Challenge: A Multidisciplinary Degree Program in Bioinformatics

Mudasser F. Wyne
University of Michigan-Flint, Flint, Michigan USA

Abstract— **Bioinformatics is a new field that is poorly served by any of the traditional science programs in Biology, Computer science or Biochemistry. Known to be a rapidly evolving discipline, Bioinformatics has emerged from experimental molecular biology biochemistry as well as from the artificial intelligence, database, pattern recognition, and algorithms disciplines of computer science. While institutions are responding to this increased demand by establishing graduate programs in bioinformatics, entrance barriers for these programs are high, largely due to the significant prerequisite knowledge which is required, both in the fields of biochemistry and computer science. Although many schools currently have or are proposing graduate programs in bioinformatics, few are actually developing new undergraduate programs. In this paper I explore the blend of a multidisciplinary approach, discuss the response of academia and highlight challenges faced by this emerging field.**

Index Terms—**Bioinformatics, Database, Artificial Intelligence, Genetics.**

I. MULTIDISCIPLINARY PROGRAMM

In this section we highlight the meaning and importance of a multidiscipline program. Multidisciplinary program refers to the involvement of several different areas of study, though not necessarily in an integrated manner. Another definition for multidisciplinary program is a program in which two or more disciplines are brought together preferably in such a way that they interact with one another and each is able to affect the other's perspective. While knowledge structures are not exclusively interdisciplinary phenomena, the capacity to create meaningful connections across the knowledge domain is significantly facilitated by the introduction of multidisciplinary perspectives.

Multidisciplinary program differs from more traditional discipline specific programs. Discipline specific programs may be particularly valuable for those seeking specialization in a particular area. These programs may be overly fragmented and limited in scope to meet the goals of a student who is willing to explore other areas of interest in addition to pursuing their respective degree. Discipline specific degrees often fail to demonstrate how a particular discipline interfaces with another. As a consequence, students who are presented with information in an isolated manner tend to acquire knowledge in a specific discipline, for example computer science, and may fail to perceive or even question the overlapping values raised by other disciplines such as biology or chemistry.

Alternatively, multidisciplinary learning creates knowledge that is more holistic than the knowledge built in discipline-specific studies. Multidisciplinary approaches, while arguably less effective than traditional approaches for building the depth of single-subject knowledge, emphasize breadth of knowledge and seek meaningful connections between disciplines. Many argue that learners in multidisciplinary programs are guided beyond simpler forms of knowledge acquisition to a deeper assimilation of cross-disciplinary concepts. [12, 13].

II. AN OVERVIEW OF BIOINFORMATICS

Bioinformatics is defined as the application of tools of computation and analysis to the capture and interpretation of biological data. More recently, bioinformatics, classically known as computational biology, has expanded from mathematical, statistical and computing methods for solving biological problems to include analyzing and predicting the composition and structure of biomolecules and the modeling of how these basic building blocks interact in complex biological systems and regulatory networks [2].

Although, the term *bioinformatics* has a range of interpretations, the core activities of bioinformatics are widely acknowledged: storage, organization, retrieval and analysis of biological data. As an expert in the field of bioinformatics, one can provide services to the wider scientific community in the form of database, designer and administrator, as well as by developing analytical tools to extract and analyze data. Biological researchers ask questions related to the data; we strive to reply with comprehensible answers and accurate estimates of their reliability. By its nature, bioinformatics is a cross disciplinary field; biologists, computer scientists, mathematicians and chemists work together, each bringing a unique point of view. This makes for a stimulating and challenging environment, requiring constant and consistent effort to ensure good communication between people from each discipline [3].

Informatics has helped launch molecular biology into the genome era. The use of informatics to organize, manage and analyze genomic data (the genetic material of an organism) has become an important element of biology and medical research. A new science discipline “bioinformatics” combines computer science, mathematics, chemistry and biology to meet the many computational challenges in modern molecular biology and medical research. The two major themes in bioinformatics—data management and knowledge discovery heavily rely on effectively adopting techniques developed by computer scientist and mathematician. In addition, the internet also played a critical role. The World Wide Web has allowed researchers throughout the world to instantaneously share and access biological data captured in online community databases. The Information technologies ultimately produced the necessary speedup for collaborative research efforts in biology and thus helped genome researchers in timely completion of their projects. [4].

III WHY BIOINFORMATICS?

The notion that innovation is central to a study of computing seems extremely important. Innovation is the creation and adoption of novel practices to address newer challenges in this rapidly developing and ever growing world of science. Innovation can take various forms including novel products, new processes, contemporary functionality, innovative research insights, or unique business models. Innovation is a clear theme in government and industry, where it is seen as a source of wealth, competitiveness and professional success [1]. However, innovation is not simply constrained to the invention of novel technologies; it must include changing practice within organizations, communities and programs offered at academic institutes. The important point for our current discussion is that innovation in programs of study can be learned. A spirit of innovation can permeate throughout our courses, starting from the first year and building up expertise in innovation through to the final year. Starting early would help win new hearts and minds to bioinformatics. We believe the prospect of participating in, and causing innovations is highly appealing to students who want to make a difference in the world. Biology provides some of the most important and most complex, scientific challenges of our times. These problems include and are definitely not limited to, understanding of the human genome, discovery of the structure and functions of proteins that the genes encode, and usage of this information efficiently for drug design. Most of these problems are intensive at the least, from a computational perspective [7].

IV. DEGREE PROGRAMS

Over the past decade, bioinformatics has become crucial to any biomedical research enterprise. Most of the top schools already have, or are developing, research and educational programs in bioinformatics. Funding agencies have mounted special opportunities in bioinformatics. Groups of people who formerly would not – or could not – have now spoken about common research interests [10]. Both academia and the research community have responded to the need of the hour and engaged in various activities. A few universities, including American, Canadian, British, Swiss, and many others have started their own bioinformatics programs as well.

Most universities in USA offer degree programs in bioinformatics. The majority of these programs are at the graduate level and a few at the undergraduate level. Some universities are offering special tracks and options for bioinformatics in their regular computer science and other academic programs. On the international scene a few universities, including British and Swiss, offer master's degrees, and many universities have started their own bioinformatics programs. Unfortunately, efforts to coordinate corresponding programs have, on the most part, not been successful. Some of the most relevant efforts to create a common teaching structure are those implemented by the Universities of Bielefeld and Stockholm—both participating in the 5-Star consortium---and the University of Manchester which leads a bioinformatics project for the definition of standard teaching topics. [5, 15]. The Asia-Pacific region spans the Asian and Australasian continents as well as the Pacific-rim countries. As such, the seeds of bioinformatics in this region have been sown as early as 1989 in India, followed by Japan and Australia in 1991. While bioinformatics research, service and education have reached laudable heights in these countries, as well as in Singapore, Taiwan, Korea, Malaysia, New Zealand and Russia, several other countries (Thailand, Indonesia, and the Philippines to name three) are also making considerable progress [14].

A. Graduate Programs

Bioinformatics is a multidisciplinary field. Being a relatively new area of study, it is thus also, an intense focus of research. For all these reasons, most educational opportunities in the area are for graduates. Consequently, graduate programs in bioinformatics are beginning to emerge at universities worldwide [7]. Entrance requirements for such programs, require students with a specific prerequisite program of undergraduate study that is rarely made available to them as part of an organized schema. Graduate bioinformatics programs must currently accept students with undergraduate degrees in either computer science or biology and have sequences of remedial or prerequisite courses designed to complement the knowledge already acquired by the students as undergraduates. Students holding undergraduate degrees in computer science generally need to spend the majority of their first year of graduate study taking focused remedial courses in basic chemistry, biochemistry, molecular biology and genetics. Students holding undergraduate degrees in biology generally spend the majority of their first year of graduate study in course work covering introductory computer programming and data structures, databases and AI. The second year of a graduate bioinformatics program is generally dominated by preexisting graduate courses in computer science (AI, database, pattern recognition, and genetic algorithms) and biology (genetics, molecular biology, physiology, or ecology). Finally, students from either background require a course sequence covering contemporary algorithms and research techniques in bioinformatics. It is unlikely for this amount of material to be accommodated in a two-year course of study without significant preparation at the undergraduate level.

B. Undergraduate Programs

Although many schools currently have, or are in the process of proposing graduate programs in bioinformatics, few are creating undergraduate programs. In forming a graduate program, it is realized that there is the fundamental problem: where to find qualified graduate students. This quickly leads to the development of an undergraduate curriculum that is believed to serve students heading both to graduate school and to industry. There is of course, a wave of activity in the development of academic programs in bioinformatics [l, 3, 8].

Because of the demanding entrance requirements, graduate programs alone prove inadequate in providing the number of bioinformatics specialists that industry will require, partly because of the amount of remedial course work necessary. New undergraduate programs that incorporate a more specific (and shorter) biology sequence along with a more focused computer science foundation, need not be developed. Some of the traditional core courses in some computer science programs (such as assembly language programming) may need to be regarded as electives to allow students to increase their knowledge in the contemporary areas of IT knowledge applicable to bioinformatics (such as AI, knowledge representation, pattern recognition, and data mining). Four-year programs need to provide opportunities and direction to students to meet the market demand for bioinformatics professionals and to better prepare students for entrance into graduate-level bioinformatics programs. As bioinformaticians must be equally versed in the languages of biology and computer science, this effort will require a fundamental multidisciplinary approach.

For a typical under graduate program in bioinformatics, students would need to take 14 to 17 credits in mathematics (Calculus I and II, Discrete structures, statistics and may be bio-statistics), 18 credits in chemistry (General chemistry I and II, Organic chemistry I and II and Bio-chemistry), 18 credits in biology (General Biology I and II, Genetics, Cell biology and genomics) and 23 credits in computer science (Programming I and II, Data Structure, Database, AI, Data Warehousing and Advance Bioinformatics). There is also a need for developing courses like Bio-ethics and Bioinformatics to be jointly designed and taught by biology and computer science faculty. Most of the courses mentioned above are generally offered in under graduate programs in biology, chemistry and computer science.

V. CHALENGES OF BIOINFORMATICS PROGRAM

In a multidisciplinary degree program like bioinformatics, perhaps the greatest benefit is the opportunity to view issues from a multiplicity of perspectives. What could be more challenging? Today the rapidly changing intersection of science, society, and management provides a persistent intellectual challenge at one of the most dynamic times in the history of the world [11]. This means that understanding the cause and effect of emerging problems in various disciplines of science and finding ways to resolve them, is a worthy challenge.

In developing a multidisciplinary program, one first needs to assemble a team of faculty members from each department involved in the program, specifically, computer science, mathematics, chemistry and biology. There are advantages of such an effort, however they are not without a set of challenges that need to be addressed. These challenges, in preparing and managing such a multidisciplinary program, include (a) The time required in learning about parallel disciplines with respect to issues, course sequences, prerequisites etc. and their vocabulary. (b) Not all team members are of the same intellectual caliber or have the same commitment to the program. This means that team leadership must be willing to shoulder more of the load and at times, make tough decisions about the continuation of support or membership. (c) The challenge of actually proposing and developing such a program (as hard as that is) is often subordinate to the challenge of incorporating social and political insights into the form the science takes. This can cause major problems in dealing with team members who refuse to address broader issues related to the program. (d) Individual variations in depth of commitment to the program often slow the program development. Everyone is busy and individuals have different priorities. Explicit recognition of this helps minimize disruptions. (e) Disagreements over ownership of the program will arise, and unless the guidelines are agreed to in advance, there will be strife. (f) Finally, the commitment of time and energy into understanding other disciplines invariably detracts from the time and commitment put into maximizing one's own mastery of a single discipline.

One must also not forget that the time, effort, and costs associated with building the team, conducting the activities and developing the program, will exceed the institutional attention span of most funding agencies and decision makers. It takes significantly longer due to the effort and time needed to reach an agreement on a central theme, to decide on specific activities and other logistical details (such as course designs, prerequisites, course offering sequence and time), and then to cooperate. Everyone needs to share the same general vision but have specific parts of the overall program that they as individual departments or programs have ‘‘ownership’’ of. Finally what follows is the program proposal. As with any other proposal, the work plan needs to be concise, seamless, and well crafted to be successful. This requires special leadership by one person or a small group, in order to successfully assemble the final product.

VI CONCLUSION

Bioinformatics is a developing multidisciplinary science. The involvement of different sciences (such as computer science, biology, chemistry and mathematics) holds great promise. This century's major research and development efforts will likely be in the biological and health sciences. Computer science departments planning to diversify their offerings can thus only gain through early entry into bioinformatics. Even using minimal resources, such efforts are wise, since computer science graduates will enhance their employment qualifications. It is still premature to say whether bioinformatics will eventually become an integral part of computer science or whether it will develop into an independent and yet multidisciplinary degree program with biology, chemistry and mathematics as key partners. Regardless of the outcome, in the long run computer scientists are sure to ultimately benefit from being active and assertive partners with biologists.

REFERENCES

[1] J. Cohen, “Computer Science and Bioinformatics,” *Communications of the ACM*, vol. 48, No. 3, pp. 72-78, March 2005.

[2] B. A. Oakley and D. M. Hanna, “A Review of Nanobioscience and Bioinformatics Initiatives in North America”, IEEE Transactions on Nanoscience, Vol. 3, No. 1, March 2004.

[3] J. Baker and J. Thornton, “Software Engineering Challenges in Bioinformatics”, In the proceedings of the 26th International Conference on Software Engineering (ICSE'04), 2004.

[4] S. Ng and L. Wong, “Accomplishments and Challenges in Bioinformatics”, IT Professionals, pp. 44- 50, January/February 2004.

[5] A. Valencia, “ The European Perspective on bioinformatics”, IEEE Intelligent Systems, pp. 18-19, November/December 2001.

[6] T. Doom, M. Raymer, D. Krane and O. Garcia, “Crossing the Multidisciplinary Barrier: A Baccalaureate Computer Science Option in Bioinformatics”, IEEE Transactions on Education, Vol. 46, No. 3., August 2003.

[7] A. Zomaya, “Grid Computing: Opportunities for Bioinformatics Research”, Proceeding of the 3rd International conference on Information Technology and Applications (ICITA-05), July 4-7, 2005.

[8] J. Cohen, “Bioinformatics—An Introduction for Computer Scientists”, ACM Computing Surveys, Vol. 36, No. 2, pp. 122–158, June 2004.

[9] R. Hughey and Kevin Karplus, “Bioinformatics: A new field in Engineering Education”, In the proceedings of the 31st ASEE/IEEE Frontiers in Education Conference, pp. F2B 15-19, Oct. 10-13, 2001.

[10] S. Davidson, “Tale of Two Cultures: Are there Databases Research Issues in Bioinformatics”, In the proceedings of the 14th International Conference on Scientific and Statistical Database Management (SSDBM'02), July 2002.

[11] R. J. Naiman, “A Perspective on Multidisciplinary Science”, ECOSYSTEMS, Vol. 9, pp. 192-295, Springer Verlag, 1999.

[12] L. Ivanitskaya, D. Clark, G. Montgomery and R. Primeau, “Multidisciplinary Learning: Process and Outcomes”, Journal of Innovative Higher Education, Vol. 27, No. 2, Winter 2002.

[13] K. Lake, “Integrated Curriculum”, School Improvement Research Series, Closeup #16, 1994.

[14] S. Miyano, “The Asia Pecific Regional on bioinformatics”, IEEE Intelligent Systems, pp. 19-21, November/December 2001.

[15] D. Councell, “A Review of Bioinformatics Education in Uk”, Briefing in Bioinformatics, Vol. 4, No. 1, 7-21, March 2003.

AUTHORS

M. F. Wyne received BSc in Electrical Engineering, MSc in Engineering and PhD in Computer Science from University of Birmingham, UK. He is with the University of Michigan-Flint since 2002. Manuscript received March1, 2006.