

What Bioinformatics Applications Demand of Abduction and Induction

Chris Bryant

School of Computing
The Robert Gordon University
Aberdeen



Contents

- Part A
 - Some general requirements
- Part B
 - Some more specific ones pertaining to revising biological networks
- Conclusion
- Acknowledgements

Throughout this talk I will refer to

AI systems that Integrate Abduction and Induction
as

AI² systems

Caveat

An apology may be needed

- Danger that I may state what may be obvious
- Wish list rather a list of solutions
- I appreciate that what follows is easy for me to say but may not be anywhere near so easy for you to deliver
- But I think AIAI systems will need to meet these requirements if they are to be applied to bioinformatics

Part A

Some General Requirements

Remove Language Restrictions where Possible

We cannot afford to be too purist

- Particular types of inference place restrictions on the how data/knowledge is represented and what may be computed
- To be of use for Bioinformatics, AIAI systems should only impose such restrictions when and where it is necessary
- E.g. AIAI systems that use Prolog should, as far as possible, handle all of Prolog
 - Not just pure Prolog
 - But procedural aspects too

Noise Tolerance

It is difficult to avoid noise in bioinformatics

- Data in biological databases is often far from a gold standard
- A high proportion of results from an experimental method in biology can be incorrect
- To be of use for Bioinformatics, AIAI systems need to handle noise in:
 - Classes of training examples
 - Background knowledge
 - Existing scientific models may be incomplete or contain errors

Part B

More Specific Requirements Pertaining to Revising Biological Networks

Networks are Ubiquitous in Biology

- Several of types of networks in biology, e.g.:
 - Metabolic pathways
 - Signalling pathways
 - Regulatory pathways
- Significant numbers of each type of network
- Many need revising or extending
- Various sources of experimental biological data that indirectly suggest refinements
- What would this require of AIAI systems?

Graph Refinements

- Biological networks can be represented by graphs
- Logics such as first-order predicate logic are sufficiently expressive to represent a graph elegantly and concisely
- Need AIAI systems because of the need to revise or extend graphs based on indirect observations about them
- What kind of refinements?

Types of Refinements Required

- Find a missing arc label
- Find multiple missing arc labels
- Find a missing arc
- Construct a part of a network
- Construct an entire network

Conclusion

- General Requirements
 - Remove Language Restrictions where Possible
 - AIAI systems must be robust w.r.t. noise
- Specific Requirements for Revising Biological Networks
 - Need AIAI systems that can revise or extend bio-networks based on indirect observations about them

Acknowledgements

- People who worked on the Robot Scientist
 - S.H. Muggleton
 - S.G. Oliver
 - D.B. Kell
 - P. Reiser
 - R.D. King
- Funding for the Robot Scientist
 - EPSRC grant GR/M56067
 - BBSRC grant 2/E11552