The Knowledge Engineering Review, Vol. 00:0, 1–24. © 2004, Cambridge University Press
DOI: 10.1017/S000000000000000 Printed in the United Kingdom

Technical Forum Group on Agents in Bioinformatics[1]

Emanuela Merelli and Michael Luck

Dipartimento di Matematica e Informatica, Università di Camerino, Italy E-mail: emanuela.merelli@unicam.it
School of Electronics and Computer Science, University of Southampton, UK E-mail: mml@ecs.soton.ac.uk

Abstract

The scope of the Technical Forum Group (TFG) on Agents in Bioinformatics (BIOAGENTS) was to inspire collaboration between the agent and bioinformatics communities with the aim of creating an opportunity to propose a different (agent-based) approach to the development of computational frameworks both for data analysis in bioinformatics and for system modelling in computational biology.

During the day, the participants examined the future of research on agents in bioinformatics primarily through 12 invited talks selected to cover the most relevant topics. From the discussions, it became clear that there are many perspectives to the field, ranging from bio-conceptual languages for agent-based simulation, to the definition of bio-ontology-based declarative languages for use by information agents, and to the use of Grid agents, each of which requires further exploration. The interactions between participants encouraged the development of applications that describe a way of creating agent-based simulation models of biological systems, starting from an hypothesis and inferring new knowledge (or relations) by mining and analysing the huge amount of public biological data. In this report we summarise and reflect on the presentations and discussions.

1 Introduction

It is increasingly clear that significant improvements can be achieved in the bioinformatics field by designing and implementing new ICT tools that are able to distribute, at least partially, the computation burden, while reducing the need for the transfer of huge amounts of data. From this point of view, it is often felt that software agents can play a major role.

The scope of the Technical Forum Group (TFG) is to promote collaboration between the agent and bioinformatics communities, with the aim of creating synergies for modelling complex systems in the fields of bioinformatics and computational biology. As suggested by the AgentLink II Roadmap in Luck, 2003, one of the most promising and emerging application domains for agent technologies is the biological sciences, for which two different areas are identified:

- multi-agent systems for simulating and modelling biological systems; and
- multi-agent systems supporting the automation of information-gathering and information-inference processes.

The TFG meeting mainly focused on the following three areas:

- the process of modelling biological systems;

[1] This report is a personal view of the first AgentLink Technical Forum Group on Agents in Bioinformatics, informed by talks and postmeeting contributions from participants: Giuliano Armano, Andrea Doms, Nicola Cannata, Flavio Corradini, Mark d'Inverno, Phil Lord, Andrew Martin, Luciano Milanesi, Steffen Moeller, Terry Payne

- the process of performing data analysis; and
- agent-based systems, tools, and languages for bioinformatics.

It also aimed at receiving contributions from other European projects regarding agent technology and bioinformatics, to compare and transfer knowledge and results. Thus, the main purpose of the meeting was to bring together researchers working on agents, or bioinformatics, to discuss relevant issues and approaches related to using multi-agent systems in the fields of bioinformatics and computational biology.

1.1 Motivations

Bioinformatics and computational biology are emerging disciplines that use information technology to organise, analyse and distribute biological information in order to answer complex biological questions. In particular, Bioinformatics typically refers to activities that involve researching, developing, or applying computational tools and techniques aimed at dealing with biological, medical, behavioural or health data, including those to acquire, store, organise, archive, analyse, or visualise such data. On the other hand, computational biology refers to the development and use of analytical data and theoretical methods, mathematical modelling and simulation techniques aimed at studying biological, behavioural, and social systems. The amount of available information is constantly increasing, and it is difficult to exploit the available data from all sources. Many of the available data are interrelated, but it is currently difficult to identify, select, clean, or use all relevant data, as different tools use different data formats and with different semantics. There is a need to devise methods aimed at learning and discovering knowledge by *intelligently* combining these distributed data and information sources. Moreover, some classical problems might better be tackled by resorting to a suitable computational paradigm that uses various interaction protocols, such as cooperation or competition, to achieve an appropriate result.

1.2 Agents in Bioinformatics

Agent technology deals with entities typically equipped with information management and coordination capabilities. It is worth pointing out that an *act of communication* between two agents is feasible only if a suitable ontology exists, shared by both agents. This restriction guarantees agreement on the semantics of the exchanged data. Moreover, whenever an agent acquires additional information, it can integrate it with its personal knowledge base. Each agent is responsible for the consistency and the correctness of this operation. The notion of agents in bioinformatics suggests supporting the integration of information by designing domain-aware information agents for knowledge management and problem-solving within a biological domain. By contrast, the notion of agents in computational biology suggests designing agent-based systems, tools and languages for modelling the biological processes (pathways) themselves.

Agents may be useful for applications that imply: repetitive and time-consuming activities; knowledge management, such as integration of different knowledge sources; and modelling of complex, dynamic systems. All of these are typical in bioinformatics. In particular, the kinds of resources available in the bioinformatics domain, with numerous databases and analysis tools independently administered in geographically distinct locations, lend themselves almost ideally to the adoption of a multi-agent approach. Here, the environment is open and distributed with resources entering and leaving the system. There are large numbers of interactions between entities for various purposes, and the need for automation is substantial and pressing. Some early work in this direction, using agents for genome analysis, is demonstrated by the GeneWeaver project in the UK (Bryson et al., 2001), and work using DECAF in the US (Decker et al., 2001). Other work has considered agents more generally in bioinformatics, such as in the context of the UK's myGrid eScience project, developing a Bioinformatics Grid testbed. In Italy, too, preliminary results have been provided by the BioAgent project (Corradini et al., 2004), while for biological systems simulation, early work demonstrates the use of agent technology to model intracellular signalling pathways (Cellulat in Gonzalez, 2003, CellMAS in Corradini 2005), and for visual tools for cell modelling (CellAK in Webb and White, 2004).

2 The BIOAGENTS TFG Meeting

The BIOAGENTS talks were organized in four sessions: introduction to agent and Grid technologies; future challenges for computing technology in bioinformatics; recent experiences in using agents in bioinformatics; and case study proposals.

The first session aimed at introducing both agent and Grid technologies through experience developed in the AgentLink, myGrid and Grid.IT projects with presentations from Michael Luck, Terry Payne and Luciano Milanesi. In the second session, Nicola Cannata and Andrew Martin, two bioinformaticians, provided a stimulating exercise in defining several scenarios in which agent technology could be exploited, while Flavio Corradini, a computer scientist working on formal modelling of complex systems, spoke about his experience in using formal and semi-formal methods for specifying complex systems such as those in Bioinformatics. In session three, Andreas Doms, Phillip Lord, Steffen Moeller and Mark d'Inverno presented some results of ongoing projects in the broader field. Finally, Giuliano Armano and Nicola Cannata proposed two pieces of work with open issues.

AgentLink[2]

Michael Luck described AgentLink II, a network of excellence funded by the European Commission under its Information Society Technologies Fifth Framework Programme, which ran until mid-2003 to foster activity in the research and development of agent-based computing. As already mentioned in this report, bioinformatics and computational biology represent two of the most promising application domains.

myGrid[3]

Terry Payne introduced the myGrid project, which claims to provide a personalised environment for bioscientists, to help them to automate, repeat and therefore better achieve their experiments. myGrid aims to provide middleware for bio-eScientists to manage, investigate and analyse the increasing deluge of genomics data and to support convergence of data and literature archives. Furthermore, within myGrid, agent technology has been considered as one possible way to achieve personalisation and service discovery, automated delegation of tasks and responsibilities, handling and making decisions based on incoming notifications, and negotiation of behaviour.

Grid.it [4]

Luciano Milanesi reported on the Italian Grid.it, a project enabling platforms for high-performance computational grids oriented to scalable virtual organizations. Within the project, a special working group (WP12) is dedicated to grid applications for biology; one of the applications under study is the mapping of protein surfaces to functional determinants. A description of a protein site through a surface that models the shape conferred by the exposed residues is an effective tool for the analysis of proteins that may highlight similarities and relationships not detectable through comparisons of primary, secondary and tertiary structure. In the project, software has been developed to identify which amino acids subtend a certain surface; when a particular surface pattern is detected, we may be interested in checking from which amino acids it is formed. To that end, we can scan the protein surface to find out from which amino acids it is composed and, of these, which are important for protein function. The use of agents will help in searching the protein domain information, and in verifying how the these amino acids are actually arranged in the protein domain.

[2] www.agentlink.org
[3] www.mygrid.org.uk
[4] www.grid.it

Cellular processes modelling
As a future challenge, Nicola Cannata proposed modelling cellular processes by using agent technology. To this end, he is analysing the complexity of biological systems by showing the cell system and cellular processes. Modelling complex systems implies a deep understanding of the system both in terms of its structure and its behaviour Kitano, 2002. Once all components, their functions, their topological relationships, as well as parameters of each relation, have been identified, we need to analyse the system behaviour to understand the mechanisms that are behind the robustness and stability of the system, and functions of the interactions among components. Here, agent technology can be exploited to develop a suitable conceptual framework for simulation. The proposed exercise, described in detail during the talk, helped to analyse the main components of cell processes by identifying the main actors of the system, their roles, their functions, and their behaviour with a view to using agent technology.

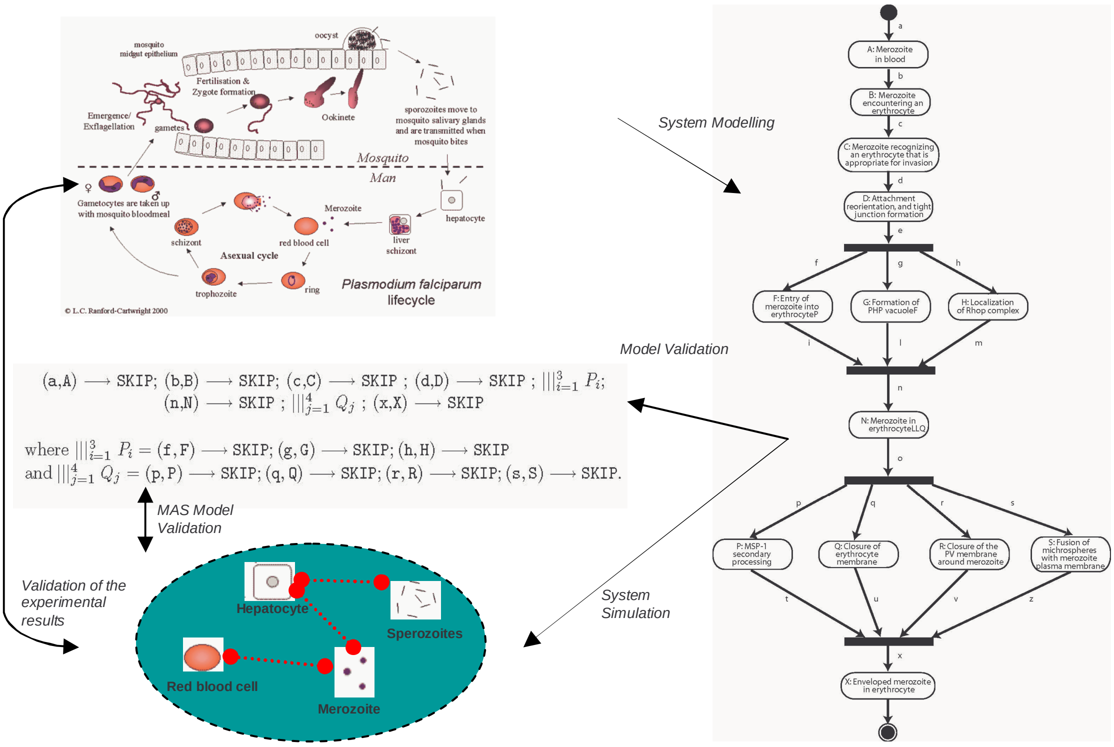
Analysis of Mutant Proteins: An Exercise in Motivation
Andrew Martin proposed considering the problem of analysing the effects of mutations on protein structure. He said that many diseases are caused by DNA mutations which lead to protein mutations: Cystic fibrosis, Favism (G6PD), Niemann Picks disease, OTC deficiency (urea cycle - hyper-ammonemia - brain damage), Cancer (p53, BRCA-1, APC, MYH, etc). Often biologists who study protein mutations attempt to analyse the protein structure, since the structure determines function. As an example, he posed the following questions: "Verify SNPs and confirm whether they are coding, leading to a protein mutation. If so, where is the mutation in the protein sequence and is a structure known? If it is, how does the mutation affect the structure?". The automation of such as workflow requires middleware suitable for supporting the specification, execution and coordination of very complex activities.

The use of information agents, in the context of the semantic web, could significantly help to retrieve and integrate meaningful information from heterogeneous and distributed data repositories.

Technology itself is not a problem. However, if something is too complex, or is perceived to be too complex, then why should biologists bother? They need to see a direct benefit: to have success, bioinformaticians and computer scientists must work closely and be driven by the needs of the biologist. The problem here is one of motivation — persuading the biologist who may have collected some interesting data and put it up on the web (e.g. one of the several hundred web sites listing mutations for a specific protein), to adopt standards and ontologies that can be used by agents and the semantic web.

Formal and Semi-Formal Methods for (Bioinformatics) Modelling
Referring to Nicola Cannata's talk, Flavio Corradini argued that the design of *incredibly* complex systems, needs suitable models, both to represent particular aspects of the biological system itself and to analyse the system from different viewpoints (e.g. static/structural, dynamic and functional). In fact, the introduction of models to describe a biological system helps an understanding of the biological system itself (by identifying the system structure, critical roles and responsibilities, functions and interactions, which are generally poorly identified). Of course, to create models we need languages and suitable notations for biological domains.

In the literature, a wide range of formal and semi-formal languages and notations can be found. These depend on the level considered, on the properties in which the designer is interested, and on the tools available to perform the analysis and verify properties. Proving properties in biological models can mean verifying properties related to the system/process behaviour (e.g. safety properties; liveness properties; simulations of system dynamics; checking for causal relationships, etc). Any property can be formally proved by using well known methods such as equivalence checking, model checking, simulation and model synthesis.


mosquito midgut epithelium
oocyst
Fertilisation & Zygote formation
Emergence/ Exflagellation
gametes
Ookinete
sporozoites move to mosquito salivary glands and are transmitted when mosquito bites
Mosquito
Man
Gametocytes are taken up with mosquito bloodmeal
♀
♂
schizont
Merozoite
red blood cell
liver schizont
hepatocyte
Asexual cycle
trophozoite
ring
Plasmodium falciparum lifecycle
© L.C. Ranford-Cartwright 2000
System Modelling
a
A: Merozoite in blood
b
B: Merozoite encountering an erythrocyte
c
C: Merozoite recognizing an erythrocyte that is appropriate for invasion
d
D: Attachment reorientation, and tight junction formation
e
f
g
h
F: Entry of merozoite into erythrocyteP
G: Formation of PHP vacuoleF
H: Localization of Rhop complex
i
l
m
Model Validation
n
N: Merozoite in erythrocyteLLQ
o
p
q
r
s
P: MSP-1 secondary processing
Q: Closure of erythrocyte membrane
R: Closure of the PV membrane around merozoite
S: Fusion of michrospheres with merozoite plasma membrane
t
u
v
z
x
X: Enveloped merozoite in erythrocyte
MAS Model Validation
Validation of the experimental results
System Simulation
Hepatocyte
Sperozoites
Red blood cell
Merozoite

$$\mathtt{(a,A)} \longrightarrow \mathtt{SKIP};\ \mathtt{(b,B)} \longrightarrow \mathtt{SKIP};\ \mathtt{(c,C)} \longrightarrow \mathtt{SKIP}\ ;\ \mathtt{(d,D)} \longrightarrow \mathtt{SKIP}\ ;\ |||_{i=1}^{3}\ P_i;$$
$$\mathtt{(n,N)} \longrightarrow \mathtt{SKIP}\ ;\ |||_{j=1}^{4}\ Q_j\ ;\ \mathtt{(x,X)} \longrightarrow \mathtt{SKIP}$$

where $|||_{i=1}^{3}\ P_i = (\mathtt{f},\mathtt{F}) \longrightarrow \mathtt{SKIP}; (\mathtt{g},\mathtt{G}) \longrightarrow \mathtt{SKIP}; (\mathtt{h},\mathtt{H}) \longrightarrow \mathtt{SKIP}$
and $|||_{j=1}^{4}\ Q_j = (\mathtt{p},\mathtt{P}) \longrightarrow \mathtt{SKIP}; (\mathtt{q},\mathtt{Q}) \longrightarrow \mathtt{SKIP}; (\mathtt{r},\mathtt{R}) \longrightarrow \mathtt{SKIP}; (\mathtt{s},\mathtt{S}) \longrightarrow \mathtt{SKIP}.$

Figure 1 Engineering Bioinformatics

Flavio proposed a scenario, that has been considered in previous work jointly made with Emanuela et. al. (Amici, 2004 and Corradini, 2004), where a semi-formal notation based on UML Activity Diagrams (see Figure 1a) is used to describe the activity workflow for a biological process: *the malaria parasite invading human host erythrocytes* (see Figure 1b). The resulting UML description, on the one hand, is translated to a formal notation (process algebra) (see Figure 1c) to verify suitable properties such as the function and structure of the resulting system (see Figure 1d) (Amici, 2004) and, on the other hand, is translated to a low level description (implementation) to simulate the biological process. In addition, the implementation part of the process itself makes use of agent-oriented technologies to support composition, amalgamation, dynamicity and mobility (Corradini, 2004).

Agents for the analysis of polygenic diseases

Steffen Moeller presented some ideas and concepts on using agents for the analysis of polygenic diseases and discussed preliminary results on combining RNA and protein expression levels, genotyping and intergenomics by using BioAgent, a programming environment based on mobile middleware (Merelli, 2002). He maintained that (disease-associated) genes and molecular pathways and the determination of consensuses of genetics with transcriptome/proteome analyses of human data with animal models of the disease are challenges of bioinformatics. Agent technology must support uniform access to local and public data (through a facilitator, i.e. a wrapper of web services or local tools e.g. EDITtoTrEMBL 1998).

In this view, agent technology must help in understanding the links between data sources and understanding the links to disease, providing reasoning over these data to yield a model of the disease in terms of the minimal number of genes/pathways that explain the maximal number of observations of the disease. This is done by gathering annotations of protein or genomic sequences (Gaasterland, 1996 and Bryson, 2000) and establishing consensus of information from protein domain databases and transmembrane protein sequence annotation.

Moeller suggests following the following protocol in using agents: submit only a single task, not hundreds of thousands; create and rank hypotheses; do not expect to find the absolute truth; bring agent technology closer to the interpretation of raw data generated in the wet-lab; and provide the results of agents to humans who do not know the truth either (implement heuristics).

In the present version, the proposed multi-agent system (Bioagent in Merelli, 2002) allows us to use: BioAgents for detection of promotor regions (based on First Exon Finder); Web services agents for SNP selection, conversion between genetic and physical distances, intergenomic consensus regions of disease association, localisation of genes on 2D gels and links to GeneOntology; LIMS for storage of expression and genotyping data; GNU R based analysis, also with BioConductor. All tools link to and from EnsEMBL. In the near future, agents could be used to suggest new wet-lab experiments to be performed, to address preferred investigation of particular regions of 2D gels (zoom gels, MS-identification of spots, search for predicted variants), to suggest investigation of genes that are not on a microarray chip and intelligently support the huge computational effort required, which could also benefit from load sharing technology in the context of Grid computing.

Towards A Semantic Web for Bioinformatics

Andreas Doms started his talk from the following consideration: in biology data grows superlinearly. Nowadays, DNA sequences in the human genome are equal to 3.2 Gbp (equivalent in size to 6 complete years of the New York Times); GenBank consists of more than 37 million sequences and more than 41 thousand million nucleotides; PubMed contains 14 million abstracts; SWISSPROT has 130,000 annotated protein sequences; TrEMBL has 850,000 protein sequences; and the PDB contains more than 25,000 protein structures. In consequence, we need powerful bioinformatics tools to support biologists in searching for meaningful information. Doms then proposed two tools: PROVA, a rule-based Java scripting language for the bionformatics semantic web; and GoPubMed, an ontology-based literature search and mining tool. PROVA can be considered a powerful tool if biologists can be persuaded in the way Andrew Martin aimed in his talk 2. In fact, PROVA supports the specification of workflows, by providing: rules for reasoning over data; rules for accessing data in flat files, databases, and other services; and rules for computations. On the other hand, by using the GO ontology, GoPubMed, allows one to retrieve and select meaningful information from a (generally) long list of abstracts obtained with a simple keyword search. GoPubMed submits a user's keywords to PubMed and retrieves the relevant articles. It extracts GO terms mentioned in the abstracts and from all the GO terms creates the induced ontology (the minimal subset of GO, which comprises all GO terms found in the documents) and displays it. The user browses the induced ontology to explore the PubMed results.

myGrid: Middleware for In Silico Biology

Phillip Lord briefly recalled that Bioinformatics analyses typically involve visiting many data resources and analytical tools. The resources are often highly heterogenous, semi-structured or un-structured, and distributed. This is largely because bioinformatics has grown up as a 'cottage industry' and is mostly web delivered. Integrating these resources is often difficult, both from a programmatic point of view, and also because of the heterogeneity. On the whole, this has been done by *screen scraping* and explicit Perl programming, which is brittle, and often done by non-expert programmers (making it worse). In addition, the data are heterogenous. Even things like identifiers are non standard.

Three key components developed in the myGrid project try to help this situation, as follows. Firstly, SOAPLAB provides a quick and easy way to publish legacy applications as web services.

This solves the problem of non-standard programming interfaces and removes the difficultly associated with screen scraping. Secondly, a workflow enactment engine enables the development of workflows thar are structurally simpler than a full programmatic environment, and also enables services to be strung together. Finally, to take advantage of these two items, we need an effective development environment to enable biologists themselves to develop the workflows and pipelines they need. From a single source of data, we query lots of different databases, lots of different resources. Effectively we are trying to find out as much as possible about the resource (in this case some DNA) as possible. Then we need to present all of these results back to the bioinformatician. Lord points out that semantic discovery of huge amounts of data and services is a very problematic issue, and that agents might help alleviate it.

Agent-based system for data analysis and simulation

Mark d'Inverno presented a project for modelling and simulation of stem cells between art and science (d'Inverno, 2004). It is an interdisciplinary project aiming to experimentally investigate new theories of stem cells. To the question of why a cell simulation should be developed, he gave many answers: 1) ethical, 2) it is difficult to identify cells in adult body, 3) even if you could, you would only ever see one possible behaviour, 4) mechanical forces can affect behaviour, 5) one need to kill cells to look at them, 6) images are heavily stained and magnified, and 7) looking at slides only in two dimensions. The role of simulation is to allow us to see things we cannot in the laboratory, to study the wholeness of a dynamic system, to examine the theoretical simplifications key to understanding fundamental properties, to develop insights into emergent/global phenomena, to suggest reasons for disease and medical experiments, and to run lots of experiments. In summary, the role of simulation is fundamentally to challenge current thinking (e.g. no such thing as stem cells).

The approach proposed is based on formal models. It is encompassing other theoretical approaches, with a strong link between the formal model and the simulation; it is a multi-agent approach (continuous, autonomous and intuitive), emergent in system behaviour, and seeking to provide an interdisciplinary visualisation of simulations. In particular, the agent approach suggests many questions: What can agents perceive? What is their state? What are their goals, strategies and intentions? Can they signal it? How can they interact? How can they communicate? And so on.

During the discussion it became clear that it is important to distinguish the role of the bioinformatician from that of computer science with respect to the biologist. The computer scientist, the bioinformatician and the biologist must form a team that works together. The computer scientist aims to create new models, methods and languages useful to solve complex problems for the biologist; but the biologist speaks a very difficult language to be understood by a computer scientist. The bioinformatician can be seen as an interpreter that helps a computer scientist to understand the computational problems behind biological systems, and to design new suitable computational models. After this step, the bioinformatician is able to develop new powerful tools for biologists.

Cases Study 1: protein secondary structure prediction by agents

Giuliano Armano proposed studying protein secondary structure prediction (SSP), a complex and difficult problem, by using a pool of agents. He introduced a multiple-expert architecture, where each expert embodies a genetic classifier (the guard) and a feed-forward artificial neural network (the embedded classifier), the former being devoted to controlling the activation of the latter. He would experiment with the use of agents both to design and implement solutions where strategies, mechanisms and policies must be highly reconfigurable. This would support the development of open systems, also able to integrate remote sources of information, or remote predictors.

Cases Study 2: BioLims as an agent-based virtual laboratory
Nicola Cannata proposed developing a virtual biological laboratory as an integration of management systems. The BioLIMS system lies in the concept of a virtual laboratory proposed as Cluster and Grid Computing for Solving Large Structural Biology Problems in the US. (Tutorial Marinescu 2002).

Just some of the activities in which an agent can assist a biologist include: replying an experiment performed earlier, planning a new experiment, controlling data processing (possibly remote experiments), evaluating the quality of partial results and getting advice when needed, engaging in collaborative efforts, and accessing the environment from a mobile device via the Web.

3 Future directions for agents in bioinformatics

It is clear that the combination of agents and bioinformatics presents a twofold opportunity. On the one hand, the domain of bioinformatics, with its extensive and growing resources of databases and analysis tools, provides an almost ideal domain for the application of agent technologies. It offers the possibility for deploying and testing agent systems in a real-world setting with the possibility of making substantial contributions to society. On the other hand, there is a distinct and identified need for good solutions to improve the performance of existing bioinformatics systems, and agents may be able to contribute to that improvement. In this sense, there is a very strong synergy between the two domains.

This picture is both enhanced and complicated by the introduction of relevant infrastructural technologies that facilitate both bioinformatics and agent-based computing. For example, the Grid has become increasingly important to both communities, and suggests a convergence to a service-oriented vision of bioinformatics underpinned by Grid-based virtual organisations.

However, there are still significant challenges. Researchers from both communities generally require education in the other, and work must be undertaken to ensure that any solutions across both areas satisfy both needs. In many cases, the language of discourse is so distinct that discussion of key issues becomes problematic. Additionally, the introduction of new technologies like the Grid requires further efforts, both in terms of understanding and adoption, and in terms of its immaturity in fully-deployed systems. Maturity at the interface is thus the key challenge. While many agent techniques may be used to address the concerns of the bioinformaticians, the lack of a complete understanding across domains suggests that it may still be too early to develop more sophisticated systems than the current generation of essentially management and mediation systems.

As identified in the AgentLink II roadmap, a potential longer-term application of multi-agent systems technologies is the use of agents engaged in reasoned argument to achieve resolution about ambiguous, or conflicting, experimental evidence, in a manner similar to the way in which human scientists do currently. This area of automated eScience is probably a decade or more from achievement, but will draw on the agent negotiation and argumentation mechanisms developed for distributed resource allocation problems, such as those found in eCommerce. This is still some distance away, but the TFG provides an opportunity to try to advance further down the road towards that goal.

References

R. Amici, D. Cacciagrano, F. Corradini, and E. Merelli. A process algebra view of coordination models with a case study in computational biology. In *Proceedings of First International Workshop on Coordination and Petri Nets, PNC'04*, 2004.

K. Bryson, M. Luck, M. Joy, and D. Jones. Applying agents to bioinformatics in Geneweaver. In *Cooperative Information Agents IV*, Lecture Notes in Artificial Intelligence, pages 60–71. Springer-Verlag, 2000.

F. Corradini, L. Mariani, and E. Merelli. An agent-based approach to tool integration. *Journal of Software Tools Technology Transfer*, 6:231–244, 2004.

F. Corradini, E. Merelli, and M. Vita. A multi-agent system for modelling the oxidation of carbohydrate cellular process. *Lecture Notes in Computer Science*, N. 3481, pages 1265–1273. Springer/Verlag, 2005.

M. d'Inverno and J. Prophet. Creative conflict in interdisciplinary collaboration: intepretation, scale and emergence. In Ernest Edmonds and Ross Gibson, editors, *Interaction: Systems, Theory and Practice*, pages 251–270. ACM, 2004.

T. Gaasterland and C. Sensen. Fully automated genome analysis that reflects user needs and preferences. a detailed introduction to the magpie system architecture. *Biochimie*, 78:302–310, 1996.

H Kitano. *Foundations of Systems Biology.* MIT Press, 2002.

M. Luck, P. McBurney, and C Preist. Agent technology: Enabling next generation computing (a roadmap for agent based computing). Technical report, AgentLink II, 2003.

E. Merelli, R. Culmone, and L Mariani. Bioagent: A mobile agent system for bioscientists. In *NETTAB — Agents in Bioinformatics*, 2002.

K. Webb and T White. Cell modeling using agent-based formalisms. In *Autonomous Agent and Multi-Agent Systems*, 2004.