第１回　情報生命科学特別講義　2011 年 10 月 7 日（金）

Maximizing expected accuracy estimations in bioinformatics

Hamada Michiaki / 浜田道昭
The University of Tokyo, Graduate School of Frontier Sciences
Computational Biology
東京大学大学院新領域創成科学研究科生命情報科学専攻

Many estimation problems in bioinformatics are formulated as point estimation problems in a high-dimensional discrete space. In general, it is difficult to design reliable estimators for this type of problem, because the number of possible solutions is immense, which leads to an extremely low probability for every solution--even for the one with the highest probability. Therefore, maximum score and maximum likelihood estimators do not work well in this situation although they are widely employed in a number of applications. Maximizing expected accuracy (MEA) estimation, in which accuracy measures of the target problem and the entire distribution of solutions are considered, is a more successful approach. In this class, I provide an extensive review of algorithms and software based on MEA. I describe how a number of algorithms used in previous studies can be classified from the viewpoint of MEA. I believe that this classification will be useful not only for users wishing to utilize software to solve the estimation problems appearing in this class but also for developers wishing to design algorithms on the basis of MEA.