation scheme for the closest substring problem. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 99–107, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1462] B. Ma, L. Wang, and M. Li. Fixed topology alignment with recombination. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 174–188, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1463] M. G. Maaß. Linear bidirectional on-line construction of affix trees. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 320–334, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1464] K. Madlener and F. Otto. Using string-rewriting for solving the word problem for finitely presented groups. *Inf. Process. Lett.*, 24(5):281–284, 1987.

[1465] M. Maes. On a cyclic string to string correction problem. *Inf. Process. Lett.*, 35(2):73–78, 1990.

[1466] H. M. Mahmoud, M. Régnier, and R. T. Smythe. Analysis of Boyer-Moore-Horspool string-matching heuristic. Report 9634, The George Washington University, 1996.

[1467] H. M. Mahmoud and R. T. Smythe. An asymptotically-optimal variation of the Knuth-Morris-Pratt string matching algorithm. Report 95-4, The George Washington University, 1995.

[1468] D. Maier. The complexity of some problems on subsequences and supersequences. *J. Assoc. Comput. Mach.*, 25(2):322–336, 1978.

[1469] D. Maier and J. A. Storer. A note on the complexity of superstring problem. Report 233, Computer Science Laboratory, Princeton University, 1977.

[1470] M. G. Main. Detecting leftmost maximal periodicities. *Discret. Appl. Math.*, 25:145–153, 1989.

[1471] M. G. Main and R. J. Lorentz. An $O(n \log n)$ algorithm for recognizing repetition. Report CS-79-056, Washington State University, Pullman, 1979.

[1472] M. G. Main and R. J. Lorentz. An $O(n \log n)$ algorithm for finding all repetitions in a string. *J. Algorithms*, 5(3):422–432, 1984.

[1473] M. G. Main and R. J. Lorentz. Linear time recognition of square free strings. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 272–278. Springer-Verlag, Berlin, 1985.

[1474] H. G. Mairson. Generating words in a context-free language uniformly at random. *Inf. Process. Lett.*, 49(2):95–99, 1994.

[1475] M. E. Majster and A. Ryser. Efficient on-line construction and correction of position trees. *SIAM J. Comput.*, 9(4):785–807, 1980.

[1476] G. S. Makanin and H. Abdulrab. On general solution of word equations. Rapport LITP 93.42, Universités Paris 6 et 7, 1993.

[1477] G. S. Makanin and H. Abdulrab. Parametric word equations. Rapport LITP 93.43, Universités Paris 6 et 7, 1993.

[1478] G. S. Makanin and H. Abdulrab. Transformation algorithm for word equations in four variables. Rapport LITP 93.49, Universités Paris 6 et 7, 1993.

[1479] G. S. Makanin and H. Abdulrab. Transformations of word equations in three variables = rouen's function. Rapport LITP 93.44, Universités Paris 6 et 7, 1993.

[1480] G. S. Makanin and H. Abdulrab. On general solution of word equations. In *Results and Trends in Theoretical Computer Science*, number 812 in Lecture Notes in Computer Science, pages 251–263. Springer-Verlag, Berlin, 1994.

[1481] G. S. Makanin, H. Abdulrab, and M. N. Maksimenko. Bunches of formal parametric equations. Rapport LITP 93.62, Universités Paris 6 et 7, 1993.

[1482] E. Mäkinen. On the subtree isomorphism problem for ordered trees. *Inf. Process. Lett.*, 32(5):271–273, 1989.

[1483] V. Mäkinen. Compact suffix array. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 305–319, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1484] M. N. Maksimenko. Algorithme quadratique de calcul de la solution générale d'équations en mots à une variable. Rapport LITP 93.14, Universités Paris 6 et 7, 1993.

[1485] M. N. Maksimenko. Algorithme quadratique de calcul de la solution générale d'équations en mots à une variable. *RAIRO Inform. Théor.*, 29(4):277–284, 1993.

[1486] K. Maly. Compressed tries. *Commun. ACM*, 19(7):409–415, 1976.

[1487] G. Manacher. A new linear-time on-line algorithm finding the smallest initial palindrome of a string. *J. Assoc. Comput. Mach.*, 22(3):346–351, 1975.

[1488] G. Manacher. An application of pattern matching to a problem in geometrical complexity. *Inf. Process. Lett.*, 5(1):6–7, 1976.

[1489] U. Manber. *Introduction to algorithms.* Addison-Wesley, Reading, MA, 1989.

[1490] U. Manber. Finding similar files in a large file system. In *Proceedings of the USENIX Winter 1994 Technical Conference*, pages 1–10, San Francisco, CA, 1994.

[1491] U. Manber. A text compression scheme that allows fast searching directly in the compressed file. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 113–124, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1492] U. Manber. A text compression scheme that allows fast searching directly in the compressed file. *ACM Trans. Inform. Syst.*, 15(2):124–136, 1997.

[1493] U. Manber and R. A. Baeza-Yates. An algorithm for string matching with a sequence of don't cares. *Inf. Process. Lett.*, 37(3):133–136, 1991.

[1494] U. Manber and G. Myers. Suffix arrays: a new method for on-line string searches. In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 319–327, San Francisco, CA, 1990.

[1495] U. Manber and G. Myers. Suffix arrays: a new method for on-line string searches. *SIAM J. Comput.*, 22(5):935–948, 1993.

[1496] U. Manber and S. Wu. An algorithm for approximate string matching with non uniform costs. Report TR-89-19, Department of Computer Science, University of Arizona, Tucson, AZ, 1989.

[1497] U. Manber and S. Wu. Approximate string-matching with arbitrary costs for text and hypertext. In *Proceedings of the IAPR Workshop on Structural and Syntactic Pattern Recognition*, pages 22–33, Bern, Switzerland, 1992.

[1498] U. Manber and S. Wu. Some assembly required: Approximate pattern matching. *Byte*, 17(12):281–292, 1992.

[1499] U. Manber and S. Wu. GLIMPSE: A tool to search through entire file systems. In *Proceedings of the USENIX Winter 1994 Technical Conference*, pages 23–32, San Francisco, CA, 1994.

[1500] S. Marcus, A. Mateescu, G. Paun, and A. Salomaa. On symmetry in strings, sequences and languages. *Int. J. Comput. Math.*, 54(1–2), 1994.

[1501] J. Markovitz, T. Ahlswede, and M. Evans. Semantically significant patterns in dictionary definitions. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 112–119, 1986.

[1502] L. Marsan and M.-F. Sagot. Algorithms for extracting structured motifs using a suffix tree with application to promoter and regulatory site consensus identification. *J. Comput. Bio.*, 7(3/4):345–360, 2000.

[1503] L. Marsan and M.-F. Sagot. Extracting structured motifs using a suffix tree - Algorithms and application to consensus identification. In S. Minoru and R. Shamir, editors, *Proceedings of the 4th Annual International Conference on Computational Molecular Biology (RECOMB)*, Tokyo, Japan, 2000. ACM Press.

[1504] J. C. Martin. Generalized morse sequences on n symbols. *Proc. Amer. Math. Soc.*, 54:379–383, 1976.

[1505] P. Martin-Löf. The definition of random sequences. *Inf. Control*, 9(2):602–619, 1966.

[1506] H. M. Martinez. An efficient method for finding repeats in molecular sequences. *Nucleic Acids Res.*, 11(13):4629–4634, 1983.

[1507] H. M. Martinez. A flexible multiple sequence alignment program. *Nucleic Acids Res.*, 16(5A):1683–1691, 1988.

[1508] A. Marzal and E. Vidal. Computation of normalized edit distance and apllications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):926–932, 1993.

[1509] W. J. Masek and M. S. Paterson. A faster algorithm for computing string edit distances. *J. Comput. Syst. Sci.*, 20(1):18–31, 1980.

[1510] W. J. Masek and M. S. Paterson. How to compute string-edit distances quickly. In D. Sankoff and J. B. Kruskall, editors, *Time Warps, String Edits, and Macromolecules: the Theory and Practice of Sequence Comparison*, chapter 14, pages 337–349. Addison-Wesley, Reading, MA, 1983.

[1511] A. Mateescu and A. Salomaa. Nondeterminism in patterns. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 661–668, Caen, France, 1994. Springer-Verlag, Berlin.

[1512] Y. Matias, S. Muthukrishnan, S. C. Sahinalpk, and J. Ziv. Augmenting suffix trees with applications. In G. Bilardi, G. F. Italiano, A. Pietracaprina, and G. Pucci, editors, *Proceedings of the 6th Annual European Symposium*, number 1461 in Lecture Notes in Computer Science, pages 67–78, Venice, Italy, 1998. Springer-Verlag, Berlin.

[1513] Wataru Matsubara, Kazuhiko Kusano, Akira Ishino, Hideo Bannai, and Ayumi Shinohara. New lower bounds for the maximum number of runs in a string. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 140–145, Czech Technical University in Prague, Czech Republic, 2008.

[1514] A. Mayne and E. B. James. Information compression by factorising common strings. *Comput. J.*, 18:157–160, 1975.

[1515] E. M. McCreight. A space-economical suffix tree construction algorithm. *J. Algorithms*, 23(2):262–272, 1976.

[1516] C. A. Mead, R. D. Pashley, L. D. Britton, Y. T. Daimon, and S. F. Sando. 128-bit multicomparator. *IEEE J. Solid-State Circuits*, 11(5):692–695, 1976.

[1517] G. Mehldau and G. Myers. A system for pattern matching applications on biosequences. Report TR-91-31, Department of Computer Science, University of Arizona, Tucson, AZ, 1991.

[1518] G. Mehldau and G. Myers. A system for pattern matching applications on biosequences. *Comput. Appl. Biosci.*, 9(3):299–314, 1991.

[1519] K. Mehlhorn. *Data structures and algorithms 1: sorting and searching.* Springer-Verlag, Berlin, 1984.

[1520] D. P. Mehta and S. Sahni. Computing display conflicts in string visualization. Report 024, University of Florida, 1991.

[1521] D. P. Mehta and S. Sahni. A data structure for circular string analysis and visualization. Report 025, University of Florida, 1991.

[1522] D. P. Mehta and S. Sahni. Computing display conflicts in string and circular string visualization. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 244–261, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1523] D. P. Mehta and S. Sahni. A data structure for circular string analysis and visualization. *IEEE Trans. Comput.*, 42(8):992–997, 1993.

[1524] D. P. Mehta and S. Sahni. Computing display conflicts in string visualization. *IEEE Trans. Comput.*, 43(3):350–361, 1994.

[1525] J. Meidanis. Distance and similarity in the presence of nonincreasing gap-weighting functions. In R. Baeza-Yates and U. Manber, editors, *Proceedings of the 2nd South American Workshop on String Processing*, pages 27–37, Valparaíso, Chile, 1995.

[1526] J. Meidanis and J. C. Setubal. Multiple alignment of biological sequences with gap flexibility. In R.A. Baeza-Yates, E. Goles, and P.V. Poblete, editors, *Proceedings of the 2nd Latin American Symposium*, number 911 in Lecture Notes in Computer Science, pages 411–426, Valparaíso, Chile, 1995. Springer-Verlag, Berlin.

[1527] J. Meidanis, M.E. Walter, and Z. Dias. Transposition distance between a permutation and its reverse. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 70–79, Valparaiso, Chile, 1997. Carleton University Press.

[1528] G. Melançon. Factorizing infinite words unsing Maple. *MapleTech*, 4(1):34–42, 1997.

[1529] G. Melançon. Lyndon words and singular factors of sturmian words. *Theor. Comput. Sci.*, 218(1):41–59, 1999.

[1530] B. Melichar. Approximate string matching by finite automata. In V. Hlaváč and R. Šára, editors, *Computer Analysis of Images and Patterns*, number 970 in Lecture Notes in Computer Science, pages 342–349. Springer-Verlag, Berlin, 1995.

[1531] B. Melichar. Space complexity of linear time approximate string matching. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 28–36, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[1532] B. Melichar. String matching with k differences by finite automata. In *Proceedings of the 13th International Conference on Pattern Recognition*, volume II., pages 256–260, Vienna, Austria, 1996. IEEE Computer Society Press.

[1533] B. Melichar. A simple method of complete indexing. In *Proceedings DATASEM*, pages 183–188, Brno, Czech Republic, 1997.

[1534] B. Melichar and J. Holub. 6D classification of pattern matching problems. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 24–32, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[1535] L. Melkemi and R. Tchuente. Un réseau linéaire pour la reconnaissance de mots snans caaré. *RAIRO Inform. Théor. Appl.*, 22:147–161, 1988.

[1536] C. Menico. Faster string searches. *Dr. Dobbs J.*, pages 74–75, 1989.

[1537] T. H. Merrett and H. Shang. Trie methods for representing text. *Lecture Notes in Computer Science*, 730:130–??, 1993.

[1538] T. H. Merrett and H. Shang. Zoom tries: A file structure to support spatial zooming. In *Sixth International Symposium on Spatial Data Handling*, volume 2, pages 792–804, Edinburgh, Scotland, 1994.

[1539] T. H. Merrett, H. Shang, and X. Zhao. Database structures, based on tries, for text, spatial and general data. In *Proceedings of International Symposium on Cooperative Database Systems for Advanced Applications*, pages 316–24, Kyoto, Japan, 1996.

[1540] Y. Métivier. Une condition suffisante de reconnaissabilbité dans un monoïde partiellement commutatif. *RAIRO Inform. Théor.*, 20(2):121–127, 1986.

[1541] Y. Métivier. On recognizable subsets of free partially commutative monoids. *Theor. Comput. Sci.*, 58(1–3):201–208, 1988.

[1542] H. W. Mewes and K. Heumann. Genome analysis: pattern search in biological macromolecules. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 261–285, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1543] B. Meyer. Incremental string matching. *Inf. Process. Lett.*, 21(5):219–227, 1985.

[1544] B. Meyer and C. Baudoin. *Méthodes de programmation.* Eyrolles, Paris, France, 1978.

[1545] M. Middendorf. The shortest common subsequence problem is np-complete. *Theor. Comput. Sci.*, 108(2):365–369, 1993.

[1546] M. Middendorf. More on the complexity of common superstrings and supersequence problem. *Theor. Comput. Sci.*, 125(2):205–228, 1994.

[1547] M. Middendorf. On the approximation of finding various minimal, maximal, and consistent sequences. In D.-Z. Du and X.-S. Zhang, editors, *Proceedings of the 5th International Symposium on Algorithms and Computation*, number 834 in Lecture Notes in Computer Science, pages 306–314, Beijing, P.R. China, 1994. Springer-Verlag, Berlin.

[1548] M. Middendorf. Shortest common superstrings and scheduling with coordinated starting times. *Theor. Comput. Sci.*, 191(1–2):205–214, 1998.

[1549] F. Mignosi. Infinite words with linear subword complexity. *Theor. Comput. Sci.*, 65(2):221–242, 1989.

[1550] F. Mignosi. On the number of factors of Sturmian words. *Theor. Comput. Sci.*, 82(1):71–84, 1991.

[1551] F. Mignosi and A. Restivo. On negative informations in language theory. In C. S. Calude and M. J. Dinneen, editors, *Combinatorics, Computation and Logic. Proceedings of DMTCS'99 and CATS'99*, Lecture Notes in Computer Science, pages 60–72, Auckland, New-Zealand, 1999. Springer-Verlag, Singapore.

[1552] F. Mignosi, A. Restivo, and S. Salemi. Periodicity and the golden ratio. *Theor. Comput. Sci.*, 204(1–2):153–167, 1998.

[1553] R.L. Milidiú, A.A. Pessoa, and E.S. Laber. Practical use of the warm-up algorithm on length-restricted coding. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 80–94, Valparaiso, Chile, 1997. Carleton University Press.

[1554] P. L. Miller, P. M. Nadkarni, and N. M. Carriero. Parallel computation and FASTA: confronting the problem of parallel database search for a fast sequence comparison algorithm. *Comput. Appl. Biosci.*, 7(1):71–78, 1991.

[1555] W. Miller. Building multiple alignments from pairwise alignments. *Comput. Appl. Biosci.*, pages ?–?, 1993.

[1556] W. Miller and E. W. Myers. A file comparison program. *Softw. Pract. Exp.*, 15(11):1025–1040, 1985.

[1557] W. Miller and E. W. Myers. Sequence comparison with concave weighting functions. *Bull. Math. Biol.*, 50(2):97–120, 1988.

[1558] W. Miller and E.W. Myers. A simple row-replacement method. *Softw. Pract. Exp.*, 18(7):597–611, 1988.

[1559] V. Mitrana and A. Salomma G. Păun, G. Rozenberg. Pattern systems. *Theor. Comput. Sci.*, 154(2):183–201, 1996.

[1560] M. Miyazaki, A. Shinohara, and M. Takeda. An improved pattern matching algorithm for strings in terms of straight-line programs. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 1–11, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1561] M. Miyazaki, A. Shinohara, and M. Takeda. An improved pattern matching algorithm for strings in terms of straight-line programs. *J. Discret. Algorithms*, 1(1):187–204, 2000.

[1562] B. Moayer and K. S. Fu. A tree system approach for fingerprint pattern recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8:376–387, 1986.

[1563] S. P. Mohanty. Shortest string containing all permutations. *Discret. Math.*, 31:91–95, 1980.

[1564] M. Mohri. Minimization of sequential transducers. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 151–163, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1565] M. Mohri. Matching patterns of an automaton. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 286–297, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1566] M. Mohri. String-matching with automata. *Nordic J. Comput.*, 4(2):217–231, 1997.

[1567] P. Møller-Nielsen and J. Staunstrup. Experiments with a fast string searching algorithm. *Inf. Process. Lett.*, 18(3):129–135, 1984.

[1568] B. Monien. Deterministic two-way one-head pushdown automata are very powerful. *Inf. Process. Lett.*, 18(3):239–242, 1984.

[1569] D. Moore and W. F. Smyth. Computing the covers of a string in linear time. In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 511–515, Arlington, VA, 1994.

[1570] D. Moore and W. F. Smyth. An optimal algorithm to compute all the covers of a string. *Inf. Process. Lett.*, 50(5):239–246, 1994.

[1571] D. Moore and W. F. Smyth. A correction to "an optimal algorithm to compute all the covers of a string". *Inf. Process. Lett.*, 54(2):101–103, 1995.

[1572] D. Moore, W. F. Smyth, and D. Miller. Counting distinct strings. *Algorithmica*, 23(1):1–13, 1999.

[1573] H. L. Morgan. Spelling correction in systems programs. *Commun. ACM*, 13(2):90–94, 1970.

[1574] T. Mori and S. J. Stolfo. Approximate string matching on the DADO2 parallel computer. Report CUCS-361-88, Computer Science Department, Columbia University, NY, 1988.

[1575] K. Morimoto, H. Iriguchi, and J.-I. Aoe. A method for compressing trie structure. *Softw. Pract. Exp.*, 24(3):265–288, 1994.

[1576] S. A. Morozov, A. D. Cherkay, and I. I. Shagurin. Hardware for character string matching. *J. Comput. Syst. Sci.*, 31(1):71–81, 1993.

[1577] J. H. Morris, E. Schmidt, and P. Wadler. Experience with an applicative string processing language. In *Proceedings of the 7th Annual ACM Symposium on Principles of Programming Languages*, pages 32–46, Las Vegas, NV, 1980.

[1578] J. H. Morris, Jr and V. R. Pratt. A linear pattern-matching algorithm. Report 40, University of California, Berkeley, 1970.

[1579] J. M. Morris. Programming by expressions refinment: the KMP algorithm. In W. H. J. Feijen, A. J. M. van Gasteren, D. Gries, and J. Misra, editors, *Beauty is Our Business*, chapter 37, pages 327–338. Springer-Verlag, Berlin, 1990.

[1580] D. R. Morrison. PATRICIA - practical algorithm to retrieve coded in alphanumeric. *J. Assoc. Comput. Mach.*, 15(4):514–534, 1968.

[1581] R. F. Mott, T. B. L. Kirkwood, and R. N. Curnow. An accurate approximation to the distribution of the length of the longest matching word between two random DNA sequences. *Bull. Math. Biol.*, 52:773–784, 1990.

[1582] L. Mouchard. Canonical decompositions of quasiperiodic words. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 74–84, Perth, WA, Australia, 1998.

[1583] L. Mouchard. *Superpositions dans les séquences.* Thèse de doctorat, Université de Rouen, 1998.

[1584] L. Mouchard. Genome research from the computer scientist point of view. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, page 1, Perth, WA, Australia, 1999.

[1585] L. Mouchard. Normal forms of quasiperiodic strings. *Theor. Comput. Sci.*, 249(2):313–324, 2000.

[1586] A. Mukherjee. Hardware algorithms for determining similarity between two strings. *IEEE Trans. Comput.*, 38(4):600–603, 1989.

[1587] A. Mukhopadhyay. Hardware algorithms for non-numeric computation. In *Proceedings of the 5th International Symposium on Computer Architecture*, pages 8–16, IEEE, New York (published as a special issue of Computer Architecture News (ACM SIGARCH) 6(7)), 1978.

[1588] A. Mukhopadhyay. Hardware algorithms for non-numeric computation. *IEEE Trans. Comput.*, C-28(6):384–394, 1979.

[1589] A. Mukhopadhyay. A fast algorithm for the longest-common-subsequence problem. *Inf. Sci.*, 20:69–82, 1980.

[1590] A. Mukhopadhyay. Hardware algorithms for string processing. In N. B. Guy Rabbat, editor, *Proceedings of the IEEE International Conference on Circuits and Computers*, pages 508–511, 1980.

[1591] Y. Mukouchi. Characterization of pattern languages. In *Proceedings of the 2nd Int. Workshop on Algorithmic Learning Theory*, pages 93–104. IOS Press, 1992.

[1592] M. Murata. Three-way Needleman-Wunsch algorithm. *Methods Enzymol.*, 183:365–375, 1990.

[1593] M. Murata, J. S. Richardson, and J. L. Sussman. Simultaneous comparison of three protein sequences. *Proc. Natl. Acad. Sci. U.S.A.*, 82:3073–3077, 1985.

[1594] R. Muth and U. Manber. Approximate multiple string search. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 75–86, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1595] R. Muthukrishnan. New results and open problems related to non-standard stringology. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 298–317, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1596] S. Muthukrishnan. Detecting false matches in string matching algorithms. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 164–178, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1597] S. Muthukrishnan. Detecting false matches in string matching algorithms. *Algorithmica*, 18(4):512–520, 1997.

[1598] S. Muthukrishnan. Simple optimal parallel multiple pattern matching. *J. Algorithms*, 34(1):1–13, 2000.

[1599] S. Muthukrishnan and R. Hariharan. A constant time optimal parallel algorithm for two dimensional matching. Manuscript, 1992.

[1600] S. Muthukrishnan and K. Palem. Highly efficient dictionary matching in parallel. Report 632, Courant Institut, New York University, NY, 1993.

[1601] S. Muthukrishnan and K. Palem. Highly efficient dictionary matching in parallel. In *Proceedings of the 5th ACM Symposium on Parallel Algorithms and Architectures*, pages 69–78, Velen, Germany, 1993. ACM Press.

[1602] S. Muthukrishnan and K. Palem. Non-standard stringology: Algorithms and complexity. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 770–779, Montréal, Canada, 1994. ACM Press.

[1603] S. Muthukrishnan and H. Ramesh. String matching under a general matching relation. In R. Shyamasundar, editor, *Proceedings of the 12th Conference on Foundations of Software Technology and Theoretical Computer Science*, number 652 in Lecture Notes in Computer Science, pages 356–367, New Dehli, India, 1992. Springer-Verlag, Berlin.

[1604] S. Muthukrishnan and H. Ramesh. String matching under a general matching relation. *Inf. Comput.*, 122(1):140–148, 1995.

[1605] P. Mužátko. Approximate regular expression matching. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 37–41, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[1606] E. W. Myers. Efficient applicatives data types. In *Proceedings of the 11th Annual ACM Symposium on Principles of Programming Languages*, pages 66–75, Salt Lake City, UT, 1984. ACM Press.

[1607] E. W. Myers. Incremental alignment algorithms and their applications. Report TR-90-25, Department of Computer Science, University of Arizona, Tucson, AZ, 1986.

[1608] E. W. Myers. An $O(ND)$ difference algorithm and its variations. *Algorithmica*, 1:251–266, 1986.

[1609] E. W. Myers. A four-russian algorithm for regular expression pattern matching. Report TR-88-34, Department of Computer Science, University of Arizona, Tucson, AZ, 1988.

[1610] E. W. Myers. A sublinear algorithm for approximate keyword searching. Report TR-90-25, Department of Computer Science, University of Arizona, Tucson, AZ, 1990.

[1611] E. W. Myers. An overview of sequence comparison algorithm in molecular biology. Report TR-91-29, Department of Computer Science, University of Arizona, Tucson, AZ, 1991.

[1612] E. W. Myers. Approximate matching of network expression with spacers. Report TR-92-5, Department of Computer Science, University of Arizona, Tucson, AZ, 1992.

[1613] E. W. Myers. A four russians algorithm for regular expression pattern matching. *J. Assoc. Comput. Mach.*, 39(2):430–448, 1992.

[1614] E. W. Myers. Approximately matching context-free languages. Report TR-94-22, Department of Computer Science, University of Arizona, Tucson, AZ, 1994.

[1615] E. W. Myers. A sublinear algorithm for approximate keyword searching. *Algorithmica*, 12(4/5):345–374, 1994.

[1616] E. W. Myers. Approximately matching context-free languages. *Inf. Process. Lett.*, 54(2):85–92, 1995.

[1617] E. W. Myers. Approximately matching context-free languages. In R. Baeza-Yates and U. Manber, editors, *Proceedings of the 2nd South American Workshop on String Processing*, pages 38–52, Valparaíso, Chile, 1995.

[1618] E. W. Myers and X. Huang. An $O(N^2 \log N)$ restriction map comparison and search algorithm. Report TR-90-11, Department of Computer Science, University of Arizona, Tucson, AZ, 1990.

[1619] E. W. Myers and W. Miller. Approximate matching of regular expressions. Report CS-88-20, Department of Computer Science, Penn State University, 1988.

[1620] E. W. Myers and W. Miller. Optimal alignment in linear space. *Comput. Appl. Biosci.*, 4(1):11–17, 1988.

[1621] E. W. Myers and W. Miller. Approximate matching of regular expressions. *Bull. Math. Biol.*, 51:7–37, 1989.

[1622] E. W. Myers and W. Miller. Row replacement algorithms for screen editors. *ACM Trans. Program. Lang. Syst.*, 11(1):33–56, 1989.

[1623] E. W. Myers and W. Miller. Chaining multiple-alignment fragments in sub-quadratic time. In *Proceedings of the 6th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 38–47, San Francisco, CA, 1995.

[1624] E. W. Myers and D. Mount. Computer program for the IBM personal computer that searches for approximate matches of short oligonucleotide sequences in long target DNA sequences. *Nucleic Acids Res.*, 14:501–508, 1986.

[1625] E. W. Myers, P. Oliva, and K. Guimãraes. Reporting exact and approximate regular expression matches. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 91–103, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1626] E. W. Myers, S. Selznick, Z. Zhang, and W. Miller. Progressive multiple alignment with constraints. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 220–225, Santa Fe, NM, 1997. ACM Press.

[1627] G. Myers. An $o(n)$ bit-vector algorithm for approximate string matching. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 1–13, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1628] G. Myers. A fast bit-vector algorithm for approximate string matching based on dynamic programming. *J. Assoc. Comput. Mach.*, 46(3):395–415, 1999.

[1629] J.-F. Myoupo and A. Wabbi. Improved systolic algorithms for substrings statistics. In *Proceedings of the 2nd International Conference on Algorithms and Architextures for Parallel Processing*, pages 9–13, Singapore, 1996.

[1630] M. Nakanishi, M. Hasidume, M. Ito, and A. Hashimoto. A linear-time algorithm for computing characteristic strings. In D.-Z. Du and X.-S. Zhang, editors, *Proceedings of the 5th International Symposium on Algorithms and Computation*, number 834 in Lecture Notes in Computer Science, pages 315–323, Beijing, P.R. China, 1994. Springer-Verlag, Berlin.

[1631] N. Nakatsu, Y. Kambayashi, and S. Yajima. A longest common subsequence algorithm suitable for similar text strings. *Acta Inform.*, 18:171–179, 1982.

[1632] M. Naor. String matching with preprocessing of text and pattern. In J. Leach Albert, B. Monien, and M. Rodriguez Artalejo, editors, *Proceedings of the 18th International Colloquium on Automata, Languages and Programming*, number 510 in Lecture Notes in Computer Science, pages 739–750, Madrid, Spain, 1991. Springer-Verlag, Berlin.

[1633] M. Naor and D. Brutlag. On suboptimal alignments of biological sequences. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 179–196, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1634] G. Navarro. An optimal index for pat arrays. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 214–227, Recife, Brazil, 1996. Carleton University Press.

[1635] G. Navarro. Multiple approximate string matching by counting. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 95–111, Valparaiso, Chile, 1997. Carleton University Press.

[1636] G. Navarro. A partial deterministic automaton for approximate string matching. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 95–111, Valparaiso, Chile, 1997. Carleton University Press.

[1637] G. Navarro. *Approximate Text Searching.* PhD thesis, Dept. of Computer Science, University of Chile, 1998.

[1638] G. Navarro. Improved approximate pattern matching on hypertex. In C. L. Lucchesi and A. V. Moura, editors, *Proceedings of the 3rd Latin American Symposium*, number 1380 in Lecture Notes in Computer Science, pages 352–357, Campinas, Brazil, 1998. Springer-Verlag, Berlin.

[1639] G. Navarro. Very fast and simple approximate string matching. *Inf. Process. Lett.*, 72(1–2):65–70, 1999.

[1640] G. Navarro. Improved approximate pattern matching on hypertext. *Theor. Comput. Sci.*, 237(1–2):455–463, 2000.

[1641] G. Navarro and R. Baeza-Yates. Improving an algorithm for approximate pattern matching. Technical Report TR/DCC-98-5, Dept. of Computer Science, University of Chile, 1998.

[1642] G. Navarro and R. Baeza-Yates. A practical q-gram index for text retrieval allowing errors. *CLEI Electonic Journal*, 1(2), 1998.

[1643] G. Navarro and R. Baeza-Yates. Very fast and simple approximate string matching. Technical Report TR/DCC-98-6, Dept. of Computer Science, University of Chile, 1998.

[1644] G. Navarro and R. Baeza-Yates. Fast multi-dimensional approximate pattern matching. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 243–257, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1645] G. Navarro and R. Baeza-Yates. A new indexing method for approximate string matching. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 163–185, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1646] G. Navarro and R. Baeza-Yates. A hybrid indexing method for approximate string matching. *J. Discret. Algorithms*, 1(1):205–239, 2000.

[1647] G. Navarro, J. P. Kitajima, B. A. Ribeiro-Neto, and N. Ziviani. Distributed generation of suffix arrays. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 102–115, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1648] G. Navarro and M. Raffinot. A bit-parallel approach to suffix automata: Fast extended string matching. Technical Report TR/DC–98–1, Department of Computer Science, University of Chile, 1998.

[1649] G. Navarro and M. Raffinot. A bit-parallel approach to suffix automata: Fast extended string matching. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 14–33, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1650] G. Navarro and M. Raffinot. Fast and flexible string matching by combining bit-parallelism and suffix automata. Technical Report TR/DC–98–4, Department of Computer Science, University of Chile, 1998.

[1651] G. Navarro and M. Raffinot. Fast regular expression search. In J.S. Vitter and C.D. Zaroliagis, editors, *Proceedings of the 3rd Workshop on Algorithm Engineering*, number 1668 in Lecture Notes in Computer Science, pages 199–213, London, UK, 1999. Springer-Verlag, Berlin.

[1652] G. Navarro and M. Raffinot. A general practical approach to pattern matching over Ziv-Lempel compressed text. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 14–36, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1653] G. Navarro, E. Sutinen, J. Tanninen, and J. Tarhio. Indexing text with approximate q-grams. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 350–363, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1654] G. Navarro and J. Tarhio. Boyer-Moore string matching over ziv-lempel compressed text. In R. Giancarlo and D. Sankoff, editors, *Proceedings of* *the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 166–180, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1655] Gonzalo Navarro. Combining text compression and string matching: The miracle of self-indexing. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, page 1, Czech Technical University in Prague, Czech Republic, 2009.

[1656] N. Nedjah. Minimal deterministic left-to-right pattern-matching automata. *SIGPLAN Not.*, 33(1):40–47, 1998.

[1657] N. Nedjah, C. D. Walter, and S. E. Eldridge. Efficient automata-driven pattern-matching for equational programs. *Softw. Pract. Exp.*, 29(9):793–813, 1999.

[1658] S. B. Needleman and C. D. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, 48:443–453, 1970.

[1659] M. Neff and B. K. Boguraev. Dictionaries, dictionary grammars and dictionary entry parsing. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pages 91–101, 1989.

[1660] M. R. Nelson, S. L. Kardia, and C. F. Sing. The combinatorial partitioning method. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 293–304, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1661] J. Néraud. Computing the rank of a finite set of word is np-complete. Rapport LITP 88.28, Universités Paris 6 et 7, France, 1988.

[1662] J. Néraud. On the deficit of a finite set of words. Rapport LITP 88.56, Universités Paris 6 et 7, France, 1988.

[1663] J. Néraud. Elementariness of a finite set of words is co-np-complete. *RAIRO Inform. Théor. Appl.*, 24(5):459–470, 1990.

[1664] J. Néraud. On the deficit of a finite set of words. *Semigroup Forum*, 41:1–21, 1990.

[1665] J. Néraud. On the rank of the subsets of a free monoid, deciding whether a finite set of words has rank at most two. In L. Budach, editor, *Proceedings Fundamental of Computation Theory*, number 529 in Lecture Notes in Computer Science, pages 360–369, Gosen near Berlin, Germany, 1991.

[1666] J. Néraud. On the rank of the subsets of a free monoid. *Theor. Comput. Sci.*, 99(2):231–241, 1992.

[1667] J. Néraud. Algorithms for detecting morphic images of a word. *Inf. Comput.*, 120(1):126–148, 1993.

[1668] J. Néraud. Deciding whether a finite set of words has rank at most two. *Theor. Comput. Sci.*, 112(2):311–337, 1993.

[1669] J. Néraud. Detecting morphic images of a word: improving the general algorithm. Rapport LITP 93.54, Universités Paris 6 et 7, 1993.

[1670] J. Néraud. New algorithms for detecting morphic images of a word. In *Proceedings of the 18th Symposium on Mathematical Foundations of Computer Science*, number 711 in Lecture Notes in Computer Science, pages 588–597, Gdansk, Poland, 1993. Springer-Verlag, Berlin.

[1671] J. Néraud. A special class of equations in words: efficient computation of the solutions. *Bull. Belg. Math. Soc.*, 1(2):253–283, 1994.

[1672] J. Néraud. Detecting morphic images of a word: On the rank of a pattern. *Acta Inform.*, 32:477–489, 1995.

[1673] J. Néraud. Pour une description structurelle des sous-ensembles du monoïde libre : Images homomorphes, concepts combinatoires et résultats de complexité. Thèse d'habilitation, Rapport LIR95.02 Informatique Fondamentale, Université de Rouen, France, 1995.

[1674] J. Néraud. Detecting the morphic images of a word: improving the general algorithm. *RAIRO Inform. Théor. Appl.*, 31(1):1–14, 1997.

[1675] J. Néraud and M. Crochemore. A string matching interpretation of the equation $x^m y^n = z^p$. *Theor. Comput. Sci.*, 92(1):145–164, 1992.

[1676] J. Nesbit. The accuracy of approximate string matching algorithms. *Journal of Computer-Based Instruction*, 13(3):80–83, 1986.

[1677] C. Nevill-Manning and Ian H. Witten. Protein is incompressible. In *Proceedings Data Compression Conference*, pages 257–266, 1999.

[1678] P. Nicodeme. *Alignement avec des familles de séquences protéiques*. PhD thesis, Université Paris VII, 1997.

[1679] P. Nicodeme. SSMAL: similarity searching with alignment graphs. *Bioinformatics*, 14(6):508–515, 1998.

[1680] P. Nicodeme and M. Régnier. Towards 2D pattern matching complexity. Manuscript, 1992.

[1681] P. Nicodeme, B. Salvy, and P. Flajolet. Motif statistics. In *Proceedings Annual European Symposium*, volume 1643 of *Lecture Notes in Computer Science*, pages 194–211. Springer-Verlag, Berlin, 1999.

[1682] P. T. Nielsen. A note on bifix-free sequences. *IEEE Trans. Inf. Theory*, 19:704–707, 1973.

[1683] P. T. Nielsen. On the expected duration of a search for a fixed pattern in random data. *IEEE Trans. Inf. Theory*, 19:702–704, 1973.

[1684] K. Nilsen. Garbage collection of strings and linked data structures in real time. *Softw. Pract. Exp.*, 18(7):613–640, 1988.

[1685] S. Nilsson and M. Tikkanen. Implementing a dynamic compressed trie. In K. Mehlhorn, editor, *Proceedings Workshop on Algorithm Engineering*, pages 1–3, Saarbrucken, Germany, 1998.

[1686] M. Nivat. Infinite words, infinite trees, infinite computations. *Math. Cent. Tracts*, 109:1–52, 1979.

[1687] R. Nussinov. Efficient algorithms for searching for exact repetition of nucleotide sequences. *J. Mol. Evol.*, 19(3–4):283–285, 1983.

[1688] R. Nussinov, G. Pieczenik, J. R. Griggs, and D. J. Kleitman. Algorithm for loop matchings. *SIAM J. Appl. Math.*, 35:68–82, 1978.

[1689] M. Nykänen and E. Ukkonen. Finding lowest common ancestors in arbitrarily directed trees. *Inf. Process. Lett.*, 50(6):307–310, 1994.

[1690] W. Obst. Delta technique and string-to-string correction. In G. Goos and J. Hartmanis, editors, *Proceedings of the 1st European Software Engineering Conference*, number 289 in Lecture Notes in Computer Science, pages 64–68. Springer-Verlag, Berlin, 1987.

[1691] E. Ohlebusch. A filter method for the weighted local similarity search problem. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 191–205, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1692] T. Okuda, E. Tanaka, and T. Kasai. A method of correction of garbled words based on the levenshtein metric. *IEEE Trans. Comput.*, 25:172–177, 1976.

[1693] B. J. Oommen. Constrained string editing. *Inf. Sci.*, 40:267–284, 1987.

[1694] B. J. Oommen. Recognition of noisy subsequences using constrained edit distances. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9:676–685, 1987.

[1695] B. J. Oommen. String editing with substitution, insertion, deletion, squahing and expansion operations. Manuscript, 1992.

[1696] B. J. Oommen and E. T. Floyd. An improved algorithm for the recognition of noisy subsequences. In *Proceedings of the IASTED International Symposium on Artificial Intelligence Applications and Neural Networks*, pages 145–147, Zurich, Switzerland, 1991.

[1697] B. J. Oommen and K. Zhang. The normalized string editing problem revsited. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(6):669–672, 1996.

[1698] J. A. Orenstein. Multidimensional tries used for associative searching. *Inf. Process. Lett.*, 14(4):150–157, 1982.

[1699] F. Otto. Solvability of word equations modulo finite special and confluent string rewriting systems is undecidable in general. *Inf. Process. Lett.*, 53(5):237–242, 1995.

[1700] J. Owens, D. Chatterjee, R. Nussinov, A. K. Konopka, and J. V. Maizel, Jr. A fixed-point alignment technique for detection of recurrent and common sequence motifs associated with biological features. *Comput. Appl. Biosci.*, 4:73–77, 1988.

[1701] O. Owolabi. Efficient pattern searching over large dictionaries. *Inf. Process. Lett.*, 47(1):17–21, 1993.

[1702] O. Owolabi and D. R. McGregor. Fast approximate string matching. *Softw. Pract. Exp.*, 18(4):387–393, 1988.

[1703] R. Paige and S. Koenig. Finite differencing of computable expressions. *ACM Trans. Program. Lang. Syst.*, 4(3):402–454, 1982.

[1704] R. Pajarola and P. Widmayer. Pattern matching in compressed raster images. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 228–242, Recife, Brazil, 1996. Carleton University Press.

[1705] Geneviève Paquin. Infinite smooth Lyndon words. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 126–139, Czech Technical University in Prague, Czech Republic, 2008.

[1706] L. Parida. Some results on flexible-pattern discovery. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 33–45, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1707] L. Parida, A. Floratos, and I. Rigoutsos. An approximation algorithm for alignment of multiple sequences using motif discovery. *J. Comb. Optim.*, 3(2/3):247–275, 1999.

[1708] L. Parida, Y. Gao, D. Platt, A. Floratos, and I. Rigoutsos. Pattern discovery on character sets and real-valued data: Linear bound on irredundant motifs and an efficient polynomial time algorithm. In *Proceedings ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 297–308, San Francisco, CA, 2000.

[1709] L. Parida, I. Rigoutsos, A. Floratos, D. Platt, and Y. Gao. Pattern discovery on character sets and real-valued data: Linear bound on irredundant motifs and an efficient polynomial time algorithm. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 297–308, San Francisco, CA, 2000.

[1710] J. Park and K. George. Parallel string matching algorithms based on dataflow. In *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*, pages ?–?, Hawaii, 1999. IEEE Computer Society Press.

[1711] K. Park. Analysis of two-dimensional approximate pattern matching algorithms. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 335–347, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1712] K. Park. Analysis of two-dimensional approximate pattern matching algorithms. *Theor. Comput. Sci.*, 201(1–2):263–273, 1998.

[1713] K. Park. Recent developments in index data structures. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, page 3, Perth, WA, Australia, 1999.

[1714] K. Park and D. K. Kim. String matching in hypertext. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 318–329, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1715] K.-H. Park, J.-I. Aoe, K. Morimoto, and D. J. Evans. An algorithm for dynamic processing of dawg's. *Int. J. Comput. Math.*, 54(3–4):155–173, 1994.

[1716] S.-H. Park, W. W. Chu, J.-H. Yoon, and C.-C. Hsu. Fast retrieval of similar sub-sequences under time warping. In *16th IEEE Internationall Conference on Data Engineering (ICDE)*, pages ??–??, San Diego, CA, 2000.

[1717] H. A. Partsch and F. A. Stomp. A fast pattern matching algorithm derived by transformational and assertional reasoning. *Form. Asp. Comput.*, 2:109–122, 1991.

[1718] H. A. Partsch and N. Völker. Another case study on reusability of transformational developments - pattern matching according to Knuth, Morris, and Pratt. In M. Broy and W. Wirsing, editors, *Methods of Programming: Selected Papers on the CIP-Project*, number 544 in Lecture Notes in Computer Science, pages 35–48. Springer-Verlag, Berlin, 1991.

[1719] S. Pascarella and P. Argos. Analysis of insertion/deletion in proteins. *J. Mol. Biol.*, 224:461–471, 1990.

[1720] M. Paterson and V. Dančík. Longuest common subsequences. In I. Prívara, B. Rovan, and P. Ružička, editors, *Proceedings of the 19th Symposium on Mathematical Foundations of Computer Science*, number 841 in Lecture Notes in Computer Science, pages 127–142, Košice, Slovakia, 1994. Springer-Verlag, Berlin.

[1721] P. K. Pearson. Fast hashing of variable-length text strings. *Commun. ACM*, 33(6):677–680, 1990.

[1722] W. R. Pearson. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.*, 183:63–98, 1990.

[1723] W. R. Pearson. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics*, 11:635–650, 1991.

[1724] W. R. Pearson and D. J. Lipman. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.*, 85:2444–2448, 1988.

[1725] W. R. Pearson and W. Miller. Dynamic programming algorithms for biological sequence comparison. *Methods Enzymol.*, 210:575–601, 1992.

[1726] J. P. Pecuchet. *Equations avec constantes et algorithme de Makanin.* PhD thesis, Université de Rouen, France, 1981. Thès de 3e Cycle.

[1727] C. Pedersen, R. Lyngso, and J. Hein. Comparison of coding DNA. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 153–173, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1728] I. Pe'er, R. Shamir, and R. Sharan. Incomplete directed perfect phylogeny. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 143–153, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1729] H. Peltola, H. Soderland, J. Tarhio, and E. Ukkonen. Algorithms for some string matching problems arising in molecular genetics. In *Proceedings of the 9th IFIP World Computer Congress*, pages 53–64, Paris, France, 1983.

[1730] F. C. N. Pereira. Machine learning for efficient natural-language processing. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, page 11, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1731] C. H. Perleberg. Three longest substring algorithms. In *Proceedings of the 1st South American Workshop on String Processing*, pages ?–?, Belo Horizonte, Brazil, 1993.

[1732] C. H. Perleberg. Single character seraching methods and the shift-or pattern-matching algorithm. *Inf. Process. Lett.*, 50(5):269–275, 1994.

[1733] D. Perrin. Words over a partially commutative alphabet. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 329–340. Springer-Verlag, Berlin, 1985.

[1734] D. Perrin. Automates et algorithmes sur les mots. Rapport LITP 88.51, Universités Paris 6 et 7, France, 1988.

[1735] D. Perrin. Automates et algorithmes sur les mots. *Ann. Télécommun.*, 44:20–33, 1989.

[1736] D. Perrin. Partial commutations. In G. Ausiello, M. Dezani-Ciancaglini, and S. Ronchi Della Rocca, editors, *Proceedings of the 16 th International Colloquium on Automata, Languages and Programming*, number 372 in Lecture Notes in Computer Science, pages 637–651, Stresa, Italy, 1989. Springer-Verlag, Berlin.

[1737] D. Perrin. Finite automata. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science, vol b, Formal Models and Semantics*, pages 1–57. Elsevier, Amsterdam, The Netherlands, 1990.

[1738] D. Perrin and J.-E. Pin. Mots infinis. Rapport LITP 93.40, Universités Paris 6 et 7, 1993.

[1739] J. L. Petersen. The ambiguity of primitive words. In P. Enjalbert, E. W. Mayr, and K. W. Wagner, editors, *Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science*, number 775 in Lecture Notes in Computer Science, pages 679–690, Caen, France, 1994. Springer-Verlag, Berlin.

[1740] J. L. Peterson. Computer programs for detecting and correcting spelling errors. *Commun. ACM*, 23(12):676–687, 1980.

[1741] P. A. Peterson and M. C. Loui. The general maximum matching algorithm of Micali and Vazirani. *Algorithmica*, 3:511–533, 1988.

[1742] A. Pettorossi, M. Proietti, and S. Renault. How to extend partial deduction to derive the KMP string-matching algorithm from a naive specification (poster abstract). In *Proceedings of the 13th Joint International Conference and Symposium on Logic Programming*, page 539, Bonn, Germany, 1996.

[1743] P. A. Pevzner. Multiple alignment, communication cost, and graph matching. *SIAM J. Appl. Math.*, 52(6):1763–1779, 1992.

[1744] P. A. Pevzner. Multiple alignment with guaranteed errors bounds and communication costs. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 205–213, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1745] P. A. Pevzner. DNA physical mapping and alternating eulerian cycles in colored graphs. *Algorithmica*, 13(1/2):77–105, 1995.

[1746] P. A. Pevzner. *Computational Molecular Biology: An Algorithmic Approach*. MIT Press, 2000.

[1747] P. A. Pevzner and R. Lipshutz. Towards DNA sequencing by hybridization. Manuscript, 1993.

[1748] P. A. Pevzner and R. Lipshutz. Towards DNA sequencing chips. In I. Prívara, B. Rovan, and P. Ružička, editors, *Proceedings of the 19th Symposium on Mathematical Foundations of Computer Science*, number 841 in Lecture Notes in Computer Science, pages 143–158, Košice, Slovakia, 1994. Springer-Verlag, Berlin.

[1749] P. A. Pevzner and S.-H. Sze. Combinatorial approaches to finding subtle signals in DNA sequences. In *Proceedings International Conference on Intelligent Systems for Molecular Biology*, pages 269–278. AAAI Press, 2000.

[1750] P. A. Pevzner and M. Vingron. Multiple sequence comparison and consistency on multipartite graphs. *Adv. Appl. Math.*, 16:1–22, 1995.

[1751] P. A. Pevzner and M. S. Waterman. Matrix longest common subsequence problem, duality and hilbert bases. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 79–89, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1752] P. A. Pevzner and M. S. Waterman. A fast filtration algorithm for the substring matching problem. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 197–214, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1753] P. A. Pevzner and M. S. Waterman. Generalized sequence alignment and duality. *Adv. Appl. Math.*, 14:139–171, 1993.

[1754] P. A. Pevzner and M. S. Waterman. Generalized sequence alignment and duality. *Adv. Appl. Math.*, 14(2):139–171, 1993.

[1755] P. A. Pevzner and M. S. Waterman. Multiple filtration and approximate pattern matching. *Algorithmica*, 13(1/2):135–154, 1995.

[1756] C. A. Phillips and T. Warnow. The asymmetric median tree - a new model for building consensus trees. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 234–252, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1757] Marcin Piatkowski and Wojciech Rytter. Asymptotic behaviour of the maximal number of squares in standard sturmian words. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 237–248, Czech Technical University in Prague, Czech Republic, 2009.

[1758] R. Y. Pinter. Efficient string matching with don't care pattern. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, pages 11–29. Springer-Verlag, Berlin, 1985.

[1759] G. Pirillo. Infinite words and biprefix codes. *Inf. Process. Lett.*, 50(6):293–295, 1994.

[1760] K. Pirklbauer. A study of pattern-matching algorithms. *Struct. Program.*, pages 89–98, 1992.

[1761] W. Plandowski. Testing equivalence of morphisms on context-free languages. In J. van Leeuwen, editor, *Proceedings of the 2nd Annual European Symposium*, number 855 in Lecture Notes in Computer Science, pages 460–470, Utrecht, The Netherlands, 1994.

[1762] W. Plandowski. Satisfiability of word equations with constants is in NEXPTIME. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 473–482, 1999.

[1763] D. R. Powell, L. Allison, and T. I. Dix. A versatile divide and conquer technique for optimal string alignment. *Inf. Process. Lett.*, 70(3):127–139, 1999.

[1764] D. R. Powell, D. L. Dowe, L. Allison, and T. I. Dix. Discovering simple DNA sequences by compression. In *Pacific Symposium on Biocomputing '98*, pages 597–608. World Scientic, 1998.

[1765] P. A. Powell. Sequence similarity algorithms. Technical report, Univ. Minnesota, dept. CS, 1989.

[1766] P. A. Powell. Sequence similarity algorithms and a linear systolic VLSI implementation. Technical report, Univ. Minnesota, dept. CS, 1989.

[1767] V. Pratt. Improvements and applications for the weiner repetition finder. Manuscript, 1975.

[1768] P. Pritchard. A problem involving subsequences. In W. H. J. Feijen, A. J. M. van Gasteren, D. Gries, and J. Misra, editors, *Beauty is Our Business*, chapter 42, pages 359–364. Springer-Verlag, Berlin, 1990.

[1769] J. Prosise. High-performance string searching: How a smart algorithm can make a big difference. *PC Magazine*, 15(15), 1996.

[1770] B. Prum, F. Rodolphe, and E. de Turckheim. Finding words with unexpected frequencies in deoxyribonucleic acid sequences. *J. Roy. Statist. Soc. Ser. B.*, 57:205–220, 1995.

[1771] N. Prunella, S. Liuni, M. Attimonelli, and G. Pesole. Fastpat: a fast and efficient algorithm for string searching in dna sequences. *Comput. Appl. Biosci.*, 9(5):541–545, 1993.

[1772] Simon J. Puglisi, William F. Smyth, and Munina Yusufu. Fast optimal algorithms for computing all the repeats in a string. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 161–169, Czech Technical University in Prague, Czech Republic, 2008.

[1773] P. W. Purdom and C. A. Brown. Fast many-to-one matching algorithms. In J.-P. Jouannaud, editor, *Proceedings of the 1st RTA*, number 202 in Lecture Notes in Computer Science, pages 407–416, Dijon, France, 1985. Springer-Verlag, Berlin.

[1774] R. W. Quong. Fast average-case pattern matching by multiplexing sparse tables. *Theor. Comput. Sci.*, 92(1):165–179, 1992.

[1775] Y. Rabani and Z. Galil. On the space complexity of some algorithms for sequence comparison. *Theor. Comput. Sci.*, 95(2):231–244, 1992.

[1776] M. O. Rabin. Discovering repetitions in strings. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 279–288. Springer-Verlag, Berlin, 1985.

[1777] M. O. Rabin. Optimal parallel pattern matching through randomization. In R. Capocelli, A. De Santis, and U. Vaccaro, editors, *Proceedings Sequences: Methods in Communication, Security and Computer Science*, pages ?–? Springer-Verlag, Berlin, 1991.

[1778] M. O. Rabin and D. Scott. Finite automata and their decision problems. *IBM J. Res. Dev.*, 3:114–125, 1959.

[1779] M. O. Rabin and D. Scott. Finite automata and their decision problems. In E. F. Moore, editor, *Sequential Machines: Selected Papers*, pages 63–91. Addison-Wesley, Reading, MA, 1964.

[1780] M. Raffinot. Asymptotic estimation of the average number of terminal states in dawgs. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 140–148, Valparaiso, Chile, 1997. Carleton University Press.

[1781] M. Raffinot. On the multi backward dawg matching algorithm (MultiBDM). In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 149–165, Valparaiso, Chile, 1997. Carleton University Press.

[1782] M. Raffinot. *Structures pour le localisation de motifs.* Thèse de doctorat, Université de Marne-la-Vallée, 1999.

[1783] S. Rahmann and E. Rivals. Exact and efficient computation of the expected number of missing and common words in random texts. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 375–387, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1784] K. J. Räihä and E. Ukkonen. The shortest common supersequence problem over binary alphabet is np-complete. *Theor. Comput. Sci.*, 16(2):187–198, 1981.

[1785] T. Raita. Tuning the Boyer-Moore-Horspool string searching algorithm. *Softw. Pract. Exp.*, 22(10):879–884, 1992.

[1786] T. Raita. On guards and symbol dependencies in substring search. *Softw. Pract. Exp.*, 29(11):931–941, 1999.

[1787] R. Raman. Algorithms beyond compare: exploiting word-level parallelism. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 13–15, Perth, WA, Australia, 1998.

[1788] R. Ramesh and I. V. Ramakrishnan. Parallel tree pattern matching. *J. Symbol. Comput.*, 9:485–501, 1990.

[1789] R. Ramesh and I. V. Ramakrishnan. Nonlinear pattern matching in trees. *J. Assoc. Comput. Mach.*, 39(2):295–316, 1992.

[1790] R. Ramesh, R. M. Verma, T. Krishnaprasad, and I. V. Ramakrishnan. Term matching on parallel computers. In T. Ottmann, editor, *Proceedings of the 14th International Colloquium on Automata, Languages and Programming*, number 267 in Lecture Notes in Computer Science, pages 336–346, Karlsruhe, Germany, 1987. Springer-Verlag, Berlin.

[1791] R. Ramesh, R. M. Verma, T. Krishnaprasad, and I. V. Ramakrishnan. Term matching on parallel computers. *J. Log. Program.*, 6:213–228, 1989.

[1792] S. Ranka and T. Heywood. Two-dimensional pattern matching with k mismatches. *Pattern Recognit.*, 24(1):31–40, 1991.

[1793] S. Ranka and S. Sahni. String editing on an SIMD hypercube multicomputer. *J. Parallel Distr. Comput.*, 9(4):411–418, 1990.

[1794] R. Ravi and J. D. Kececioglu. Approximation algorithms for multiple sequence alignment under a fixed evolutionary tree. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 330–339, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1795] D. R. Raymond and H. J. Fawcett. Playing detective with full text searching software. Report OED-90-03, University of Waterloo, 1990.

[1796] Sébastien Rebecchi and Jean-Michel Jolion. On the uniform distribution of strings. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 116–125, Czech Technical University in Prague, Czech Republic, 2008.

[1797] M. Régnier. Knuth-Morris-Pratt algorithm: an analysis. Rapport 966, INRIA, Rocquencourt, France, 1989.

[1798] M. Régnier. Knuth-Morris-Pratt algorithm: an analysis. In *Proceedings of the 14th Symposium on Mathematical Foundations of Computer Science*, Lecture Notes in Computer Science, pages 431–444, Porabka-Kozubnik, Poland, 1989. Springer-Verlag, Berlin.

[1799] M. Régnier. Performance of string searching algorithm under various probabilistic models. Rapport 1565, INRIA, Rocquencourt, France, 1991.

[1800] M. Régnier. Enumeration of bordered words, le langage de la vache-quirit. *RAIRO Inform. Théor.*, 26:303–317, 1992.

[1801] M. Régnier. Fast two dimensional pattern matching. Chapter 28 in RR-INRIA-1779: Algorithms Seminar 1991-1992, 1992.

[1802] M. Régnier. A language approach to string searching evaluation. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 15–26, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1803] M. Régnier. A unified approach to word statistics. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 207–213, New York, NY, 1998. ACM Press.

[1804] M. Régnier. A unified approach to word occurrence probabilities. *Discret. Appl. Math.*, 104:259–280, 2000.

[1805] M. Régnier and L. Mouchard. Periods and quasiperiods characterization. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 388–396, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1806] M. Régnier and L. Rostami. A unifying look at d-dimensional periodicities and space coverings. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 215–227, Padova, Italy, 1993. Springer-Verlag, Berlin.

[1807] M. Régnier and L. Rostami. A simple (but optimal) 2d-witness algorithm. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 243–256, Recife, Brazil, 1996. Carleton University Press.

[1808] M. Régnier and W. Szpankowski. On the approximate pattern occurrence in a text. In *Proceedings Compression and Complexity of SEQUENCES'97*. IEEE Press, 1997.

[1809] M. Régnier and W. Szpankowski. On pattern frequency occurrences in a Markovian sequence. *Algorithmica*, 22:631–649, 1998.

[1810] T. A. Reichert, D. N. Cohen, and A. K. C. Wong. An application of information theory to genetic mutations and the matching of polypeptide sequences. *J. Theor. Biol.*, 42:245–261, 1973.

[1811] G. Reinert and S. Schbath. Compound Poisson and Poisson process approximations for occurrences of multiple words in Markov chains. *J. Comput. Bio.*, 5:223–254, 1998.

[1812] G. Reinert, S. Schbath, and M. S. Waterman. Probabilistic and statistical properties of words: An overview. *J. Comput. Bio.*, 7:1–46, 2000.

[1813] K. Reinert, H.-P. Lenhof, P. Mutzel, K. Mehlhorn, and J. D. Kececioglu. A branch-and-cut algorithm for multiple sequence alignment. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 241–250, Santa Fe, NM, 1997. ACM Press.

[1814] E. M. Reingold and K. J. Urban. K-M-P string matching without pain. Manuscript, 1996.

[1815] E. M. Reingold, K. J. Urban, and D. Gries. K-M-P string matching revisited. *Inf. Process. Lett.*, 64(5), 1997.

[1816] A. Restivo. A note on multiset decipherable codes. *IEEE Trans. Inf. Theory*, 35(3):662–663, 1989.

[1817] A. Restivo. Permutation properties and the fibonacci semigroup. *Semigroup Forum*, 38:337–345, 1989.

[1818] A. Restivo. Coding sequences with constraints. In R. Capocelli, editor, *Sequences: Combinatorics, Compression, Security and Transmission*, pages 530–539. Springer-Verlag, Berlin, 1990.

[1819] A. Restivo and S. Salemi. Some decision results on nonrepetitive words. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages 289–295. Springer-Verlag, Berlin, 1985.

[1820] A. Restivo, S. Salemi, and T. Sportelli. Completing codes. *RAIRO Inform. Théor.*, 23(2):135–147, 1989.

[1821] D. Revuz. Minimization of acyclic deterministic automata in linear time. *Theor. Comput. Sci.*, 92(1):181–189, 1992.

[1822] D. Revuz and M. Zipstein. DZ: A text compression algorithm for natural languages. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 644 in Lecture Notes in Computer Science, pages 193–204, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1823] S. W. Reyner. An analysis of a good algorithm for the subtree problem. *SIAM J. Comput.*, 6(4):730–732, 1977.

[1824] S. V. Rice, H. Bunke, and T. A. Nartker. Classes of cost functions for string edit distance. *Algorithmica*, 18(2):271–280, 1997.

[1825] T. Richter. A new algorithm for the ordered tree inclusion problem. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 150–166, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1826] C. Rick. A new flexible algorithm for the longest common subsequence problem. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 340–351, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1827] C. Rick. Simple and fast linear space computation of longest common subsequences. *Inf. Process. Lett.*, 75(6):275–281, 2000.

[1828] I. Rigoutsos and A. Califano. Searching in parallel for similar strings. *IEEE Comput. Sci. & Eng.*, 1(2):60–75, 1994.

[1829] I. Rigoutsos and A. Floratos. Combinatorial pattern discovery in biological sequences: The Teiresias algorithm. *Bioinformatics*, 14(1):55–67, 1998.

[1830] I. Rigoutsos and A. Floratos. Motif discovery in biological sequences without alignment or enumeration. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 221–227, New York, NY, 1998. ACM Press.

[1831] E. M. Riseman and R. W. Elrich. Contextual word recognition using binary digrams. *IEEE Trans. Comput.*, C-20(4):397–403, 1971.

[1832] E. M. Riseman and A. R. Hanson. A contextual postprocessing system for error correction using binary n-grams. *IEEE Trans. Comput.*, 5:480–493, 1974.

[1833] J.-L. Risler, M.-O. Delorme, H. Delacroix, and A. Henaut. Amino acid substitutions in structurally related proteins: A pattern recognition approach. Determination of a new and efficient scoring matrix. *J. Mol. Biol.*, 204:1019–1029, 1998.

[1834] J. Rissanen. Complexity of strings in the class of Markov sources. *IEEE Trans. Inf. Theory*, 32(4):526–532, 1986.

[1835] E. S. Ristad and P. N. Yianilos. String similarity. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(5):522–532, 1998.

[1836] S. Ristov and E. Laporte. Ziv Lempel compression of huge natural language data tries using suffix arrays. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 196–211, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1837] S. Ristov and E. Laporte. Ziv Lempel compression of huge natural language data tries using suffix arrays. *J. Discret. Algorithms*, 1(1):241–256, 2000.

[1838] K. Risvik. Approximate word sequence mathing over spase suffix trees. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 65–79, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[1839] E. Rivals, O. Delgrange, J. P. Delahaye, M. Dauchet, M. O. Delorme, A. Henaut, and E. Ollivier. Detection of significant patterns by compression algorithms: The case of approximate tandem repeats in DNA sequences. *Comput. Appl. Biosci.*, 13(2):131–136, 1997.

[1840] R. L. Rivest. Partial-match retrieval algorithms. *SIAM J. Comput.*, 5(1):19–50, 1976.

[1841] R. L. Rivest. On the worst case behavior of string searching algorithms. *SIAM J. Comput.*, 6(4):669–674, 1977.

[1842] Y. Robert and M. Tchuente. Calcul en temps linéaire d'une plus longue sous-suite commune à deux chaîne sur une architecture systolique. *Comptes Rendus Acad. Sci. Paris Sér. I Math.*, 7:269–271, 1984.

[1843] Y. Robert and M. Tchuente. A systolic array for the longest common subsequence problem. *Inf. Process. Lett.*, 21(4):191–198, 1985.

[1844] D. C. Roberts. A computer system for text retrieval: design concept development. Report RD-77-10011, Office of Research and Development, Central Intelligence Agency, Washington, DC, 1977.

[1845] D. C. Roberts. A specialized computer architecture for text retrieval. In *Proceedings of the 4th Workshop on Computer Architecture*, pages 51–59, 1982.

[1846] S. Robin and J.-J. Daudin. Exact distribution of word occurrences in a random sequence of letters. *J. Appl. Probab*, 36:179–193, 1999.

[1847] E. Roche. Smaller representations for finite-state transducers and finite-state automata. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 352–365, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1848] E. Roche. Compact factorization of finite-state transducers and finite-state automata. *Nordic J. Comput.*, 4(2):187–216, 1997.

[1849] E. Rocke. Using suffix trees for gapped motif discovery. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 335–349, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1850] E. Rocke and M. Tompa. An algorithm for finding novel gapped motifs in DNA sequences. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 228–233, New York, NY, 1998. ACM Press.

[1851] M. Rodeh. A fast test for unique decipherability based on suffix trees. *IEEE Trans. Inf. Theory*, 28:648–651, 1982.

[1852] M. Rodeh, V. Pratt, and S. Even. Linear algorithm for data compression via string matching. *J. Assoc. Comput. Mach.*, 28(1):16–24, 1991.

[1853] K. Rohde. A multiple alignment routine for proteins: use, limits and experiences. In *Abstract of: Application of data bases and microcomputers in molecular biology, Jena*, page 47. ?, 1988.

[1854] A. Rosenfeld and A. Kak. *Digital picture processing*. Academic Press, New York, NY, 1982.

[1855] R. Ross and K. Winklmann. Repetitive strings are not context free. Report CS-81-070, Washington State University, Pullman, 1981.

[1856] R. Ross and K. Winklmann. Repetitive strings are not context free. *RAIRO Inform. Théor.*, 16(3):191–199, 1982.

[1857] P. Roth. Every binary pattern of length six is avoidable on the two-letter alphabet. *Acta Inform.*, 29:95–106, 1992.

[1858] M. A. Roytberg. A search for common pattern in many sequences. *Comput. Appl. Biosci.*, 8:57–64, 1992.

[1859] A. Rubinov and V. Timkovsky. String non-inclusion optimization problems. *SIAM J. Discret. Math.*, 11(3):456–467, 1998.

[1860] F. Ruskey and J Sawada. An efficient algorithm for generating necklaces of fixed density. *SIAM J. Comput.*, 29(2):671–684, 1999.

[1861] F. Ruskey and J Sawada. An fast algorithm to generate necklaces. In *Proceedings of the 11th ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 256–262, San Francisco, CA, 2000.

[1862] F. Ruskey and J Sawada. Generating necklaces and strings with forbidden substrings. In D.-Z. Du, P. Eades, V. Estivill-Castro, X. Lin, and A. Sharma, editors, *Proceedings of the 6th Annual International Conference on Computing and Combinatorics*, number 1858 in Lecture Notes in Computer Science, pages 330–339, Sydney, Australia, 2000. Springer-Verlag, Berlin.

[1863] P. Ružička and I. Prívara. On tree pattern unification problems. In Z. Ésik, editor, *Proceedings of the Fundamental of Computation Theory*, number 710 in Lecture Notes in Computer Science, pages 418–429, Szeged, Hungary, 1993. Springer-Verlag, Berlin.

[1864] W. L. Ruzzo and M. Tompa. A linear time algorithm for finding all maximal scoring subsequences. In *Seventh International Conference on Intelligent Systems for Molecular Biology*, pages 234–241, Heidelberg, Germany, 1999.

[1865] W. Rytter. A correct preprocessing algorithm for Boyer-Moore string searching. *SIAM J. Comput.*, 9(3):509–512, 1980.

[1866] W. Rytter. The complexity of two-way pushdown automata and recursive programs. In A. Apostolico and Z. Galil, editors, *Combinatorial Algorithms on Words*, volume 12 of *NATO Advanced Science Institutes, Series F*, pages ?–? Springer-Verlag, Berlin, 1985.

[1867] W. Rytter. The space complexity of the unique decipherability problem. *Inf. Process. Lett.*, 23(1):1–3, 1986.

[1868] W. Rytter. On efficient computations of costs of paths of a grid graph. *Inf. Process. Lett.*, 29(2):71–74, 1988.

[1869] W. Rytter. On the parallel transformations of regular expressions to non-deterministic finite automata. *Inf. Process. Lett.*, 31(2):103–109, 1989.

[1870] W. Rytter and K. Diks. On optimal parallel computations for sequences of brakets. In R. Capocelli, editor, *Sequences: Combinatorics, Compression, Security and Transmission*, pages 92–105. Springer-Verlag, Berlin, 1990.

[1871] K. Sadakane. A fast algorithm for making suffix arrays and for Burrows-Wheeler transformation. In J. A. Storer and M. Cohn, editors, *Proceedings Data Compression Conference*, pages 129–138, Snowbird, UT, 1998. IEEE Computer Society Press.

[1872] M. Sagot and E. W. Myers. Identifying satellites in nucleic acid sequences. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 234–242, New York, NY, 1998. ACM Press.

[1873] M.-F. Sagot. *Ressemblance lexicale et structurale entre macromolécules – Formalisation et approches combinatoires.* PhD thesis, Université de Marne-la-Vallée, 1996. Rapport IGM 96-37.

[1874] M-F. Sagot. Spelling approximate repeated or common motifs using a suffix tree. In C. L. Lucchesi and A. V. Moura, editors, *Proceedings of the 3rd Latin American Symposium*, number 1380 in Lecture Notes in Computer Science, pages 374–390, Campinas, Brazil, 1998. Springer-Verlag, Berlin.

[1875] M.-F. Sagot, V. Escalier, A. Viari, and H. Soldano. Searching for repeated words in a text allowing for mismatches and gaps. In R. Baeza-Yates and U. Manber, editors, *Proceedings of the 2nd South American Workshop on String Processing*, pages 101–116, Valparaíso, Chile, 1995.

[1876] M.-F. Sagot and A. Viari. A double combinatorial approach to discovering patterns in biological sequences. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 186–208, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1877] M.-F. Sagot and A. Viari. Flexible identification of structural objects in nucleic acid sequences: palindromes, mirror repeats, pseudoknots and triple helices. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in

Lecture Notes in Computer Science, pages 224–246, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1878] M.-F. Sagot, A. Viari, and H. Soldano. Multiple sequence comparison: a peptide matching approach. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 366–385, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[1879] M.-F. Sagot, A. Viari, and H. Soldano. Multiple sequence comparison – a peptide matching approach. *Theor. Comput. Sci.*, 180(1–2):115–137, 1997.

[1880] S. C. Sahinalp and U. Vishkin. On a parallel-algorithms method for string matching problems (overview). In M. Bonuccelli, P. Crescenzi, and R. Petreschi, editors, *Algorithms and Complexity*, number 778 in Lecture Notes in Computer Science, pages 22–32, Rome, Italy, 1994. Springer-Verlag, Berlin.

[1881] S.C. Sahinalp and U. Vishkin. Symmetry breaking for suffix tree construction. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 300–309, Montréal, Canada, 1994.

[1882] Y. Sakakibara, M. Brown, R. Hughey, I. Saira Mian, K. Sjölander, R. C. Underwood, and D. Haussler. Recent methods for RNA modeling. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 289–306, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[1883] H. Sakoe and S. Chiba. A similarity evaluation of speech patterns by dynamic programming (in japanese). Technical report, Institute of Electronic Communications Engineering of Japan, 1970. Report.

[1884] H. Sakoe and S. Chiba. A similarity evaluation of speech patterns by dynamic programming (in japanese). In *Proceedings of the International Congress of Acoustics*, pages ?–?, Budapest, Hungary, 1971. Paper 20C.

[1885] H. Sakou, T. Miyatake, S. Kashioka, and M. Ejiri. A position recognition algorithm for semiconductor alignment based on structural pattern matching. *IEEE Trans. Acoust. Speech Signal Process.*, 37(12):2148, 1989.

[1886] Mikaël Salson, Thierry Lecroq, Martine Léonard, and Laurent Mouchard. Dynamic Burrows-Wheeler transform. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 13–25, Czech Technical University in Prague, Czech Republic, 2008.

[1887] G. Salton. Automatic information retrieval. *IEEE Trans. Comput.*, 13:41–55, 1980.

[1888] G. Salton. *Automatic text processing.* Addison-Wesley, Reading, MA, 1989.

[1889] H. Samet. Distance transform for images represented by quadtrees. *IEEE Trans. Pattern Anal. Mach. Intell.*, 4(3):298–303, 1982.

[1890] B. Sandak, R. Nussinov, and H. J. Wolfson. Docking of conformationally flexible proteins. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 271–287, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[1891] D. Sankoff. Matching sequences under deletion/insertion constraints. *Proc. Natl. Acad. Sci. U.S.A.*, 69:4–6, 1972.

[1892] D. Sankoff. Simultaneous solution of the rna folding, alignment and protosequence problems. *SIAM J. Appl. Math.*, 45:810–825, 1985.

[1893] D. Sankoff. Edit distance for genome comparison based on non-local operations. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 121–135, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[1894] D. Sankoff. The early introduction of dynamic programming into computational biology. In *Advances in Mathematical Sciences: CRM's 25 Years*, volume 11, pages ?–? American Mathematical Society, 1997.

[1895] D. Sankoff. The early introduction of dynamic programming into computational biology. *Bioinformatics*, 16:41–47, 2000.

[1896] D. Sankoff and J. B. Kruskal. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison.* Addison-Wesley, Reading, MA, 1983.

[1897] D. Sankoff and J. B. Kruskal. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison.* Cambridge University Press, 1999. 2nd edition.

[1898] D. Sankoff, M.-N. Parent, I. Marchand, and V. Ferretti. On the Nadeau-Taylor theory of conserved chromosome segments. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 262–274, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[1899] M. Santibanez and K. Rohde. A multiple alignment program for protein sequences. *Comput. Appl. Biosci.*, 3:111–114, 1987.

[1900] T. Sato. Fast full text search using tree structured [TS] file. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 42–63, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[1901] W. Saurin and P. Marliere. Comparaison de plusieurs séquences protéiques par reconnaissance de blocs conservés. *Comptes Rendus Acad. Sci. Paris*, 303:541–546, 1986.

[1902] S. A. Savari. Redundancy of the lempel-ziv string matching code. *IEEE Trans. Inf. Theory*, 44(2):787–791, 1998.

[1903] R. Schaback. On the expected sublinearity of the Boyer-Moore algorithm. *SIAM J. Comput.*, 17(4):648–658, 1988.

[1904] J. D. Schaffer and A. Morishima. An adaptive crossover distribution mechanism for genetic algorithms. In *Proceedings of the 2nd International Conference on Genetic Algorithms and Their Applications*, pages 36–40, 1987.

[1905] S. Schbath. Compound Poisson approximation of word counts in DNA sequences. *?*, 1:1–16, 1995.

[1906] S. Schbath. An efficient statistic to detect over- and under-represented words in DNA sequences. *J. Comput. Bio.*, 4:189–192, 1997.

[1907] S. Schbath. An overview on the distribution of word counts in Markov chains. *J. Comput. Bio.*, 7:193–201, 2000.

[1908] S. Schbath, B. Prum, and E. de Turckheim. Exceptional motifs in different Markov chain models for a statistical analysis of DNA sequences. *J. Comput. Bio.*, 2:417–437, 1995.

[1909] R. Scheihing. Combinatorial problems in strings. M.sc. thesis, Department of Computer Science, University of Chile, 1996.

[1910] C. Schensted. Largest increasing and decreasing subsequences. *Can. J. Math.*, 13:179–191, 1961.

[1911] B. Schieber and U. Vishkin. On finding lowest common ancestors: simplification and parallelization. *SIAM J. Comput.*, 17(6):1253–1262, 1988.

[1912] J. P. Schmidt. All highest scoring paths in weighted grid graphs and their application to finding all approximate repeats in strings. In *Proceedings of the 3rd Israel Symposium on Theory of Computing and Systems*, pages 67–77. IEEE Computer Society Press, 1995.

[1913] J. P. Schmidt. All highest scoring paths in weighted grid graphs and their application to finding all approximate repeats in strings. *SIAM J. Comput.*, 27(4):972–992, 1998.

[1914] U. Schmidt. *Motifs évitables dans les mots.* Thèse, report litp 86-63, Universités Paris 6 et 7, 1986.

[1915] U. Schmidt. Avoidable patterns on two letters. *Theor. Comput. Sci.*, 63(1):1–17, 1989.

[1916] P. Schnoebelen. Refined compilation of pattern-matching for functional languages. *Sci. Comp. Program.*, 11(2):133–159, 1988.

[1917] M. Schöniger and M. S. Waterman. A local algorithm for DNA sequence alignment with inversions. *Bull. Math. Biol.*, 54:521–536, 1992.

[1918] A. Schroeder. A statistical approach to the study of program behavior via reference string analysis. In K. M. Chandy and M. Reiser, editors, *Computer Performance*, pages 381–196. North-Holland, 1977.

[1919] G. D. Schuler, S. F. Altschul, and D. J. Lipman. A workbench for multiple sequence alignment construction and analysis. *PROTEINS: Struct., Funct. Genet.*, 9:180–190, 1991.

[1920] G. E. Schulz and R. H. Schirmer. Protein evolution. In *Principles of protein structure*, pages 166–205. Springer-Verlag, Berlin, 1979.

[1921] K. U. Schulz. Makanin's algorithm for word equations - two improvements and a generalization. In K. U. Schulz, editor, *Proceedings Word Equations and Related Topics*, number 572 in Lecture Notes in Computer Science, pages 85–150. Springer-Verlag, Berlin, 1990.

[1922] K. U. Schulz. Word unification and transformation of generalized equations. *J. Autom. Reasoning*, 11:149–184, 1993.

[1923] R. Sedgewick. *Algorithms.* Addison-Wesley, Reading, MA, 1988.

[1924] J. Seiferas and Z. Galil. Real time recognition of substring repetition and reversal. *Math. Syst. Theory*, 11:111–146, 1977.

[1925] R. C. Sekar, R. Ramesh, and I. V. Ramakrishnan. Adaptative pattern matching. Report, Suny at Stony Brook, 1992.

[1926] R. C. Sekar, R. Ramesh, and I. V. Ramakrishnan. Adaptative pattern matching. In W. Kuich, editor, *Proceedings of the 19th International Colloquium on Automata, Languages and Programming*, number 623 in Lecture Notes in Computer Science, pages 247–260, Wien, Austria, 1992. Springer-Verlag, Berlin.

[1927] R. C. Sekar, R. Ramesh, and I. V. Ramakrishnan. Adaptative pattern matching. *SIAM J. Comput.*, 24(6):1207–1234, 1995.

[1928] S. M. Selkow. The tree to tree editing problem. *Inf. Process. Lett.*, 6(6):184–186, 1977.

[1929] P. H. Sellers. An algorithm for the distance between two finite sequences. *J. Comb. Theory Ser. A*, 16:253–258, 1974.

[1930] P. H. Sellers. On the theory and computation of evolutionary distance. *SIAM J. Appl. Math.*, 26:787–793, 1974.

[1931] P. H. Sellers. Pattern recognition in genetic sequences. *Proc. Natl. Acad. Sci. U.S.A.*, 76:3041, 1979.

[1932] P. H. Sellers. The theory and computation of evolutionary distances: Pattern recognition. *J. Algorithms*, 1(4):359–373, 1980.

[1933] P. H. Sellers. Pattern recognition in genetic sequences by mismatch density. *Bull. Math. Biol.*, 46:501–514, 1984.

[1934] P. H. Sellers. Pattern recognition in DNA. In R. Miura, editor, *Some Mathematical Questions in Biology: DNA Sequence Analysis*, volume 17 of *Mathematics in the Life Sciences*, pages 19–28. Am. Math. Soc., 1986.

[1935] I. Semba. An efficient string searching algorithm. *J. Inform. Process.*, 8:101–109, 1985.

[1936] D. Semé and J.-F. Myoupo. A parallel solution of the sequence alignment problem using BSR model. In A. El-Amawy and S. Q. Zheng, editors, *Proceedings of the 10th International Conference on Parallel and Distributed Computing Systems*, pages 357–362, New Orleans, LO, 1997.

[1937] J. Senellart. Fast pattern matching in indexed texts. *Theor. Comput. Sci.*, 237(1–2):239–262, 2000.

[1938] J. C. Setubal and J. Meidanis. *Introduction to Computional Molecular Biology.* PWS Publishing Company, 1997.

[1939] H. Shang and T. H. Merrett. Tries for approximate string matching. *IEEE Trans. Knowl. Data Eng.*, 8(4):540–547, 1996.

[1940] B. A. Shapiro and K. Zhang. Comparing multiple RNA secondary structures using three comparisons. *Comput. Appl. Biosci.*, 6(4):309–318, 1990.

[1941] L. G. Shapiro and R. M. Haralick. Structural descriptions and inexact matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 3(5):504–519, 1981.

[1942] D. Shasha and K. Zhang. Fast algorithms for the unit cost editing distance between trees. *J. Algorithms*, 11(4):581–621, 1990.

[1943] J. Shawe-Taylor. Fast string matching in a stationary ergodic source. Report RHBNC CSD-TR-633, University of London, 1990.

[1944] J. Shawe-Taylor and J.-Y. Kim. Fast two dimensional pattern matching. Report CSD-TR-91-25, University of London, 1991.

[1945] F. Shi. Fast approximate string matching with q-blocks sequences. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 257–271, Recife, Brazil, 1996. Carleton University Press.

[1946] F. Shi and D.-G. Shin. Centroid trees with application to string processing. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '99*, pages 40–60, Czech Technical University, Prague, Czech Republic, 1999. Collaborative Report DC–99–05.

[1947] Shibata, Kida, Fukamachi, Takeda, Shinohara, Shinohara, and Arikawa. Speeding up pattern matching by text compression. In *CIAC: Italian Conference on Algorithms and Complexity*, 2000.

[1948] Y. Shibata, T. Matsumoto, M. Takeda, A. Shinohara, and S. Arikawa. A Boyer-Moore type algorithm for compressed pattern matching. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 181–194, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1949] Y. Shibata, M. Takeda, A. Shinohara, and S. Arikawa. Pattern matching in text compressed by using antidictionaries. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 37–49, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1950] Y. Shibata, M. Takeda, A. Shinohara, and S. Arikawa. Pattern matching in text compressed by using antidictionaries. *J. Discret. Algorithms*, 1(1):257–269, 2000.

[1951] T. Shibuya and H. Imai. New flexible approaches for multiple sequence alignment. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 267–276, Santa Fe, NM, 1997. ACM Press.

[1952] Y. Shiloach. A fast equivalence-checking algorithm for circular lists. *Inf. Process. Lett.*, 8(5):236–238, 1979.

[1953] Y. Shiloach. Fast canonization of circular strings. *J. Algorithms*, 2(2):107–121, 1981.

[1954] T. Shinohara. Polynomial time inference of extended regular pattern languages. In *RIMS Symposia on Software Science and Engineering*, number 147 in Lecture Notes in Computer Science, pages 115–127. Springer-Verlag, Berlin, 1982.

[1955] T. Shinohara. Polynomial time inference of pattern languages and its applications. In *Proceedings of the 7th IBM Symposium on Mathematical Foundations of Computer Science*, pages 191–209, 1982.

[1956] T. Shinohara. Inferring unions of two pattern languages. *Bull. Inform. Cybern.*, 20:83–88, 1983.

[1957] M. Shishibori, M. Okada, T. Sumimoto, and J. Aoe. An efficient trie hashing method using a compact binary tree. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 1–11, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[1958] P. R. Sibbald. Notes on a very fast substring search algorithm. *Commun. ACM*, 35(4):133–135, 1992.

[1959] P. R. Sibbald and M. J. White. How probable are antibody cross-reactions ? *J. Theor. Biol.*, 127:163–169, 1987.

[1960] J. S. Sim, C. Iliopoulos, K. Park, and W.F. Smyth. Approximate periods of strings. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 123–133, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[1961] J. S. Sim, C. Iliopoulos, K. Park, and W.F. Smyth. Approximate periods of strings. *Theor. Comput. Sci.*, 262(1–2):557–?, 2001.

[1962] J. S. Sim and K. Park. The consensus string problem for a metric is NP-complete. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop On Combinatorial Algorithms*, pages 107–113, Perth, WA, Australia, 1999.

[1963] M. Šimánek. The factor automaton. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 102–106, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1964] I. Simon. An algorithm to distinguish words efficiently by their subwords. Manuscript, 1984.

[1965] I. Simon. Sequence comparison: some theory and some practice. In M. Gross and D. Perrin, editors, *Electronic Dictionaries and Automata in Computational Linguistics*, number 377 in Lecture Notes in Computer Science, pages 79–92, Saint Pierre d'Oléron, France, 1987. Springer-Verlag, Berlin.

[1966] I. Simon. Sequence comparison: some theory and some practice. In M. Gross and D. Perrin, editors, *Electronic, Dictionaries and Automata in Computational Linguistics*, number 377 in Lecture Notes in Computer Science, pages 79–92. Springer-Verlag, Berlin, 1989.

[1967] I. Simon. String matching algorithms and automata. In R. Baeza-Yates and N. Ziviani, editors, *Proceedings of the 1st South American Workshop on String Processing*, pages 151–157, Universidade Federal de Minas Gerais, Brazil, 1993.

[1968] I. Simon. String matching algorithms and automata. In J. Karhumäki, H. Maurer, and Rozenberg G, editors, *Results and Trends in Theoretical Computer Science*, number 814 in Lecture Notes in Computer Science, pages 386–395, Graz, Austria, 1994. Springer-Verlag, Berlin.

[1969] I. Simon. String matching algorithms and automata. Rapport LITP 94.47, Universités Paris 6 et 7, France, 1994.

[1970] G. Singh, S. A. Smolka, and I. V. Ramakrishnan. Distributed algorithms for tree pattern matching. In J. van Leeuwen, editor, *Proceedings of the 2nd Workshop on Distributed Algorithms*, number 312 in Lecture Notes in Computer Science, pages 92–107, Amsterdam, The Netherlands, 1987. Springer-Verlag, Berlin.

[1971] R. Siromoney, L. Mathew, V. R. Dare, and K. G. Subramanian. Infinite lyndon words. *Inf. Process. Lett.*, 50(2):101–104, 1994.

[1972] S. Skiena and P. Sumazin. Shift error detection in standardized exams. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 264–276, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[1973] S. Skiena and G. Sundaram. Reconstructing strings from substrings (extended abstract). In F. Dehne, J.-R. Sack, N. Santoro, and S. Whitesides, editors, *Proceedings of the 3rd Workshop on Algorithms and Data Structures*, number 709 in Lecture Notes in Computer Science, pages 565–576, Montréal, Canada, 1993. Springer-Verlag, Berlin.

[1974] S. Skiena and G. Sundaram. Reconstructing strings from substrings. *J. Comput. Bio.*, 2(2):333–354, 1996.

[1975] T. Skopal, V. Snášel, and M. Krátký. Image recognition using finite automata. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '02*, pages 88–95, Czech Technical University, Prague, Czech Republic, 2002. Research Report DC–2002–03.

[1976] A. O. Slisenko. Recognition of palindromes by multihead turing machines. In V. P. Orverkov and N. A. Sonin, editors, *Problems in the Constructive Trend in Mathematics VI (Proceedings of the Steklov Institute of Mathematics 129)*, pages 30–202, 1973.

[1977] A. O. Slisenko. String-matching in real-time: some properties of the data structure. In J. Winkowski, editor, *Proceedings of the 7th Symposium on

Mathematical Foundations of Computer Science, number 64 in Lecture Notes in Computer Science, pages 493–496, Zakopane, Poland, 1978. Springer-Verlag, Berlin.

[1978] A. O. Slisenko. Determination in real time of all the periodicities in a word. *Sov. Math. Dokl.*, 21:392–295, 1980.

[1979] A. O. Slisenko. Detection of periodicities and string matching in real time. *J. Sov. Math.*, 22:1316–1386, 1983.

[1980] G. de V. Smit. A comparison of three string matching algorithms. *Softw. Pract. Exp.*, 12:57–66, 1982.

[1981] D. Smith and E. Pierzchala. An algorithm and architectures for approximate string matching. In *Proceedings of the 33rd Midwest Symposium on Circuits and Systems*, pages 736–739, 1990.

[1982] P. D. Smith. Experiments with a very fast substring search algorithm. *Softw. Pract. Exp.*, 21(10):1065–1074, 1991.

[1983] P. D. Smith. On tuning the Boyer-Moore-Horspool string searching algorithms. *Softw. Pract. Exp.*, 24(4):435–436, 1994.

[1984] R. F. Smith and T. F. Smith. Automatic generation of primary sequence patterns from set of related protein sequences. *Proc. Natl. Acad. Sci. U.S.A.*, 87:118–122, 1990.

[1985] R. F. Smith and T. F. Smith. Pattern-induced multiple-sequence alignment (PIMA) algorithm employing secondary structure dependent gap penalties for use in comparative protein modelling. *Protein Eng.*, 5:35–42, 1992.

[1986] T. F. Smith and M. S. Waterman. Identification of common molecular sequences. *J. Mol. Biol.*, 147:195–197, 1981.

[1987] T. F. Smith, M. S. Waterman, and W. M. Fitch. Comparative biosequence metrics. *J. Mol. Biol.*, 18:38–46, 1981.

[1988] Sebastian Smyczyński. Constant-memory iterative generation of special strings representing binary trees. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 183–191, Czech Technical University in Prague, Czech Republic, 2009.

[1989] W. F. Smyth. Repetitive perhaps, but not boring. In C. S. Iliopoulos, editor, *Proceedings of the 9th Australasian Workshop On Combinatorial Algorithms*, pages 64–73, Perth, WA, Australia, 1998.

[1990] W. F. Smyth. Repetitive perhaps, but not boring. *Theor. Comput. Sci.*, 249(5):343–355, 2000.

[1991] W. F. Smyth. *Computing Patterns in Strings.* Addison Wesley Pearson, 2003.

[1992] William F. Smyth, Shu Wang, and Mao Yu. An adaptive hybrid pattern-matching algorithm on indeterminate strings. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 95–107, Czech Technical University in Prague, Czech Republic, 2008.

[1993] V. Snášel. Approximate string matching by fuzzy automata. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, page 101, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[1994] V. Snášel and T. Koutný. Algebra of pattern matching automata. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 61–68, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[1995] E. Sobel and H. Martinez. A multiple sequence alignment program. *Nucleic Acids Res.*, 14:363–374, 1986.

[1996] A. Sorgen and T. Søndergaard. On a fast string searching algorithm. *Commun. ACM*, 22(12):679, 1979.

[1997] J.-C. Spehner. La reconnaissance des facteurs d'un mot dans un texte. *Theor. Comput. Sci.*, 48(1):35–52, 1986.

[1998] J.-C. Spehner. Le calcul rapide des mélanges de deux mots. *Theor. Comput. Sci.*, 47(2):181–203, 1986.

[1999] J.-C. Spehner. La reconnaissance des facteurs d'un langage fini dans un texte en temps linéaire. In T. Lepistõ and A. Salomaa, editors, *Proceedings of the 15th International Colloquium on Automata, Languages and Programming*, number 317 in Lecture Notes in Computer Science, pages 547–560, Tampere, Finland, 1988. Springer-Verlag, Berlin.

[2000] H. Spencer. Faster string functions. In *Proceedings of the USENIX Winter 1992 Conference*, pages 419–428, San Francisco, CA, 1992.

[2001] D. Spinellis. Declarative peephole optimization using string pattern matching. *SIGPLAN Not.*, 34(2):47–51, 1999.

[2002] M. A. Sridhar. Efficient algorithms for multiple pattern matching. Report Computer Sciences 661, University of Wisconsin-Madison, 1986.

[2003] Y.V. Srinivas. *Pattern matching: A sheaf-theoretic approach.* Ph. d. thesis, University of California, Irvine, 1991.

[2004] R. Staden. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.*, 10(15):4731–4751, 1982.

[2005] I. Stadnyk. Schema recombination in a pattern recognition problem. In *Proceedings of the 2nd International Conference on Genetic Algorithms and Their Applications*, pages 27–35, 1987.

[2006] J. M. Steele. Long common subsequences and the proximity of two random strings. *SIAM J. Appl. Math.*, 14(4):731–737, 1982.

[2007] Y. Steinberg and M. Gutman. An algorithm for source coding subject to a fidelity criterion, based on string matching. *IEEE Trans. Inf. Theory*, 39(3):877–886, 1993.

[2008] A. O. Steinhardt. Correlation matching by finite length sequences. *IEEE Trans. Acoust. Speech Signal Process.*, 36(4):545, 1988.

[2009] W. H. Stellhorn. A processor for direct scanning of text. In *Proceedings of the 1st Workshop on Computer Architecture for Non-numeric Processing*, pages ?–?, New York, NY, 1974. ACM Press.

[2010] G. A. Stephen. String search. Report TR-92-gas-01, University College of North Wales, 1992.

[2011] G. A. Stephen. *String searching algorithms*. World Scientific Press, 1994.

[2012] J. A. Storer. Np-completeness results concerning data compression. Report 234, Princeton University, 1977.

[2013] J. A. Storer. *Data Compression: Methods and Theory*. Computer Science Press, 1988.

[2014] J. A. Storer. *Image and Text Compression*. Kluwer Academic Press, 1992.

[2015] J. A. Storer and T. G. Szymanski. The macro model for data compression. In *Proceedings of the 10th ACM Symposium on the Theory of Computing*, pages 30–39, San Diego, CA, 1978. ACM Press.

[2016] J. A. Storer and T. G. Szymanski. Data compression via textual substitution. *J. Assoc. Comput. Mach.*, 29(4):928–951, 1982.

[2017] J. Stoye and D. Gusfield. Simple and flexible detection of contiguous repeats using a suffix tree. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 140–152, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[2018] Tinus Strauss, Derrick G. Kourie, and Bruce W. Watson. A concurrent specification of an incremental DFA minimisation algorithm. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 218–226, Czech Technical University in Prague, Czech Republic, 2008.

[2019] S. S. Sturrock. Biological sequence comparisons on a transputer network. Master's thesis, Univ. Kent, 1991.

[2020] S. S. Sturrock and J. F. Collins. Fast Protein Database Searching on the MasPar (MP1104). *?*, ?:?–?, 1994.

[2021] S. S. Sturrock and I. Salmon. Application of occam to biological sequence comparisons. In *Proc. World Occam and Transputer User Group*, number 14, pages 181–190, 1991.

[2022] S. Subbiah and S. C. Harrison. A method for multiple sequence alignments with gaps. *J. Mol. Biol.*, 209:539–548, 1989.

[2023] D. M. Sunday. A very fast substring search algorithm. *Commun. ACM*, 33(8):132–142, 1990.

[2024] E. Sutinen and J. Tarhio. On using q-gram locations in approximate string matching. In P. G. Spirakis, editor, *Proceedings 3rd Annual European Symposium*, number 979 in Lecture Notes in Computer Science, pages 327–340, Corfu, Greece, 1995. Springer-Verlag, Berlin.

[2025] E. Sutinen and J. Tarhio. Filtration with q-samples in approximate string matching. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 50–63, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[2026] O. Sýkora and I. Vrto. Tight chip area lower bounds for string matching. *Inf. Process. Lett.*, 26(3):117–119, 1987.

[2027] W. Szpankowski. Some results on V-ary assymetric tries. *J. Algorithms*, 9(2):224–244, 1988.

[2028] W. Szpankowski. Patricia tries again revisited. *J. Assoc. Comput. Mach.*, 37(4):691–711, 1990.

[2029] W. Szpankowski. On the height of digital trees and related problems. *Algorithmica*, 6:256–277, 1991.

[2030] W. Szpankowski. Suffix trees revisited: (un)expected asymptotic behaviors. Report CSD-TR-91-063, Purdue University, 1991.

[2031] W. Szpankowski. Analysis of generalized suffix trees. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 664 in Lecture Notes in Computer Science, pages 1–14, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[2032] W. Szpankowski. (un)expected behavior of typical suffix trees. In *Proceedings of the 3rd ACM-SIAM Annual Symposium on Discrete Algorithms*, pages 422–431, Orlando, FL, 1992.

[2033] W. Szpankowski. Asymptotic properties of data compression and suffix trees. *IEEE Trans. Inf. Theory*, 39, 1993.

[2034] W. Szpankowski. A generalized suffix tree and its (un)expected asymptotic behaviors. *SIAM J. Comput.*, 22(6):1176–1198, 1993.

[2035] T. G. Szymanski. A special case of the maximal common subsequence problem. Report TR-170, Computer Science Laboratory, Princeton University, 1975.

[2036] K. C. Tai. The tree-to-tree correction problem. *J. Assoc. Comput. Mach.*, 26(3):422–433, 1979.

[2037] K. Tajima. Multiple DNA and protein sequence alignment on a workstation and a supercomputer. *Comput. Appl. Biosci.*, 4:467–471, 1988.

[2038] K. Takahashi, H. Yamada, H. Nagai, and M. Hirata. Intelligent string search processor to accelearte text information retrieval. In M. Kitsuregawa and H. Tanaka, editors, *Proceedings of the 5th International Workshop on Database Machines*, pages 297–310, 1987.

[2039] K. Takahashi, H. Yamada, H. Nagai, and K. Matsumi. A new string search hardware architecture for VLSI. In *Proceedings of the 13th Annual International Symposium on Computer Architecture*, pages 20–27, Washington, DC, (published as a special issue of Computer Architecture News ACM SIGARCH), 14(2), 1986. IEEE Computer Society Press.

[2040] T. Takaoka. An on-line pattern matching algorithm. *Inf. Process. Lett.*, 22(6):329–330, 1986.

[2041] T. Takaoka. Approximate pattern matching with samples. In Ding-Zhu Du and Xiang sun Zhang, editors, *Proceedings of the 5th International Symposium on Algorithms and Computation*, number 834 in Lecture Notes in Computer Science, pages 236–242, Beijing, P.R. China, 1994. Springer-Verlag, Berlin.

[2042] T. Takaoka. A left-to-right preprocessing computation for the Boyer-Moore string matching algorithm. *Comput. J.*, 39(5):413–416, 1996.

[2043] M. Takeda. A proof of the correctness of uratani's string searching algorithm. Report RIFIS-TR-CS7, Kyushu University, Japan, 1988.

[2044] M. Takeda. A fast matching algorithm for pattern with pictures. *Bull. Inform. Cybern.*, 25(3-4):137–153, 1993.

[2045] M. Takeichi and Y. Akama. Deriving a functional Knuth-Morris-Pratt algorithm by transformation. *Inf. Process. Lett.*, 13(4):522–528, 1990.

[2046] M. Tamm. Blitzfindig: Texte schnell durchsuchen mit T-search. *c't*, 8/97:292, 1997.

[2047] S. Tang, K. Zhang, and X. Wu. Matching with matrix norm minimization. In M. Crochemore and D. Gusfield, editors, *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, number 807 in Lecture Notes in Computer Science, pages 250–258, Asilomar, CA, 1994. Springer-Verlag, Berlin.

[2048] J. Tarhio. A Boyer-Moore approach for two-dimensional matching. Report, Division of Computer Science, University of California at Berkeley, CA, 1993.

[2049] J. Tarhio and H. Peltola. String matching in the DNA alphabet. *Softw. Pract. Exp.*, 27(7):851–861, 1997.

[2050] J. Tarhio and E. Ukkonen. A greedy approximation algorithm for constructing shortest common superstrings. In J. Gruska, B. Rovan, and J. Wiedermann, editors, *Proceedings of the 12th Symposium on Mathematical Foundations of Computer Science*, number 233 in Lecture Notes in Computer Science, pages 602–610, Bratislava, Czechoslovakia, 1986. Springer-Verlag, Berlin.

[2051] J. Tarhio and E. Ukkonen. A greedy approximation algorithm for constructing shortest common superstrings. *Theor. Comput. Sci.*, 57(1):131–145, 1988.

[2052] J. Tarhio and E. Ukkonen. Approximate Boyer-Moore string matching. Report A-1990-3, Department of Computer Science, University of Helsinki, Finland, 1990.

[2053] J. Tarhio and E. Ukkonen. Boyer-Moore approach to approximate string matching. In J. R. Gilbert and R. Karlsson, editors, *Proceedings of the 2nd Scandinavian Workshop on Algorithm Theory*, number 447 in Lecture Notes in Computer Science, pages 348–359, Bergen, Norway, 1990. Springer-Verlag, Berlin.

[2054] J. Tarhio and E. Ukkonen. Approximate Boyer-Moore string matching. *SIAM J. Comput.*, 22(2):243–260, 1993.

[2055] R. E. Tarjan and A. C.-C. Yao. Storing a sparse table. *Commun. ACM*, 22(11):606–611, 1979.

[2056] R. L. Tatusov and E.V. Koonin. A simple tool to search for sequence motifs that are conserved in BLAST outputs. *Comput. Appl. Biosci.*, 10(4):457–459, 1994.

[2057] W. R. Taylor. Multiple sequence alignment by a pairwise algorithm. *Comput. Appl. Biosci.*, 3:81–87, 1987.

[2058] W. R. Taylor. A flexible method to align large numbers of biological sequences. *J. Mol. Evol.*, 28:161–169, 1988.

[2059] W. R. Taylor. Hierarchical method to align large numbers of biological sequences. *Methods Enzymol.*, 188:456–474, 1990.

[2060] S.-H. Teng and F. Yao. Approximating shortest superstrings. In *Proceedings of the 34th IEEE Annual Symposium on Foundations of Computer Science*, pages 158–165, Palo Alto, CA, 1993. IEEE Computer Society Press.

[2061] S.-H. Teng and F. Yao. Approximating shortest superstrings. *SIAM J. Comput.*, 26(2):410–417, 1997.

[2062] K. Thompson. Regular expression search algorithm. *Commun. ACM*, 11:419–422, 1968.

[2063] A. Thue. Über die gegenseitige lage gleicher teile gewisser zeichenreichen. *Nor. Vidensk. Selsk. Skr. Mat. Nat. Kl.*, 1:1–67, 1912.

[2064] W. Tichy. The string-to-string correction problem with block moves. *ACM Trans. Comput. Syst.*, 2:309–321, 1984.

[2065] V. G. Timkovsky. The complexity of subsequences, supersequences and related problem. *Kybern.*, 5:1–13, 1989.

[2066] I. Tomescu. On words containing all short subwords. *Theor. Comput. Sci.*, 197(1–2):235–240, 1998.

[2067] M. Tompa. Finding motifs using random projections. In *Proceedings Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 67–74, Montreal, Canada, 2001.

[2068] L. Toniolo. σ-irreducible strings. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 272–279, Recife, Brazil, 1996. Carleton University Press.

[2069] S. La Torre, M. Napoli, and D. Parente. Parallel word substitution. *Fundam. Inform.*, 27(1):27–36, 1996.

[2070] Avraham N. Trahtman. The road coloring and Černy conjecture. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2008*, pages 1–12, Czech Technical University in Prague, Czech Republic, 2008.

[2071] N. Tran. An easy case of sorting by reversals. In A. Apostolico and J. Hein, editors, *Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching*, number 1264 in Lecture Notes in Computer Science, pages 83–89, Aarhus, Denmark, 1997. Springer-Verlag, Berlin.

[2072] E. Trichina and S. Eriksson-Bique. Parallel solution to parentheses and pattern matching via near homomorphisms. In J. Edwards, editor, *Computer Science. Proceedings of the 22nd Australasian Computer Science Conference, ACSC'99*, Lecture Notes in Computer Science, pages 205–216, Auckland, New-Zealand, 1999. Springer-Verlag, Singapore.

[2073] J. Tromp and J. Shallit. Subword complexity of ageneralized the-morse word. *Inf. Process. Lett.*, 54(6):313–316, 1995.

[2074] Z. Troníček. Operations on DASG. In J.-M. Champarnaud, D. Maurel, and D. Ziadi, editors, *Proceedings of the 3rd International Workshop on Implementing Automata'98*, number 1660 in Lecture Notes in Computer Science, pages 82–91, Rouen, France, 1999. Springer-Verlag, Berlin.

[2075] Z. Troníček and B. Melichar. Directed acyclic subsequence graph. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 107–118, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[2076] W.-H. Tsai and S.-S. Yu. Attributed string matching with merging for shape recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 7(4):453–462, 1985.

[2077] J. Turner. Approximation algorithms for the shortest common superstring problem. Report WUCS-86-16, Department of Computer Science, Washington University, St. Louis, MO, 1986.

[2078] J. Turner. Approximation algorithms for the shortest common superstring problem. *Inf. Comput.*, 83(1):1–20, 1989.

[2079] J.T. Udding. The maximum length of a palindrome in a sequence. In W. H. J. Feijen, A. J. M. van Gasteren, D. Gries, and J. Misra, editors, *Beauty is Our Business*, chapter 49, pages 410–416. Springer-Verlag, Berlin, 1990.

[2080] N. Ukiyama and H. Imai. Parallel multiple alignments and their implementation on cm5. In *Proceedings of the 4th Genome Inform. Workshop*, pages 103–108. Universal Academy Press, 1993.

[2081] E. Ukkonen. Algorithms for approximate string matching. *Inf. Control*, 64(1–3):100–118, 1985.

[2082] E. Ukkonen. Finding approximate patterns in strings. *J. Algorithms*, 6(1–3):132–137, 1985.

[2083] E. Ukkonen. A linear time algorithm for finding approximate shortest common superstring. Report C-1987-31, Department of Computer Science, University of Helsinki, 1987.

[2084] E. Ukkonen. A linear time algorithm for finding approximate shortest common superstring. *Algorithmica*, 5:313–323, 1990.

[2085] E. Ukkonen. Approximate string matching with q-grams and maximal matches. *Theor. Comput. Sci.*, 92(1):191–212, 1992.

[2086] E. Ukkonen. Constructing suffix trees on-line in linear time. In J. van Leeuwen, editor, *Proceedings of the 12th IFIP World Computer Congress*, pages 484–492, Madrid, Spain, 1992. North-Holland.

[2087] E. Ukkonen. Approximate string matching over suffix trees. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 228–242, Padova, Italy, 1993. Springer-Verlag, Berlin.

[2088] E. Ukkonen. On-line construction of suffix trees. Report A-1993-1, Department of Computer Science, University of Helsinki, Finland, 1993.

[2089] E. Ukkonen. On-line construction of suffix trees. *Algorithmica*, 14(3):249–260, 1995.

[2090] E. Ukkonen and D. Wood. A simple on-line algorithm to approximate string matching. Report, Department of Computer Science, Univeristy of Helsinki, Finland, 1989.

[2091] E. Ukkonen and D. Wood. Approximate string matching with suffix automata. Report A-1990-4, Department of Computer Science, Univeristy of Helsinki, Finland, 1990.

[2092] E. Ukkonen and D. Wood. Approximate string matching with suffix automata. *Algorithmica*, 10(5):353–364, 1993.

[2093] S. M. Ulam. Some combinatorial problems studied experimentally on computing machines. In S. K. Zaremba, editor, *Applications of Number Theory to Numerical Analysis*, pages 1–3. Academic Press, 1972.

[2094] S. M. Ulam. Some ideas and prospects in biomathematics. *Annual Review of Biophysics and Bioengineering*, 1:277–291, 1972.

[2095] J. R. Ullmann. An algorithm for subgraph isomorphism. *J. Assoc. Comput. Mach.*, 23(1):31–42, 1976.

[2096] J. R. Ullmann. A binary n-gram technique for automatic correction of substitution, deletion, insertion and reversal errors in words. *Comput. J.*, 20(2):141–147, 1977.

[2097] N. Uratani. A fast algorithm for matching patterns. Report, IECEJ, 1988.

[2098] G. Valiente and C. Martínez. An algorithm for graph pattern-matching. In R. Baeza-Yates, editor, *Proceedings of the 4th South American Workshop on String Processing*, pages 180–197, Valparaiso, Chile, 1997. Carleton University Press.

[2099] J. van der Woude. Playing with patterns, searching for strings. *Sci. Comp. Program.*, 12(3):177–190, 1989.

[2100] J. van der Woude. Playing with patterns, searching for strings. *Sci. Comp. Program.*, 12(3):177–190, 1989.

[2101] J. van Leeuwen and A. K. Tsakalidis. An optimal pointer machine algorithm for finding nearest common ancestors. Report RUU-CS-88-17, Department of Computer Science, University of Utrecht, The Netherlands, 1988.

[2102] Lynette van Zijl and Leendert Botha. Feature extraction for image pattern matching with cellular automata. In Jan Holub and Jan Žďárek, editors, *Proceedings of the Prague Stringology Conference 2009*, pages 3–14, Czech Technical University in Prague, Czech Republic, 2009.

[2103] V. M. Velichko and N. G. Zagoruyko. Automatic recognition of 200 words. *Int. J. Man-Mach. Stud.*, 2:223–234, 1970.

[2104] R. M. Verma. Strings, trees, and patterns. *Inf. Process. Lett.*, 41(3):157–161, 1992.

[2105] R. M. Verma, T. Krishnaprasad, and I. V. Ramakrishnan. An efficient parallel algorithm for term matching. In K. V. Nori, editor, *Proceedings of the 6th Conference on Foundations of Software Technology and Theoretical Computer Science*, number 241 in Lecture Notes in Computer Science, pages 504–518. Springer-Verlag, Berlin, 1986.

[2106] R. M. Verma and I. V. Ramakrishnan. Tight complexity bounds for term matching problems. *Inf. Comput.*, 101(1):33–69, 1992.

[2107] R. M Verma and S. W. Reyner. An analysis of a good algorithm for the subtree problem, corrected. *SIAM J. Comput.*, 18(5):906–908, 1989.

[2108] M. Vihinen. An algorithm for simultaneous comparison of several sequences. *Comput. Appl. Biosci.*, 4:89–92, 1988.

[2109] M. Vihinen. Simultaneous comparison of several sequences. *Methods Enzymol.*, 183:447–458, 1990.

[2110] J. Vilo. Discovering frequent patterns from strings. Technical Report C-1998-9, Department of Computer Science, University of Helsinki, Finland, 1998.

[2111] J. Vilo, A. Brāzma, I. Jonassen, A. Robinson, and E. Ukkonen. Mining for putative regulatory elements in the yeast genome using gene expression data. In *Proceedings International Conference on Intelligent Systems for Molecular Biology*, pages 384–394. AAAI Press, 2000.

[2112] M. Vingron. Multiple sequence alignment and applications in molecular biology. Report 91-12, University of Heidelberg, Germany, 1991.

[2113] M. Vingron and P. Argos. A fast and sensitive multiple sequence alignment algorithm. *Comput. Appl. Biosci.*, 5(2):115–121, 1989.

[2114] M. Vingron and P. Argos. Determination of reliable regions in protein sequence alignments. *Protein Eng.*, 7:565–569, 1990.

[2115] M. Vingron and P. Argos. Motif recognition and alignment for many sequences by comparison of dot-matrices. *J. Mol. Biol.*, 218:33–43, 1991.

[2116] M. Vingron and P. A. Pevzner. Multiple sequence comparison and n-dimensional image reconstruction. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 243–253, Padova, Italy, 1993. Springer-Verlag, Berlin.

[2117] M. Vingron and M. S. Waterman. Sequence alignment and penalty choice. Review of concepts, studies and implications. *J. Mol. Biol.*, 235:1–12, 1994.

[2118] T. K. Vintsyuk. Speech discrimination by dynamic programming. *Comput.*, 4(1):52–57, 1968.

[2119] U. Vishkin. Optimal parallel pattern matching in strings. *Inf. Control*, 67(1–3):91–113, 1985.

[2120] U. Vishkin. Deterministic sampling - a new technique for fast pattern matching. *SIAM J. Comput.*, 20(1):22–40, 1991.

[2121] J. S. Vitter. Design and analysis of dynamic huffman codes. *J. Assoc. Comput. Mach.*, 34(4):825–845, 1987.

[2122] K. P. Vo. More ¡curses¿: the ¡screen¿ library. Report, AT&T Bell Laboratories, 1986.

[2123] R. A. Wagner. Common phrases and minimum-space text storage. *Commun. ACM*, 16(3):148–152, 1973.

[2124] R. A. Wagner. Order-n correction for regular languages. *Commun. ACM*, 17(6):265–268, 1974.

[2125] R. A. Wagner. On the complexity of extended string-to-string correction probelm. In *Proceedings of the 7th ACM Symposium on the Theory of Computing*, pages 218–223, Albuquerque, NM, 1975. ACM Press.

[2126] R. A. Wagner. On the complexity of extended string-to-string correction probelm. In D. Sankoff and J. B. Kruskall, editors, *Time Warps, String Edits, and Macromolecules: the Theory and Practice of Sequence Comparison*, pages ?–? Addison-Wesley, Reading, MA, 1983.

[2127] R. A. Wagner and M. Fischer. The string-to-string correction problem. *J. Assoc. Comput. Mach.*, 21(1):168–173, 1974.

[2128] R. A. Wagner and J. I. Seiferas. Correcting counter automaton recognizable languages. *SIAM J. Comput.*, 7(3):357–375, 1978.

[2129] B. F. Wang, G. H. Chen, and K. Park. On the set LCS and set-set LCS problems. *J. Algorithms*, 14(3):466–477, 1993.

[2130] J. T. L. Wang, B. A. Shapiro, and D. Shasha, editors. *Pattern Discovery in Biomolecular Data: Tools, Techniques and Applications.* Oxford University Press, 1999.

[2131] J. T. L. Wang, K. Zhang, K. Jeong, and D. Shasha. A system for approximate tree matching. to appear in IEEE Trans. Knowledge Data Eng., 1989.

[2132] J. T. L. Wang, K. Zhang, K. Jeong, and D. Shasha. Reference manual for ATBE: a tool for approximate tree matching. Report TR551, Courant Institue of Mathematical Science, New York University, NY, 1991.

[2133] J. T. L. Wang, K. Zhang, K. Jeong, and D. Shasha. A tool for tree pattern matching. In *Proceedings of the 1991 IEEE International Conference on Tools for AI*, pages 436–444, San Jose, CA, 1991.

[2134] L. Wang and D. Gusfield. Improved approximation algorithms for tree alignment. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 220–233, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[2135] L. Wang and T. Jiang. On the complexity of multiple sequence alignment. *J. Comput. Bio.*, 1:337–348, 1994.

[2136] L. L. Wang, P. K. Ku, and W. H. Tsai. Model-Based Guidance by Longest Common Subsequence Algorithms for Indoor Autonomous Vehicle Navigation. *Automation in Construction*, 2:123–137, 1993.

[2137] Z. Wang and K. Zhang. Finding common rna secondary structures from RNA sequences. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 258–269, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[2138] M. S. Waterman. Sequence alignments in the neighborhood of the optimum with general application to dynamic programming. *Proc. Natl. Acad. Sci. U.S.A.*, 80:3123–3124, 1983.

[2139] M. S. Waterman. General methods of sequence comparison. *Bull. Math. Biol.*, 46(4):473–501, 1984.

[2140] M. S. Waterman. Multiple sequence alignments by consensus. *Nucleic Acids Res.*, 14:9095–9102, 1986.

[2141] M. S. Waterman. Sequence alignments. In M. S. Waterman, editor, *Mathematical methods for DNA sequences*, pages 53–92. CRC Press Inc., Boca Raton, FL, 1989.

[2142] M. S. Waterman. *Mathematical methods for DNA sequences.* CRC Press Inc., Boca Raton, FL, 1991.

[2143] M. S. Waterman. Sequence comparison and statistical significance in molecular biology. In T. Lengauer, editor, *Proceedings of the 1st Annual European Symposium*, number 726 in Lecture Notes in Computer Science, page 396, Bad Honnef, Germany, 1993. Springer-Verlag, Berlin.

[2144] M. S. Waterman. Parametric and ensemble sequence alignment. *Bull. Math. Biol.*, 56:743–767, 1994.

[2145] M. S. Waterman. *Introduction to Computational Biology.* Chapman & Hall, 1995.

[2146] M. S. Waterman, R. Arratia, and D. J. Galas. Pattern recognition in several sequences: consensus and alignment. *Bull. Math. Biol.*, 46:515–527, 1984.

[2147] M. S. Waterman and T. H. Byers. A dynamic programming algorithm to find all solutions in a neighborhood of the optimum. *Math. Biosci.*, 77:179–188, 1985.

[2148] M. S. Waterman and M. Eggert. A new algorithm for best subsequence alignments with applications to tRNA-rRNA comparisons. *J. Mol. Biol.*, 197:723–728, 1987.

[2149] M. S. Waterman, M. Eggert, and E. Lander. Parametric sequence comparisons. *Proc. Natl. Acad. Sci. U.S.A.*, 89:6090–6093, 1992.

[2150] M. S. Waterman, L. Gordon, and R. Arratia. Phase transitions in sequence matches and nucleic acid structure. *Proc. Natl. Acad. Sci. U.S.A.*, 84:1239–1243, 1987.

[2151] M. S. Waterman and R. Jones. Consensus methods for DNA and protein sequence alignment. *Methods Enzymol.*, 183:221–237, 1990.

[2152] M. S. Waterman and M. D. Perlwitz. Line geometries for sequence comparison. *Bull. Math. Biol.*, 46:567–577, 1984.

[2153] M. S. Waterman, T. F. Smith, and W. A. Beyer. Some biological sequence metrics. *Adv. Math*, 20:367–387, 1976.

[2154] B. W. Watson. An introduction to the FIRE engine: A C++ toolkit for FInite automata and Regular Expressions. Report, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1994.

[2155] B. W. Watson. The performance of single-keyword and multiple-keyword pattern matching algorithms. Computing Science Notes 94/19, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1994.

[2156] B. W. Watson. A taxonomy of finite automata construction algorithms. Report, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1994.

[2157] B. W. Watson. A taxonomy of finite automata minimization algorithms. Report, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1994.

[2158] B. W. Watson. A collection of new regular grammar pattern matching algorithms. In J. Holub, editor, *Proceedings of the Prague Stringologic Club Workshop '96*, pages 64–83, Czech Technical University, Prague, Czech Republic, 1996. Collaborative Report DC–96–10.

[2159] B. W. Watson. A new regular grammar pattern matching algorithm. In J. Díaz and M. Serna, editors, *Proceedings of the 4th Annual European Symposium*, number 1136 in Lecture Notes in Computer Science, pages 364–377, Barcelona, Spain, 1996. Springer-Verlag, Berlin.

[2160] B. W. Watson. The performance of single and multiple keyword pattern matching algorithms. In N. Ziviani, R. Baeza-Yates, and K. Guimarães, editors, *Proceedings of the 3rd South American Workshop on String Processing*, pages 280–294, Recife, Brazil, 1996. Carleton University Press.

[2161] B. W. Watson. A Boyer-Moore (or Watson-Watson) type algorithm for regular tree pattern matching. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 33–38, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[2162] B. W. Watson. SPARE Parts: A C++ toolkit for String PAttern REcognition. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 47–60, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[2163] B. W. Watson. A new family of Commentz-Walter-style multiple-keyword pattern matching algorithms. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '2000, Bratislava, Slovakia*, pages 71–76, Czech Technical University, Prague, Czech Republic, 2000. Collaborative Report DC–2000–03.

[2164] B. W. Watson and R. E. Watson. A new family of string pattern matching algorithms. In J. Holub, editor, *Proceedings of the Prague Stringology Club Workshop '97*, pages 12–23, Czech Technical University, Prague, Czech Republic, 1997. Collaborative Report DC–97–03.

[2165] B. W. Watson and R. E. Watson. An early-retirement plan for the states. In J. Holub and M. Šimánek, editors, *Proceedings of the Prague Stringology Club Workshop '98*, pages 119–124, Czech Technical University, Prague, Czech Republic, 1998. Collaborative Report DC–98–06.

[2166] B. W. Watson and G. Zwaan. A taxonomy of keyword pattern matching algorithms. Computing Science Notes 92/27, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1992.

[2167] B. W. Watson and G. Zwaan. A taxonomy of keyword pattern matching algorithms. In H. A. Wijshoff, editor, *Proceedings Computing Science in the Netherlands 93*, pages 25–39, SION, Stichting Mathematish Centrum, 1993.

[2168] B. W. Watson and G. Zwaan. A taxonomy of sublinear multiple keyword pattern matching algorithms. Computing Science Report 95/13, Department of Mathematics and Computing Science, Eindhoven University of Technology, The Netherlands, 1995.

[2169] B. W. Watson and G. Zwaan. A taxonomy of sublinear multiple keyword pattern matching algorithms. *Sci. Comp. Program.*, 27(2):85–118, 1996.

[2170] P. Weiner. Linear pattern matching algorithm. In *Proceedings of the 14th Annual IEEE Symposium on Switching and Automata Theory*, pages 1–11, Washington, DC, 1973.

[2171] Z. Wen. New algorithms for the LCA problem and the binary tree reconstruction problem. *Inf. Process. Lett.*, 51(1):11–16, 1994.

[2172] Z.-X. Wen and Z.-Y. Wen. Some properties of the singular words of the fibonacci word. In *Séminaire Lotharingien de Combinatoire*, pages 115–128, Gerolfingen, 1993. Inst. Rech. Math. Av., Université Louis Pasteur.

[2173] C. Wenk. Applying an edit distance to the matching of tree ring sequences in dendrochronology. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 223–242, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[2174] R. L. Wessner. Optimal alphabetic search trees with restricted maximal height. *Inf. Process. Lett.*, 4(4):90–94, 1976.

[2175] C. T. White, R. K. Singh, P. B. Reintjes, J. Lampe, B. W. Erikson, W. D. Dettloff, V. L. Chi, and S. F. Altschul. BioSCAN: A VLSI-Based System for Biosequences Analysis. In *IEEE Int. Conf. Comp. Design: VLSI Comp. Proc.*, pages 504–509, 1991.

[2176] R. Wilber. The concave least-weight subsequence problem revisited. *J. Algorithms*, 9(3):418–425, 1988.

[2177] W. J. Wilbur and D. J. Lipman. Rapid similarity searches of nucleic acid protein data banks. *Proc. Natl. Acad. Sci. U.S.A.*, 80:726–730, 1983.

[2178] W. J. Wilbur and D. J. Lipman. The context-dependent comparison of biological sequences. *SIAM J. Appl. Math.*, 44:557–567, 1984.

[2179] P. Willet. Software and hardware techniques for string seraching in serial document databases. *World Pat. Inf.*, 10(2):120–129, 1988.

[2180] I. H. Witten. Browsing around a digital library: Today and tomorrow. In R. Giancarlo and D. Sankoff, editors, *Proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching*, number 1848 in Lecture Notes in Computer Science, pages 12–26, Montréal, Canada, 2000. Springer-Verlag, Berlin.

[2181] C. K. Wong and A. K. Chandra. Bounds for the string editing problem. *J. Assoc. Comput. Mach.*, 23(1):13–16, 1976.

[2182] J. A. Woods. More pep for Boyer-Moore grep. In *Usenet archive comp. sources. unix, egrep*, volume 4, 1987.

[2183] A. H. Wright. Approximate string matching using within-word parallelism. *Softw. Pract. Exp.*, 24(4):337–362, 1994.

[2184] K. Wright. *Inductive inference of pattern languages.* Ph. D. Thesis, University of Pittsburgh, 1989.

[2185] Wu and Wang. Two-dimensional object recognition through two-stage string matching. *IEEE Trans. Image Proc.*, 8:?–?, 1999.

[2186] S. Wu and U. Manber. Fast text searching with errors. Report TR-91-11, Department of Computer Science, University of Arizona, Tucson, AZ, 1991.

[2187] S. Wu and U. Manber. Agrep – a fast approximate pattern-matching tool. In *Proceedings USENIX Winter 1992 Technical Conference*, pages 153–162, San Francisco, CA, 1992.

[2188] S. Wu and U. Manber. Fast text searching allowing errors. *Commun. ACM*, 35(10):83–91, 1992.

[2189] S. Wu and U. Manber. A fast algorithm for multi-pattern searching. Report TR-94-17, Department of Computer Science, University of Arizona, Tucson, AZ, 1994.

[2190] S. Wu, U. Manber, and E. Myers. A subquadratic algorithm for approximate regular expression matching. *J. Algorithms*, 19(3):346–360, 1995.

[2191] S. Wu, U. Manber, and E. Myers. A subquadratic algorithm for approximate limited expression matching. *Algorithmica*, 15(1):50–67, 1996.

[2192] S. Wu, U. Manber, and E. W. Myers. Improving the running times for some string-matching problems. Report TR-91-20, Department of Computer Science, University of Arizona, Tucson, AZ, 1991.

[2193] S. Wu, U. Manber, E. W. Myers, and W. Miller. An $O(NP)$ sequence comparison algorithm. *Inf. Process. Lett.*, 35(6):317–323, 1990.

[2194] S. Wu, U. Manber, and G. Myers. A sub-quadratic algorithm for approximate limited expression matching. Report 36, Department of Computer Science, University of Arizona, Tucson, AZ, 1992.

[2195] A. D. Wyner, J. Ziv, and A. J. Wyner. On the role of pattern matching in information theory. *IEEE Trans. Inf. Theory*, 44(6):2045–2056, 1998.

[2196] H. Yamada, M. Hirata, H. Nagai, and K. Takahashi. A character string search processor. In *ISSCC Digest of Technical Papers, IEEE 34th International Solid-State Circuits Conference*, pages 272–273, 1987.

[2197] H. Yamada, M. Hirata, H. Nagai, and K. Takahashi. A high-speed string-search engine. *IEEE J. Solid-State Circuits*, SC-22(5):829–834, 1987.

[2198] Yang and Kieffer. On the performance of data compression algorithms based upon string matching. *IEEE Trans. Inf. Theory*, 44, 1998.

[2199] C. B. Yang and R. C. Lee. Systolic algorithms for the longest common subsequence problem. In *Int. Comp. Symp.*, pages 895–901, Taipei, Taiwan, Republic of China, 1984.

[2200] C. B. Yang and R. C. Lee. Systolic algorithms for the longest common subsequence problem. *J. Chinese Instit Engineers*, 10(6):691–699, 1987.

[2201] Y. P. Yang and T. Pavlidis. Optimal correspondence of string subsequences. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(11):1080–1087, 1993.

[2202] A. C. Yao. The complexity of pattern matching for a random string. *SIAM J. Comput.*, 8(3):368–387, 1979.

[2203] A. C. Yao and F. F. Yao. Dictionary look-up with small errors. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 387–394, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[2204] A. C. Yao and F. F. Yao. Dictionary look-up with one error. *J. Algorithms*, 25(1):194–202, 1997.

[2205] F. F. Yao. Speed-up in dynamic programming. *SIAM J. Algebr. Discret. Methods*, 3:532–540, 1982.

[2206] T. K. Yap, O. Frieder, and R. L. Martino. Parallel homologous sequence searching in large databases. In *Proceedings of the 5th Symposium on the Frontiers of Massively Parallel Computation*, pages 231–237, 1995.

[2207] T. K. Yap, O. Frieder, and R. L. Martino. *High Performance Computational Methods for Biological Sequence Analysis.* Kluwer Academic Publishers, 1996.

[2208] T. K. Yap, O. Frieder, and R. L. Martino. Parallel computation in biological sequence analysis. *IEEE Trans. Parallel Distr. Syst.*, 1998.

[2209] T. K. Yap, P. J. Munson, O. Frieder, and R. L. Martino. Parallel multiple sequence alignment using speculative computation. In *Proceedings of the International Conference on Parallel Processing*, pages III 60–III 67, 1995.

[2210] P. N. Yianilos. A dedicated comparator matches symbol strings fast and intelligently. *Electron.*, 56(5):113–117, 1983.

[2211] H. Yokoo. A dynamic data structure for reverse lexicographically sorted prefixes. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 150–162, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[2212] D. H. Younger. Recognition and parsing of context-free languages in time n^3. *Inf. Comput.*, 2:189–208, 1967.

[2213] H. Zantema. Longest segment problems. *Sci. Comp. Program.*, 18:39–66, 1992.

[2214] K. Zhang. A new editing based distance between unordered labeled trees. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 4th Annual Symposium on Combinatorial Pattern Matching*, number 684 in Lecture Notes in Computer Science, pages 254–265, Padova, Italy, 1993. Springer-Verlag, Berlin.

[2215] K. Zhang. On the approximation of longest common nonsupersequences and shortest common nonsubsequences. *Theor. Comput. Sci.*, 143(2):353–362, 1995.

[2216] K. Zhang. Efficient parallel algorithms for tree editing problems. In D. S. Hirschberg and E. W. Myers, editors, *Proceedings of the 7th Annual Symposium on Combinatorial Pattern Matching*, number 1075 in Lecture Notes in Computer Science, pages 361–372, Laguna Beach, CA, 1996. Springer-Verlag, Berlin.

[2217] K. Zhang. Efficient parallel algorithm for the editing distance between ordered trees. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 80–90, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[2218] K. Zhang and D. Shasha. Simple fast algorithms for the editing distance between trees and related problems. *SIAM J. Comput.*, 18(6):1245–1262, 1989.

[2219] K. Zhang and D. Shasha. Tree pattern matching. In A. Apostolico and Z. Galil, editors, *Pattern matching algorithms*, chapter 11, pages 341–371. Oxford University Press, 1997.

[2220] K. Zhang, D. Shasha, and J. T. L. Wang. Approximate tree matching in the presence of variable length don't cares. Report TR-328, Department of Computer Science, The University of Western Ontario, 1992.

[2221] K. Zhang, D. Shasha, and J. T. L. Wang. Fast serial and parallel algorithms for approximate tree matching with VLDC's. In A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, editors, *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching*, number 644 in Lecture Notes in Computer Science, pages 151–161, Tucson, AZ, 1992. Springer-Verlag, Berlin.

[2222] K. Zhang, D. Shasha, and J. T. L. Wang. Approximate tree matching in the presence of variable length don't cares. *J. Algorithms*, 16(1):33–66, 1994.

[2223] K. Zhang, R. Statman, and D. Shasha. On the editing distance between unordered labeled trees. *Inf. Process. Lett.*, 42:133–139, 1992.

[2224] K. Zhang, J. T. L. Wang, and D. Shasha. On the editing distance between undirected acyclic graphs and related problems. In Z. Galil and E. Ukkonen, editors, *Proceedings of the 6th Annual Symposium on Combinatorial Pattern Matching*, number 937 in Lecture Notes in Computer Science, pages 395–407, Espoo, Finland, 1995. Springer-Verlag, Berlin.

[2225] K. Zhang, L. Wang, and B. Ma. Computing similarity between RNA structures. In M. Crochemore and M. Paterson, editors, *Proceedings of the 10th Annual Symposium on Combinatorial Pattern Matching*, number 1645 in Lecture Notes in Computer Science, pages 281–293, Warwick University, UK, 1999. Springer-Verlag, Berlin.

[2226] K. Zhang, Y. Yang, X. Wang, J. Wang, and D. Shasha. An approximate oracle for distance in metric spaces. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 104–117, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[2227] Z. Zhang, P. Berman, and W. Miller. Alignment without low-scoring regions. *J. Comput. Bio.*, 5(2):197–210, 1998.

[2228] Z. Zhang, P. Berman, and W. Miller. Alignment without low-scoring regions. In S. Istrail, P. Pevzner, and M. Waterman, editors, *Proceedings of the 2nd Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 294–301, New York, NY, 1998. ACM Press.

[2229] Z. Zhang, W. R. Pearson, and W. Miller. Aligning a DNA sequence with a protein sequence. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 337–343, Santa Fe, NM, 1997. ACM Press.

[2230] R. F. Zhu and T. Takaoka. On improving the average case of the Boyer-Moore string matching algorithm. *J. Inform. Process.*, 10(3):173–177, 1987.

[2231] R. F. Zhu and T. Takaoka. A technique for two-dimensional pattern matching. *Commun. ACM*, 32(9):1110–1120, 1989.

[2232] B. Ziegler. Notes on a very fast substring search algorithm. *Commun. ACM*, 35(4):132–133, 1992.

[2233] R. Zimmer and T. Lengauer. Fast and numerically stable parametric alignment of biosequences. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB)*, pages 344–353, Santa Fe, NM, 1997. ACM Press.

[2234] M. Zipstein. Data compression with factor automata. *Theor. Comput. Sci.*, 92(1):213–221, 1992.

[2235] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Trans. Inf. Theory*, 23:337–343, 1977.

[2236] J. Ziv and A. Lempel. Compression of individual sequences via variable length coding. *IEEE Trans. Inf. Theory*, 24:530–536, 1978.

[2237] M. Ziv-Ukelson and A. Kershenbaum. A dictionary matching algorithm fast on the average for terms of varying lengths. In M. Farach-Colton, editor, *Proceedings of the 9th Annual Symposium on Combinatorial Pattern Matching*, number 1448 in Lecture Notes in Computer Science, pages 34–54, Piscataway, NJ, 1998. Springer-Verlag, Berlin.

[2238] J. Zobel and P. Dart. Finding approximate matches in large lexicons. *Softw. Pract. Exp.*, 25(3):331–345, 1995.

[2239] J. Zobel and P. W. Dart. Phonetic string matching: Lessons from information retrieval. In H.-P. Frei, D. Harman, P. Schäble, and R. Wilkinson, editors, *Proceedings of the 19th International Conference on Research and Development in Information Retrieval*, pages 166–172, Zurich, Switzerland, 1996. ACM Press.

[2240] J. Zobel, S. Heinz, and H. E. Williams. In-memory hash tables for accumulating text vocabularies. *Inf. Process. Lett.*, 80(6):271–277, 2001.

[2241] J. Zobel and A. Moffat. Similarity measures explored. Technical Report TR-95–3, Collaborative Information Technology Research Institute, Departments of Computer Science, RMIT and The University of Melbourne, Australia, 1995.

[2242] J. Zobel, A. Moffat, and K. Ramamohanarao. Inverted files versus signature files for text indexing. Technical Report TR-95–5, Collaborative Information Technology Research Institute, Departments of Computer Science, RMIT and The University of Melbourne, Australia, 1995.

[2243] J. Zobel, A. Moffat, and R. Sacks-Davis. An efficient indexing technique for full-text database systems. In L.-Y. Yuan, editor, *Proceedings of the 18th Conference on Very Large Databases*, pages 352–362, Vancouver, Canada, 1992.

[2244] J. Zobel, A. Moffat, and R. Sacks-Davis. Searching large lexicons for partially specified terms using compressed inverted files. In R. Agrawal, S. Baker, and D. Bell, editors, *Proceedings of the 19th Conference on Very Large Databases*, pages 290–301, Dublin, Ireland, 1993.

[2245] J. Zobel, A. Moffat, and R. Sacks-Davis. Storage management for files of dynamic records. In M. E. Orlowska and M. Papazoglou, editors, *Proceedings of the 4th Australian Database Conference*, pages 26–38, Brisbane, Australia, 1993. World Scientific.

[2246] J. Zobel, A. Moffat, R. Wilkinson, and R. Sacks-Davis. Efficient retrieval of partial documents. *Information Processing & Management*, 31(3):361–377, 1995.

[2247] M. Zuker. Suboptimal sequence alignment in molecular biology, alignment with error analysis. *J. Mol. Biol.*, 221:403–420, 1991.

[2248] M. Zuker and D. Sankoff. RNA secondary structures and their prediction. *Bull. Math. Biol.*, 46:591–621, 1984.