COMPOSING LEARNING FOR ARTIFICIAL COGNITIVE SYSTEMS (CompLACS)
Main theme of the project

- Development of components for various learning problems, that can be composed together.
- The goal is a well-founded suite of technologies that enable the construction of complex cognitive systems.
Components and learning problems

- Representation of perception and experience
- Trading off exploration and exploitation
- Autonomous exploration and skill acquisition
- Reinforcement learning
- Control and multi-component systems
“Design Scenarios”

- How can components be combined and put to use?
  - Assumption: Well chosen decomposition facilitates complex design.

- Taxonomy of components leads to a decomposition methodology for complex systems.
  - Requirements for the components
Design scenarios and components are tested on three application scenarios:

- Robot arm with cameras
- Quadrocoopers
- Web content analyser

Simulators will be distributed to groups working on components, and used as platforms to implement learning algorithms.
Application: Robot arm

- Test-bed for scenarios of motor control
  - Complex motor skills require a variety of movements chosen reactively, and motor planning and coordination on several timescales.
  - An important area of innovation is in learning to produce such behaviour by composing biologically plausible motor primitives to provide a wide range of reactive movement.
Application: Quadrocopter

- Adapted from an existing commercial radio-controlled quadrocopter
- Test-bed for scenarios with multi-agent control (swarms of quadrocopters)
Application: Web Content Analyser

- Existing live system reads ~30K news articles per day (23 languages), classifies them with dozens of SVMs, and maintains info on ~10K named entities.
- Multiple independent learning modules communicate by tagging and summarising news articles, and provide mutual training feedbacks.
- Tractable test-bed for examining stability and effectiveness of multiple interacting learning modules.
WP1: Perceptual and Experiential Representations

- Key observation: Learning grounded in experience traces can avoid hand-crafted representation spaces.

- Passive learning of experiential representations

- Active and attentional control of experiential systems
Multi-armed bandits are the simplest model involving a trade-off between exploration and exploitation.

Algorithmic and theoretical analysis have shown how to achieve near optimal performance.

Remarkably effective in tackling complex AI problems: e.g. MOGO resulted in a quantum jump in artificial GO playing performance.
WP2: Multi-armed bandits and extensions

- Use bandit algorithms to perform planning in MDPs or POMDPs.
- Extend bandit algorithms to hierarchical structures, such as trees.
- Bandit problems with many dependent arms (e.g. GP bandits).
WP3: Towards real-world reinforcement learning

How can future robots *learn* really complex tasks?

- Complex tasks generically consist of a small set continuous basic behaviors.
- These behaviors need to be learned efficiently both using demonstrations and trial & error.
- Task context needs to employ these behaviors in various situations.
WP5: Control and learning of multi-component intelligent systems

- Decompose the control problem for multi-component systems.
  - Efficient approximate control methods capitalizing on the link between stochastic optimal control and probabilistic inference.
  - Superposition principle in path integral control theory: simple skills can be combined to generate complex movements.
  - Control architectures that include reasoning among agents.
WP4: Self motivated and autonomous exploration and skill acquisition

- Autonomous exploration is necessary when supervision and explicit feedback is insufficient.

- Objective: Investigate together
  - autonomous exploration,
  - the learning of multiple skills,
  - the construction of higher-level representations.
Autonomous exploration

WHAT IF THERE’S NO AFTERLIFE? SUPPOSE THIS IS ALL WE GET.

OH, WHAT THE HECK. I’LL TAKE IT ANYWAY.

YEAH, BUT IF I’M NOT GOING TO BE ETERNALLY REWARDED FOR MY BEHAVIOR, I’D SURE LIKE TO KNOW NOW.
Project details

- 4 years project, started in March 2011.
- 8 academic partners:
  - UCL: coordination, design scenarios, representations (John Shawe-Taylor, Yee Whye Teh, Steve Hailes)
  - Bristol: application scenarios (Nello Cristianini)
  - Lille: bandits (Remi Munos)
  - Darmstadt: reinforcement learning (Jan Peters)
  - Nijmegen: control theory (Bert Kappen)
  - Berlin: control theory (Manfred Opper)
  - Royal Holloway: skill acquisition (Chris Watkins)
  - Leoben: autonomous exploration (Peter Auer)