

A Framework for Scaling Up Distributed Minds

Oisín Mac Fhearaí, Mark Humphrys and Ray Walshe

School of Computing, Dublin City University, Dublin, Ireland
{omacfh, humphrys, Ray.Walshe}@computing.dcu.ie

Abstract An important topic in artificial intelligence is the modular decomposition of problems and solutions into smaller ones. Modular, “hybrid” solutions tend to be built from components mostly created by the same programmer or the same team. What if we already have many solutions online - programs created by a number of different authors and hosted as services - and wish to compose them into a larger, distributed hybrid program which performs better than the individual components? Can this be done automatically?

The World-Wide Mind (W2M) project attempts to scale up artificial intelligence by distributing action-selecting agents (which we call “minds”) and problem environments (which we call “worlds”) as services on the internet, and by allowing minds to call other minds and thus facilitate building hybrid minds from many programs which may have been written by many authors.

This paper gives a general overview of the W2M architecture and ongoing work examining the possibility of automatically constructing hybrid minds.

The World-Wide Mind: Architecture and Implementation

To facilitate the creation of minds and worlds, and especially of hybrid minds which consult other minds when making decisions, we introduced a uniform interface which services must follow, representing messages that may be sent to minds (*getaction*) and worlds (*getstate* and *takeaction*). This, coupled with the ability to upload a mind to a server and have it immediately appear as a service online, makes experimentation and composition of minds simpler.

Although we consider the W2M platform to be useful for teaching and exploration of problems, our hope is that hybrid minds will be created which query other subminds for suggested actions (which may themselves be hybrids), and thus large-scale hierarchical problem-solving programs can be built from the work of many authors who may not understand how the other subcomponents function.

Since each level in the hybrid mind's hierarchy may have many branches, it is easy to see that computational demands can rise quickly. To cope with this problem, it must be possible to distribute minds across machines and networks.

In the first implementation of the W2M server, worlds and minds were embodied as web services and assigned a URL. These services were hosted using the

Tomcat application server, so that messages between services consisted of a web request and response, with the message content represented as XML. While this enabled connectivity across the internet, running a mind - especially a hybrid mind composed of many remote minds - in a world was slow due to overheads in the underlying servlet technology and the use of HTTP to wrap messages.

To avoid this bottleneck, the server now sends the XML messages over a simple TCP protocol with very little overhead. This allowed us to take advantage of the common case where mind/world services are located on the same machine, by avoiding the network stack completely. Distribution will still be needed to scale up to bigger hybrids, but where network access can be avoided, it should be.

To further reduce latency, a scheme was implemented whereby the user asks the mind to carry out a run with the world and receives an asynchronous stream of messages from the mind with the states seen and actions taken, as well as a score object representing how well the mind performed in this run.

Evaluation and Future Work

The system was used by undergraduate computer science students taking an A.I. module. A W2M server was used to host several hundred minds for the Tyrrell animal world. Minds were ranked by their performance in this world, which prioritises mating and survival in a simulated environment.

A requirement was added that every student must submit at least one hybrid mind which delegates to one or more subminds. A call graph feature was implemented to track calls between minds, and the scoreboard at the end of the assignment showed that nine of the top ten minds called at least one other mind.

Ongoing research looks at the selection of subminds when constructing a hybrid mind. The set of minds described above was used to perform an analysis of the world state and score data. A large number of runs were performed with the submitted minds, and a record kept of all states seen and actions taken.

Metrics were created corresponding to important subgoals, for example minimising total thirst or hunger. These metrics were used to rank minds by their performance on each goal. A hybrid mind was created which consults the best mind for each subgoal in a series of simple case-action tests, for example: "if a mate is nearby, return the action suggested by the best mating mind".

Experiments were carried out for hybrids composed of 2, 3 and 4 subminds. These tests produced a hybrid which scored 10% better than any of the submitted minds (mating 82 times versus 74 in the best existing mind).

Future work will explore ways to automatically determine which metrics most strongly influence the score, using statistical correlation rather than human intuition.

Acknowledgments This work was supported by IRCSET's Embark Initiative.