MAC5701 - Tópicos em Ciência da Computação
Plano de Estudos

Augusto Fernandes Vellozo

16 de abril de 2002

O plano de estudos montado para a disciplina MAC 5701 - Tópicos em Ciência da Computação foi baseado em artigos científicos, e capítulos de livros relevantes à área de Genômica Comparativa, em Bioinformática.

Pretendemos que com estas leituras o aluno possa embasar-se no estudo de artigos correlatos e técnicas aplicáveis à área.

Interessa-nos de forma particular o estudo, desenvolvimento e/ou combinação de algoritmos que possam eventualmente ser aplicáveis ao desenvolvimento e/ou aprimoramento de sistemas usados em Genômica Comparativa.

Alguns dos artigos em questão tratam de algoritmos usados na comparação de seqüências. Alguns tratam de algoritmos clássicos usados na procura de uma seqüência comum de maior comprimento, cada qual com melhor desempenho e aplicabilidade em casos particulares.

Outros artigos tratam do problema do estudo de eventos importantes na história evolutiva tais como inversões e demais rearranjos.

Interessa-nos em particular o estudo de inversões associados aos usuais algoritmos de alinhamento, especialmente aplicados à identificação de regiões altamente conservadas entre as seqüências de DNA de duas espécies próximas.

Para completar, há alguns poucos artigos que visam enriquecer o conhecimento geral do aluno com técnicas e conceitos aplicáveis à bioinformática, mesmo que não sejam aparentemente tão diretamente aplicáveis ao que se espera seja desenvolvolvido pelo aluno. Citamos por exemplo aplicações do uso de entropia para a extração de informação das seqüências de DNA.

Desta forma cremos que o conhecimento do aluno deverá ser bastante enriquecido, obtendo o mesmo assim uma boa base teórica e um amadurecimento que aumentarão as chances de bom desenvolvimento de seu programa de mestrado.

Ao final do semestre, será escrita uma visão geral do tema estudado e de como este estudo deverá se desenvolver. Um breve resumo da maioria dos artigos ora listados é esperado, sendo que de alguns deles se espera um estudo mais cuidadoso.

Referências

[1] A. Apostolico and C. Guerra. The longest common subsequence problem revisited. *Algorithmica*, 2(3):315–336, 1987.

[2] 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. *Inform. Process. Lett.*, 80(6):279–285, 2001.

[3] M. Crochemore and R. Vérin. Zones of low entropy in genomic sequences. *Computers and Chemistry*, 23(3-4):275–282, Jun 1999.

[4] A. L. Delcher, S. Kasif, R. D. Fleischmann, J. Peterson, O. White, and S. L. Salzberg. Alignment of whole genomes. *Nucleic Acids Research*, 27(11):2369–2376, Jun 1999.

[5] S. Hannenhalli and P. A. Pevzner. Towards a computational theory of genome rearrangements. In *Computer science today*, pages 184–202. Springer, Berlin, 1995.

[6] S. Hannenhalli and P. A. Pevzner. Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals. *J. ACM*, 46(1):1–27, 1999.

[7] J. W. Hunt and T. G. Szymanski. A fast algorithm for computing longest common subsequences. *Comm. ACM*, 20(5):350–353, 1977.

[8] G. M. Landau, E. W. Myers, and J. P. Schmidt. Incremental string comparison. *SIAM Journal on Computing*, 27(2):557–582, 1998.

[9] G. Myers. A fast bit-vector algorithm for approximate string matching based on dynamic programming. *Journal of the ACM*, 46(3):395–415, 1999.

[10] M. Schöniger and M. S. Waterman. A local algorithm for DNA sequence alignment with inversions. *Bull Math Biol*, 54(4):521–536, Jul 1992.

[11] A. Tridgell and P. Mackerras. The rsync algorithm. Technical Report TR-CS-96-05, Department of Computer Science, The Australian National University, Canberra 0200 ACT, Australia, June 1996.