Project Proposal

Kinship inference of a population is an indispensable component of understanding genetic relatedness or coefficient of relationship. It helps biologists to understand more profoundly how traits inherit from parents to their offspring. It also includes study of mating process so as to correctly construct relationship between a parent and its offspring. One of the major steps in kinship analysis is the formalization of biological problems into computational problems which can then be solved using statistical or combinatorial approaches [1].

Various biological problems pertaining to kinship inference have already been formalized including but not limited to MIN-PARENT Problem [2], MIN MATING Problem and Parentage assignment problem.

Formalization of these problems have been very helpful to biologists for analyzing consanguinity and evolution of genetic traits among different generations. We have not been able to completely analyze these processes by only finding the minimum number of parents. There are many variances of these problems which are yet to be formalized and classified into easy and hard categories.

We aim to formalize different variances of these problems using various constraints and restrictions. Our focus is mainly to formalize Full Sibling Reconstruction and Parentage Assignment Problem. We propose to define and impose a set of constraints in order to generate different variances of the above problems. For example, Full Sibling Reconstruction problem can be constrained on the mating process, whereas Parentage Assignment Problem's variances can be constructed on the basis of unbalanced sex ratios or minimizing the number of females. We are also aiming to classify these problems into easy or hard categories accordingly.

Investigation of many fundamental biological phenomena including mating systems, selection and adaptation, kin selection, and dispersal patterns depends on the growing development and application of molecular markers. The power and potential of the genotypic information obtained in these studies often rests in our ability to reconstruct genealogical relationships among individuals. Formalization of these problems will prove to be a great advantage for above studies.

References

[1] Mary V. Ashley, Tanya Y. Berger-Wolf, Isabel C. Caballero, Wanpracha Chaovalitwongse, Bhaskar DasGupta, and Saad I. Sheikh. Full Sibling Reconstruction in Wild Populations from Microsatellite Genetic Markers.

[2] Ashley, Mary V., Tanya Y. Berger-Wolf, Wanpracha Chaovalitwongse, Bhaskar Dasgupta, Ashfaq Khokhar, and Saad Sheikh. On Approximating an Implicit Cover Problem in Biology. *Algorithmic Aspects in Information and Management Lecture Notes in Computer Science* (2009): 43-54. Web.