AdaLab

- Adaptive Automated Scientific Laboratory (AdaLab)
- Adaptive Machines in Complex Environments
- Start Date: 1.4.15
Scientific Background
Computer systems capable of originating their own experiments, physically executing them, interpreting the results, and then repeating the cycle.
Scientific Goals

- We aim to integrate the scientific method with 21st century automation technology.
- We aim to make scientific discovery more efficient: cheaper, faster, better.
- Our vision is that within 10 years many scientific discoveries will be made by teams of human and robot scientists.
- This collaborations will produce scientific knowledge more efficiently than either could alone.
Scientific Goals

- We propose to develop a framework for semi-automated and automated knowledge discovery by teams of human and robot scientists.

- This framework will integrate and advance: knowledge representation, ontology engineering, semantic technologies, machine learning, bioinformatics, and automated experimentation.

- We will evaluate the AdaLab framework on an important real-world application in cell biology with biomedical relevance to cancer and ageing.
The Diauxic Shift

- Yeast (*S. cerevisiae*).
- First turn sugar into ethanol.
- Then turn ethanol into CO₂.
- Cancer
- Ageing
Eve

- Eve running
Partners
In London, near Heathrow airport.

Over 15,000 students and 1,000 academic staff from 113 different countries

Coordinates 11 FP7 EU projects, and a partner in >50 other EU projects

Secured >€18 M in European funding for the last two years

A wide range of scientific expertise, from robotic engineering to the Centre of Systems and Synthetic Biology
Role in Project: Knowledge representation

- The formal machine processable representations of the principle data and knowledge entities involved to the project, e.g. equipment, processes, participants, hypotheses, data, results.

- A declarative language for ML and probabilistic reasoning components.

- A knowledge base about the yeast diauxic shift.

- A communication mechanism between robot and human scientists.
University of Manchester

Ross D. King, Professor of Machine Intelligence, ross.king@manchester.ac.uk

- Manchester
- North of England
- Alan M. Turing
- Manchester ‘Baby’

Role in Project
- Head of Research
- Robot Scientists: Eve.
- Biological application
- Machine Learning
Partner KULeuven (Belgium)

- **PI:** Jan Ramon
- **Specificity:**
  - Data mining in graphs
  - Active learning
  - Probabilistic models
- **Related projects:**
  - MiGraNT (theory for data mining in networks)
  - InSPECTor (Proteomics experimental research)
- **Role in ChistERA AdaLab:**
  - Modeling uncertain knowledge
  - Hypothesis generation
  - Optimisation of experiment selection
  - Probabilistic inference
  - Algorithms for network data
Laboratoire d'Informatique de Paris-Nord: LIPN

- Celine Rouveirol

LIPN is associated with CNRS (UMR 7030). Research groups in *Combinatorics, Combinatorial Optimisation, Algorithmics, Logic, Software Engineering, Natural Language, Machine Learning*. The group involved in the project is: Machine Learning and applications.

Research in this team focuses on three main topics:
- Algebraic and logical models of learning,
- Collaborative and transfer learning,
- Learning structures from heterogeneous data.
Role in the Project

- Inductive Logic Programming: incremental theory revision, active learning, (deterministic) action model learning, learning from ambiguous relational examples.

- Complex systems analysis: community extraction, link prediction in multiplex networks.

- In collaboration with UPMC-LIP6, collective learning (multi-agent), distributed abduction.

- In collaboration with Evry University, inference of regulation networks from gene expression datasets.
AdaLab
AdaLab

- Adaptive Automated Scientific Laboratory (AdaLab)
- Adaptive Machines in Complex Environments
- Start Date: 1.4.15
Scientific Background
The Concept of a Robot Scientist

Computer systems capable of originating their own experiments, physically executing them, interpreting the results, and then repeating the cycle.

Background Knowledge → Hypothesis Formation → Experiment selection → Robot → Results Interpretation → Analysis → Final Theory
Scientific Goals

- We aim to integrate the scientific method with 21st century automation technology.
- We aim to make scientific discovery more efficient: cheaper, faster, better.
- Our vision is that within 10 years many scientific discoveries will be made by teams of human and robot scientists.
- This collaborations will produce scientific knowledge more efficiently than either could alone.
Scientific Goals

- We propose to develop a framework for semi-automated and automated knowledge discovery by teams of human and robot scientists.

- This framework will integrate and advance: knowledge representation, ontology engineering, semantic technologies, machine learning, bioinformatics, and automated experimentation.

- We will evaluate the AdaLab framework on an important real-world application in cell biology with biomedical relevance to cancer and ageing.
The Diauxic Shift

- Yeast (*S. cerevisiae*).
- First turn sugar into ethanol.
- Then turn ethanol into CO$_2$.
- Cancer
- Ageing
Partners
Brunel University: Coordinator

- In London, near Heathrow airport.
- Over 15,000 students and 1,000 academic staff from 113 different countries.
- Coordinates 11 FP7 EU projects, and a partner in >50 other EU projects.
- Secured >€18 M in European funding for the last two years.
- A wide range of scientific expertise, from robotic engineering to the Centre of Systems and Synthetic Biology.
Role in Project: Knowledge representation

- The formal machine processable representations of the principle data and knowledge entities involved to the project, e.g. equipment, processes, participants, hypotheses, data, results.
- A declarative language for ML and probabilistic reasoning components.
- A knowledge base about the yeast diauxic shift.
- A communication mechanism between robot and human scientists.
<table>
<thead>
<tr>
<th>Manchester</th>
<th>Role in Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>North of England</td>
<td>Head of Research</td>
</tr>
<tr>
<td>Alan M. Turing</td>
<td>Robot Scientists: Eve.</td>
</tr>
<tr>
<td>Manchester ‘Baby’</td>
<td>Biological application</td>
</tr>
<tr>
<td></td>
<td>Machine Learning</td>
</tr>
</tbody>
</table>
Partner KULeuven (Belgium)

- PI: Jan Ramon
- Specificity:
  - Data mining in graphs
  - Active learning
  - Probabilistic models
- Related projects:
  - MiGraNT (theory for data mining in networks)
  - InSPECTor (Proteomics experimental research)
- Role in ChistERA AdaLab:
  - Modeling uncertain knowledge
  - Hypothesis generation
  - Optimisation of experiment selection
  - Probabilistic inference
  - Algorithms for network data
Celine Rouveirol

LIPN is associated with CNRS (UMR 7030). Research groups in Combinatorics, Combinatorial Optimisation, Algorithmics, Logic, Software Engineering, Natural Language, Machine Learning. The group involved in the project is: Machine Learning and applications.

Research in this team focuses on three main topics:
- Algebraic and logical models of learning,
- Collaborative and transfer learning,
- Learning structures from heterogeneous data.
Role in the Project

- Inductive Logic Programming: incremental theory revision, active learning, (deterministic) action model learning, learning from ambiguous relational examples.

- Complex systems analysis: community extraction, link prediction in multiplex networks

- In collaboration with UPMC-LIP6, collective learning (multi-agent), distributed abduction.

- In collaboration with Evry University, inference of regulation networks from gene expression datasets.
The Institute of Systems and Synthetic Biology is a research unit of University of Evry and CNRS. The iSSB is located on the Genopole campus, the leading BioPark in France, near Paris.

Research areas: machine learning, computational and systems biology, bioinformatics.

AdaLab:
- Reconstruction of context-specific molecular networks about yeast diauxic shift
- Integration of molecular network data, to reconstruct active ontology and selecting experiments
Key challenges and potential impact of the project
The proposed AdaLab needs to be:

- *autonomous and perceptive to human requirements* (its scientific collaborators).
- *able to continuously learn, adapt and improve in the “real world” complex environment* of scientific research.
- capable of continuous cycles of scientific hypothesis formation and experimentation that will improve its scientific knowledge (models).
Key Challenges 2

- Integrating a systems approach, with the research involving collaboration between experts in: robotics, machine learning, logical and probabilistic inference, semantic technologies, and yeast microbiology.

- Integrating high-level reasoning about scientific knowledge with the control of low-level robotic movements to execute experiments.

- Develop a protocol for communication between human and robot scientists.
Key Challenges 3

n Scientific knowledge is inherently uncertain. Therefore within the AdaLab framework we need to develop Bayesian methods that make *inferences* and plan experiments *under uncertainty*.

n Scientific knowledge is best represented using *logic*. To *integrate logic with probabilities* we will use statistical relational learning, and develop an ontology for *representing uncertain knowledge*.

n The success of the AdaLab framework will be objectively determined by quantitative measurements of the different system components, and the scientific knowledge generated.
Key Outputs

- An AdaLab demonstrated to be greater than 20% more efficient at discovering scientific knowledge (within a limited scientific domain) than human scientists alone.
- A novel ontology for modelling uncertain knowledge.
- An efficient communication mechanism between human and robot scientists.
- New machine learning methods for the generation and efficient testing of complex scientific.
- Novel biomedical knowledge about cell biology relevant to cancer and ageing.
Potential Impact

- Science is the greatest generator of economic wealth (through developments in technology).
- Science is the greatest driver of better health (through development in biomedical science).
- The AdaLab framework will contribute to realising Europe’s 2020 strategy for smart, sustainable and inclusive growth.
Potential Impact

- Intelligent laboratories have the potential to speed up the technological progress.
- Such an increase would lead to more scientific discoveries, better technological solutions, and new products.
- For example new better drugs could be delivered to the market faster and cheaper. Currently, ~25Billion € is spent annually within the EU on pharmaceutical research. Most of this is spent on late-stage trials (which are less amenable to automation), but we conservatively estimate that ~10% is amenable to the AdaLab framework.
Planning
Work packages

- WP1: Management
- WP2: Knowledge Representation
- WP3: Automated Scientific Discovery
- WP4: Machine Learning and Inference
- WP5: Dissemination
Work plan

- Five interlinked WPs.

- The proposed AdaLab framework will be developed in three cyclic iterations.

- The main software components will be incrementally released and updated.

- All projects outputs will be made available to the research community by the end of the project.
The End of the Beginning of AdaLab