Bioinfo Project

Benjamin Haibe-Kains

January 4, 2005

Goal

Implement a normalization method for Affymetrix© data using a parlalel computer architecture.

Materials and Methods

A very widespread microarray technology is the Affymetrix© technology (`http://wwww.affymetrix.com`). It is a short olginucleotide microarray platform.

One of the most used normalization method for this platform is the RMA method [Irizarry et al., 2003, Bolstad et al., 2003]. An implementation of the RMA method is available under R [Team, 2004] through the Bioconductor project [Bates et al., 2004].

The amount of data generated by microarray technology is large due to the huge number of gene expression measures for each experiment. Moreover the number of microarray experiments is increasing rapidly and traditional programming methods show their limits. It is the reason why it would be very interesting to implement the RMA normalization method on a parallel architecture in order to treat a large number of samples with high density Affymetrix© chips.

Such a parallel computer architecture is available at the Computer Science Department at the Université Libre de Bruxelles. An R interface to Parallel Virtual Machine (PVM) is installed on Lit5 (`http://www.lit5.ulb.ac.be`).

Good Work !

References

[Bates et al., 2004] Bates, D., Bolstad, B., Carey, V., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Gentleman, R., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Leisch, F., MacDonald, J., Maechler, M., Smith, C., Smyth, G., Rossini, A., Hutchinson, F., Sawitzki, G., Tierney, L., Yang, J., and Zhang, J. (2004). Bioconductor. http://www.bioconductor.org.

[Bolstad et al., 2003] Bolstad, B. M., Irizarry, R. A., Astrand, M., and TP, T. S. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–193.

[Irizarry et al., 2003] Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003). Summaries of affymetrix genechip probe level data. *Nuceli Acids Research*, 31(4):e15.

[Team, 2004] Team, R. D. C. (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3.