Real-World Optimisation with Life-Long Learning

EPSRC grant EP/J021628/1

1 Background

Many practical problems arising in industrial domains concerned with operating sustainably, meeting demand and minimising costs cannot be solved exactly. Meta-heuristic optimisation techniques have been widely developed in academia to solve such problems with much success reported in the literature. However, there remains a worrying void between scientific research into optimisation techniques and those problems faced by end-users and addressed by commercial optimisation software vendors. From a commercial perspective, the problems addressed by academia are too simplistic compared to those faced in the real-world, failing to embrace many real-world constraints [?]. For example, even 'rich-vehicle routing' research fails to address concerns of customers regarding non-trivial loading problems, working-time legislation, rush-hour speed modelling and asymmetric distance networks. From the scientific perspective, [?] have also identified a "lack of advanced metaheuristic techniques in commercial software" which has been attributed in part to the academic community failing to demonstrate that their solutions are applicable to the needs of the commercial world, and in part to academics failing to impart their message the industrial community.

Meta-heuristics approaches can further be costly to develop as they generally require human expertise to integrate specialist knowledge into an algorithm, and expertise in heuristic methods to design and tune algorithms. Recent research has therefore focused on automated algorithm design and configuration [?] which produce tuned solvers that perform well on either individual problems or across suites of problems. One branch of this field is hyper-heuristics (HH), which operate on a space of low-level heuristics, searching for combinations of heuristics which exploit the strength and compensate for the weakness of individual known heuristics. The resulting algorithms are cheap to implement, require less human expertise, have robust performance within a problem class, and are portable across problem domains. These features compensate for some reduction in solution quality compared to tailor-made approaches, while still ensuring solutions of acceptable quality. However, most automated design approaches fail to incorporate or recognise a crucial human competence; human beings *continuously* learn from experience; by generalising observations and feedback, they are able to update their internal problem-solving models in order to continuously improve them, and adapt to changing circumstances. The failure of computational solver to exploit previous knowledge both wastes useful knowledge and potentially hinders the discovery of good solutions. Furthermore, if the characteristics of instances of problems in the domain change over time, solvers may need to be completely re-tuned or in the worst case redesigned periodically.

This proposal addresses these dual concerns raised above. We propose a novel *lifelong-learning* hyper-heuristic system ($LLL - HH$) which addresses current deficiencies inherent in current systems: it will exhibit short-term learning, producing fast and effective solutions to individual problems and at the same time, long-term learning processess will enable the system to autonomously adapt to new problem characteristics over time. It therefore exploits existing knowledge whilst simultaneously adapting to new information. Secondly, by working closely with two collaborators, a commerical routing software vendor and a Forestry expert, our research will be directly informed by real-world problems, accounting for real constraints and performance criteria, thereby producing economic impact. Advances in optimisation techniques will be driven by the development of a publically available suite of problems which reflect real-world priorities and constraints, derived from actual problem data provided through our collaborators.

2 Research Hypothesis and Objectives

Hypothesis: *A hyper-heuristic framework which is able to continuously learn over time as it gathers and records experience learned from solving problems will be more* **efficient** *and* **effective** *at solving a range of practical problems than current optimisation techniques and will address current commerical concerns regarding the applicability of academic optimisation techniques to real-world problems.* The proposal has the following aims:

- to improve the current state of the art in generic problem-solvers by developing a novel system which exploits previous knowedge, continuously learns from experience and autonomously adapts to shifting problem characteristics
- to demonstrate that the proposed system is more *efficient* and *effective* at producing high-quality solutions to real-world practical problems than previous optimisation approaches in terms of reducing costs and environmental impact

- to develop an an *information database* of problem-solving knowledge as a platform for advancing the development of optimisation techniques which are informed by real-world problems constraints and commercial priorities
- to demonstrate to end-users and commercial software vendors that meta-heuristic optimisation techniques are applicable to the needs of the commerical world in terms of encapsulating real-world constraints and being cheap to both implement and maintain.

The proposal is timely in that a climate of economic austerity it becomes increasingly necessary for industry to gain an competitive advantage by reducing costs and being able to rapidly adapt to the environment in which it operates. Systems which can be developed at reduced cost and with minimal need to expert input will further provide advantage. Increasingly, industry must also consider the environmental impact of their activities; optimisation systems which are able to incorporate environmental objectives alongside traditional objectives such as time and cost will be particularly advantageous [?].

The hyper-heuristic philosophy is often summarised as 'heuristics to choose heuristics' in that a search method is utilised to search a space consisting of a set of low-level problem specific heuristics, each of which makes an can make an incremental improvement to a solution. A domain barrier separates a HH *algorithm* from *domain-specific* heuristics, thereby ensuring the approach is generally applicable. In a recent classification [?], two classes of HH methods are identified: *online learning* techniques which learn to solve or improve a single instance of a problem and *offline* learning techniques which learn rules from a set of training instances which then generalise to new instances. Online HH strategies typically learn a function which rates the quality of a set of heuristics in terms of improving solution quality given the current problem state, repeatedly applying heuristics until the problem is solved: examples of techniques to learn the rating function include reinforcement learning, tabu search and other meta-heuristic techniques [?]. In offline learning, classifier systems and case-based reasoning have been used to learn rules from a set of training instances, which are thereafter applied to new problem instances [?]. Both online and learning approaches have proved successful in the past on specific classes of problems but both have limitations. Online methods start from a clean slate each time an instance is presented. With each problem, the HH method must learn how to rate the heuristics for this particular problem. Previously learned knowledge gleaned from solving previous problem instances is never exploited, potentially reducing algorithm efficiency. Offline techniques on the other hand require a large and representative set of training examples for use in a one-off training phase. Ensuring that the problem space is covered in the training set presents difficulties, and furthermore, if the nature of the problem space changes over time, the system has to be regularly re-trained.

HH systems should be able to capture learned information regarding the utility of a heuristic in a given situation, and re-utilise and/or adapt their knowledge over time as the nature of the problem space shifts. As in existing rule-based systems, knowledge must also be able to rapidly retrieved from the system to produce acceptable solutions to problems. We propose a novel $LLL-HH$ framework guided by previous research in the domain of mobile robotics e.g [?] . This body of research exploits mechanisms apparent in the natural immune system (IS), inspired by the observation that the immune system can be viewed as cognitive system with recognition and learning abilities, coupled with a capacity to interact with an open and changing environment. Three essential qualities are apparent in the IS; it performs efficient search; it has a large adaptive capacity; it has an endogenous selective memory. In robotics, the metaphor has been exploited to build systems which satisfy goals in dynamic environments by both evolving novel behaviours and arbitrating between sets of competing behaviours to select the most appropriate to solve an immediate task. The learning mechanism operates over two distinct timescales; rapid adaptation leads to good solutions to immediate tasks; a long-term learning process adapts the underlying system based on experience and environmental changes. We propose that similar mechanisms can enable the creation of a novel and effective $LLL-HH$framework.

Briefly, the immune system produces *antibodies* which destroy pathogenic material. A short-term learning procedure rapidly adapts antibodies produced by the bone marrow in response to pathogenic stimulus to efficiently clear infection. According to Jerne's idiotypic network theory [?], antibodies further self-organise into networks in which connections between *antibodies* further stimulate or suppress antibodies. This self-regulatory behaviour eliminates redundant antibodies while reinforcing useful ones, simultaneously supporting both memory and decision making. *Memory* is encapulated in the topological structure of the network which maintains useful antibodies. *Decision making* is induced by competitition between antibodies which emerges in a 'winning' antibody which dominates the response. Crucially, the antibody network is plastic in terms of the connections, the connection strengths and the constituent antibodies themselves, which supports learning and adaptation across multiple time-scales [?].

In the context of previously successful *robotic* research e.g. [?], an antibody represents a state-action pair which describes a *action* to be taken given a particular environmental *state*. Connections between antibodies in the idiotypic network represent *preference relationships* which signify whether an antibody's action takes priority over, or concedes to, another action. Network interactions modify the concentrations of those antibodies whose state condition matches the environment; the antibody that emerges with the highest concentration executes its action. *Dynamic* processes (such as cloning and mutation) acting on the network enable short-term learning by performing a search around existing antibodies which refines immediate decision making. *Meta-dynamic* processes provide long-term adaptation by reinforcing network connections, recruiting new nodes or pruning redundant nodes from the network. In the context of HH, we propose that a similar approach could result in a network which arbitrates between *problem-specific heuristics*, given the current *state* of a solution to a problem instance. Meta-dynamic processes would adapt the topology of the network through feedback from solving multiple problem instances thus adapting to underlying changes in the domain, while dynamic processes would enable short-term learning ensuring rapid production of solutions.

In order to judge a hyper-heuristic system successful, it must be shown to guarantee a minimum level of quality in solutions produced, guarantee a maximum time-period in which solutions can be produced, and furthermore, generalise across problem domains. Although results from existing HH algorithms are often reported on large benchmark sets of data available in the literature, this does not necessarily inspire confidence outwith academia as noted by [**?**]; [**?**] further note with respect to vehicle routing that taxonomies used in academic literature merely encourage short-sightedness with respect to the complexity of real-world problems, failing to capture practical constraints. Furthermore, solutions are often evaluated according to criteria which are not relevant to practitioners [**?**]. Therefore, we will develop problem generators for a number of problem domains which generate realistic problems, informed by input from two collaborators: two external partners, Optrak, an optimiation software vendor, and Forest Research (Forestry Commission) will support this endeavour, providing example problems and domain expertise in the scheduling, routing and forest management domains. We will ensure scientific integrity by adhering to the principles laid down in [**?**] for generating good test-suites. Specifying the characteristics of a problem is not sufficient — a problem description will be coupled with a description of the appropriate metrics for evaluating the problem, to enable more rigorous comparison of methods between researchers, and more closely emulate the real world in that the evaluation criteria are chosen by the end-user rather than the researcher — for example, including environmental metrics as well as more traditional cost-based ones. Benchmark and supplied problems will be analysed to extract features and characteristics, and used to abstract templates for generation of new real-world problems. To address the question of quality, we will implement a scientifically rigorous testing framework to evaluate the proposed approach, informed by the latest research in experimental design and analysis of stochastic algorithms, e.g [**?**, **?**]. We will evaluate the solutions according to customer-centric criteria, which may for example include quality, robustness, worst-case performance etc. Additionally, we will use ranking methods to compare the quality of solutions generated to the individual heuristics, to state-of-the-art HH approaches, and further to standard meta-heuristic optimisation techniques. The success of the project in achieving the project aims will be achieved through the following measurable objectives:

1. Develop a HH framework informed by existing work in robotics using immune-inspired network algorithms
2. Develop a set of test-suites in a range of problem domains consisting of problem generators informed by characteristics of real-world problems and constraints
3. Identify appropriate metrics to measure the efficiency and effectivenes of the HH system in producing solutions in close collaboration with end-users
4. Utilise the metrics to evaluate the efficiency and effectiveness of the proposed system in a number of problem domains including vehicle-routing, forestry management and bin-packing
5. Conduct a sensitivity analysis of the system and stress-test the approaches to determine the operational boundaries within which the approach is acceptable to end-users
6. Compare the new system to existing HH approaches and to solutions achieved from state-of-the-art optimisation approaches on individual problem instances to establish the extent of any trade-off between performance on individual problems and across problem classes
7. Evaluate the *generality* of the approach by undertaking a cross-domain analysis of results.