

Foraging in a Grid World Using Action Templates

R. A. Chaminda Ranasinghe^{1*}, A. P. Madurapperuma², N. D. Kodikara³

^{1,3} University of Colombo School of Computing, Colombo, Sri Lanka

² University of Moratuwa - Faculty of IT, Sri Lanka

chaminda.ranasinghe@dialog.lk, ajith@itfac.mrt.ac.lk, ndk@ucsc.cmb.ac.lk

Revised: 30 September 2008; Accepted: 24 September 2008

Abstract: The emergence of behavioural and structural congruence based on simple local interactions of atomic units is a fascination to the scientific community across many disciplines. The climax of behavioural congruence and emergence of behaviour is exemplified by the community life-style of ants. Each individual ant possesses the capability only to solve part of the overall puzzle while aggressively communicating in primitive methods with the spatially related neighbours to produce emergent behaviour.

The primary hypothesis of this research is that the constituent atomic actions of a complex behaviour could be successfully coordinated by a collection of collaborative and autonomous agents with the use of Action Templates.

The AAANTS (Adaptive Autonomous Agent colony interactions with Network Transparent Services) model was conceptualised and implemented as a platform to represent the biologically inspired coordination and learning model to test the research hypothesis. The domain of foraging in a grid-world was identified as the experimental basis to evaluate the AAANTS coordination model. The experiments demonstrated relative improvements in achieving behavioural congruence using the AAANTS model in relation to the traditional Monte-Carlo based methods.

Keywords: Collective Intelligence, Action Templates, Emergent Behaviour, Reinforcement Learning, Frame Representation.

INTRODUCTION

The survival of an entity in the environment is directly attributed to selecting the most appropriate and refined behaviour with respect to the rapid changes in the environment. Behaviour of this nature could be called as congruent with reference to the current demands of the environment. However, over a period of time due to the changes and demands of the environment, the existing behaviour could become incongruent or obsolete. Hence, behaviour should adapt and improve, or simply be congruent to the latest changes in the environment.

The adaptive entities in the natural world use emergent models to achieve behavioural congruence. These models begin with an innate layer of basic incongruent atomic behaviour, which based on the reinforcements and or supervisions from the environment reaches a level of refinement more aligned with the demands of the environment. Hence, dynamically and

stochastically combining atomic behaviour that are either accepted or rejected based on the reinforcements from the environment tends to provide a high level of behavioural congruence in natural systems.

The success of the naturally occurring models in delivering abundance of heterogeneous and congruent behaviour using concepts of emergence, innateness and adaptations has inspired this research. The primary hypothesis of this research is that the constituent atomic actions of a complex behaviour could be successfully coordinated by a collection of collaborative and autonomous agents with the use of Action Templates. The domain of grid-world foraging was selected to implement the use of Action Templates. The Action Templates were executed as a coordinated effort of a collection of autonomous Software Agents.

The AAANTS model could be applied to several application domains based on the generic nature of the concept. The simulations and experiments discussed in this paper were based on the domain of grid-world navigation. However, this model could also be applied to the domains of pattern recognition, robotic movement and vision navigation. The experimental results related to robotic movement and vision navigation were excluded from the scope of this publication.

The subsequent sections of the paper discuss the conceptualisation, realisation and experimentation of the AAANTS model within the domain of foraging in a grid-world. Sections 2 and 3 clearly state the objectives and inspirations that contributed to the motivation of the research. Section 4 describes the AAANTS coordination model that consists of Atomic Actions, Action Templates, Behavioural Centres and Sensory Templates. Section 5 contains a discussion of the design, implementation and execution of the grid-world experiment. The conclusion of the research with respect to the defined objectives of the research is discussed in the last section.

MOTIVATION

The age-old ambition of creating intelligence on artificial substance that is anthropomorphic in nature is still considered a dream, yet to be realised by humans. It was this curiosity that initiated the investigation into the behavioural complexity found in nature, which subsequently became the foundation of this research.

There are several theories, models and paradigms that have given inspiration and direction to the work

* corresponding author

carried out in this dissertation. Naturally occurring collective systems of individually simple animals such as populations of insects and turtles together with artificial phenomena such as traffic jams suggest that individual complexity is not a necessity for complex intelligent behaviour of colonies of such entities [1], [2].

The community life-style of ants was an inspiration to this research. It was estimated that the Ants' success story spans over several millions of years preceding the known era of human existence [3]. Each individual ant possesses the capability of only solve a part of the overall puzzle while aggressively communicating in primitive methods with the spatially related neighbours to produce emergent behaviour. Ant colonies have evolved means of performing collective tasks, which are far beyond the capacities of their individual structures. This phenomenon is demonstrated without being hard-wired together in any specific architectural pattern and without central control [1], hence void of top-down control. The consensus is that comprehension of emergent complexity in insect colonies such as ants would serve as a good foundation for the study of emergent, collective behaviour in more advanced social organisms, as well as leading to new practical methods in distributed computation [4], [5]. Therefore, the key motivation was to devise an artificial learning model, that could demonstrate collective intelligence analogous to insects.

The "Society of Mind" theory by Marvin Minsky [6], was another inspiration to this research. This theory portrays the mind as a collection of mindless components that interact and compete to provide intelligent emergent behaviour. Society of agents in the mind is triggered by external sensations where agents act individually but in a cooperative and synchronised manner.

The incarnation of a complete multi-cellular being starting from a single fertilised egg seems like a heavenly secret to all of us and is certainly a motivation to this research. It is the initial set of genes in a fertilised egg that helps a simple cellular growth to be morphed into a complex combination of organs found in a complete animal. It is amazing that every cell contains a complete footprint of all genes found in the initial cell and each cell only represents a single instance of the overall pattern. This aspect of different cells expressing same genes at different levels could be called as a sub-pattern where most patterns are in fact combinations of a small number of basic patterns [7]. Hence, a gene could be compared to a conductor leading an orchestra; the conductor makes no music on its own but with the proper participants could produce a symphony of enormous beauty and complexity [8].

RESEARCH OBJECTIVES

Congruent behaviour could be achieved through several methods. However, persistence of congruent behaviour in relation to the dynamics of the environment, and further the sustenance of congruence over a considerable period of time is still considered non-trivial based on the current

artificial models of intelligence. This research tends to take a step in the direction of sustaining behavioural congruence using coordination methods from nature based on emergence.

The primary objective of the research is to evaluate whether the bottom-up emergent methodologies could provide similar or improved results in comparison to the methodologies that prescribe behaviour composition in a top-down manner to achieve behavioural congruence in dynamic environments. There are several aspects to be focused to realise this objective. The rest of the discussion primarily focuses on building a unique coordination model to realise the stated objective within the domain of grid-world Foraging.

AAANTS COORDINATION MODEL BASED ON PATTERNS AND EMERGENCE

The "AAANTS Coordination Model" was conceptualised based on the inspiration from the natural emergent systems. The model encompasses aspects such as identifying sensory patterns, relationship among actions and sensations and team formation among agents for coordination. The interactions among agents act as perturbations and the system achieves congruence with the use of reinforcements. The resulting model consists of heuristics and algorithms that could be used to implement an agent system that demonstrates emergent behaviour. Similar work on sensory-motor coordination and identifying sensory patterns is found on research by Rolf Pfeifer et al and Stefano Nolfi et al [16], [17].

Creating Behavioural Concentres with Atomic Actions and Action Templates

The term Atomic Action (AA) could be defined as an action that cannot be further subdivided into elementary actions. For example, in humans, the contraction of a homogeneous muscle could be identified of as an AA. A given AA could produce different effects based on the intensity and the degree of temporal progress. If base duration of an atomic action a is defined as t , $a.t$ represents the elementary temporal result of executing action a . However, changes in the temporal dimension of executing the same atomic action a , would produce different end results – e.g. $[a.2t]$, $[a.3t]$, etc. Within the boundary of this research, AAs are considered innate

and could only be manipulated within the dimensions of time and intensity.

The Action Templates

The concept of the Action Template (AT) is introduced herewith as the primary method of grouping AAs to define behaviour. A template could be defined as a generalisation of related instances that determines or serves as a pattern. Further, the concept of templates is used analogously to represent the concept of a class in object-oriented programming and design

methodologies. A template could be also considered as a description of an aspect of a task. In-line with these definitions, a definite list of AAs executed in concurrency and or sequence in relation to environmental sensations could be called as an AT.

An AT would not be of any use without being instantiated. Hence, ATs are instantiated by the agents, thereafter refined for a specific task. A given AT could be modified to accomplish different tasks by a group of agents. Concurrency is a basic fact of nature for achieving complex behaviour. The survival in the environment demands concurrent threads of attention to both sensations and actuations. It should be noted that due to the need for concurrency, the AAs within a single AT could be contributed by several agents.

The methodology used by agents to collectively execute synchronised tasks without the knowledge of the overall outcome was given special emphasis during the conceptualisation stage of this research. According to Keith Deckeretal [9], the coordination problem of choosing and temporally ordering actions is more complex because the agent may only have an incomplete view of the entire task structure of which its actions are a part, the task structure may change dynamically and the agent may be uncertain about the outcomes of its actions.

The type and sequence of AAs and their synchronisation with sensations for initiation and termination uniquely differentiates ATs from each other. Hence, in summary, three aspects are important to an AT: types of AAs, maximum temporal exposure of each AA and the influence of sensations (environmental sensations and the temporal progress of other AAs within the same AT could also be served as a sensation) for the purpose of initiation and termination effects of each AA.

Figure 1, depicts an action template defined using several AAs. With reference to this diagram, each AA is constrained with a start and a finish (e.g. actions a1, a2, a3, a4 have respective start and finish defined as {s1, e1}, {s2, e2}, {s3, e3}, {s4, e4}). In addition, for each started AA instance, a timer is created for

measuring the temporal progress of the action. A started action could finish due to lapse of allocated maximum execution time or due to a trigger from a sensation. The maximum allocated time of each AA would be defined during the creation of the AT. Further, the initiation of actions would be triggered from temporal progress of other dependent actions within the same template and or sensory stimulation from the environment.

An important aspect of the AT concept is in the methodology used for action synchronisation. An AT should be first instantiated to facilitate the defined behaviour. Subsequent to the initial instantiation, the first action in the sequence would be activated. However, there could be situations where several AAs that belong to an AT are activated simultaneously at the initiation based on the stochastic nature of the action selection mechanism. An ongoing action would publish the temporal progress within the respective domain, and other participants (agents) could use this information for coordinated participation. Therefore, both the temporal progress of the other actions and the sensory information from the environment is used for action coordination. The coordination sequence is improved over a period of time due to the reinforcements received after executing an instance of a template.

The AAs could be described as innate to an intelligent entity. However, the ATs could be formed both in terms of innateness and adaptations. The innate ATs would be ready to use, though with further fine-tuning through environmental supervisions and or reinforcements. The adaptive ATs would be created through a stochastic process where innate AAs are randomly selected to form novel behavioural structures. Further, ATs would be able to form hierarchical or lateral bonds with each other, again through a stochastic process to create complex behavioural outcomes. The AAANTS model conceptualises both flavours of ATs, but the grid-world experiment only focussed on the innate ATs that are refined through reinforcements.

A similar approach is taken in leaning systems like ALECSYS [10], where the learning “brain” of an

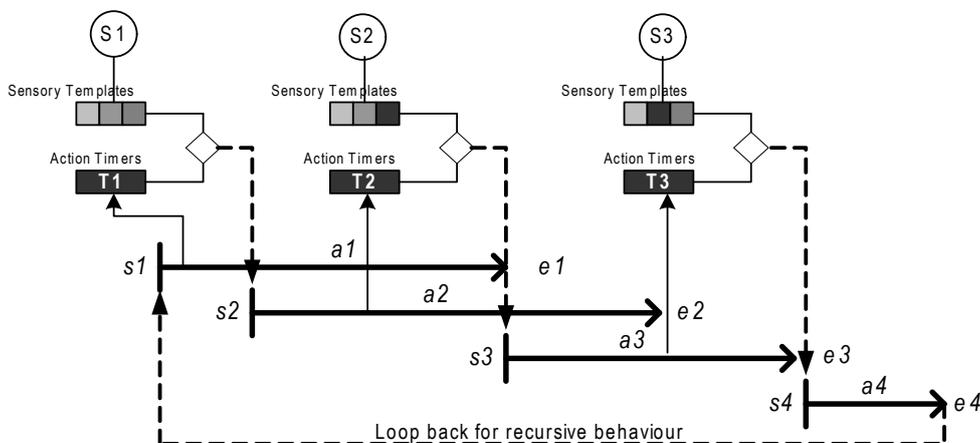


Figure 1: Action template with a defined sequence of atomic actions

agent could be designed as the composition of many learning behavioural modules. The modules are called as basic behavioural modules which are connected to sensory and motor routines that learn from external stimuli. The behavioural modules of ALECSYS could be made analogous in concept to ATs discussed above. Simply, AAs are like the bricks and templates are like different wall types of a building, where different combinations of walls could be used to create buildings of diverse architectural complexities.

The concept of the AT would also be similar in some extent to behavioural assemblages [11]. According to Tucker Balch [11], groups of behaviours are referred to as behavioural assemblages. One way that behavioural assemblages may be used in solving complex tasks is to develop an assemblage for each sub-task and to execute the assemblages in an appropriate sequence. The resulting task-solving strategy could be represented as a Finite State Automaton (FSA) and the technique is referred to as temporal sequencing.

Behavioural Concentres

The groups of actions in an AT that consist of AAs are the basis for building complex behaviour. A sequence of AAs (depicted in figure 1) that are executed in a coordinated manner is referred to as a Behavioural Act (BA). The concept of a BA is similar to the definition found in myrmecology for a collection of elementary actuations [12]. For example, in Figure 1, actions a1, a2, a3 and a4 represent a BA. Further, a collection of closely linked BAs could be defined as a Role where a Task could be differentiated as a similar sequence of BAs that are coordinated.

A popular method of depicting a behavioural repertory is by the use of an ethogram, which incorporates repertory of a caste, transition probabilities of acts and the time distributions spent on each act [12]. The Figure 2 represents an ethogram that depicts the roles within a group of entities and the states and actions that facilitate the transitions. It should be noted that some actions (actions a5, a11 & a17 in the ethogram – Figure 2) enable a role to be navigated to states of another role.

Roles could also be described in terms of cohesion and coupling of ATs. There exists high cohesion among the AAs that belong to an AT. Further, one or many ATs are required to define a given role. The ATs that belong to a specific role should have higher coupling within them than with others external to the role.

The Action Breakdown Structure (ABS) of the AAANTS coordination model would be a good approach to explain the rest of the behavioural complexity of the AAANTS coordination model. The ABS conceptualised within the AAANTS model is represented in Figure 3. The structure is segmented into two primary layers of functionality based on the innateness and adaptability. The actuation layer represents the raw AAs that are innate in nature and less complex. As examples, the basic contraction of muscles, release of enzymes and hormones, change of chemical composition in animals

are analogies to these types of actions. Hence, AAs are the building blocks of any complex behaviour.

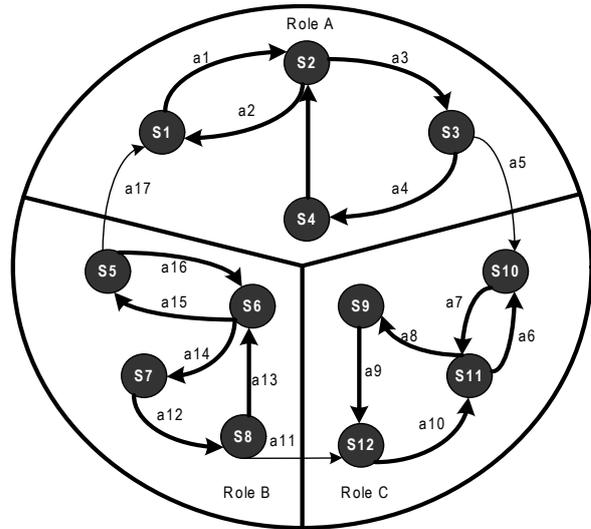


Figure 2: Ethogram representing transitions of behavioural acts based on different roles

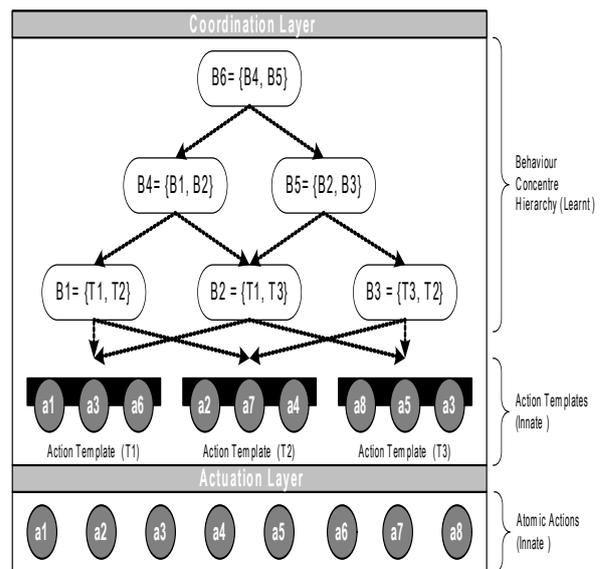


Figure 3: Conceptual action breakdown structure of the AAANTS coordination model.

The ATs represented within the coordination layer in Figure 3 are responsible for grouping AAs into elementary chunks of coordinated behaviour. However, these templates would be useless without being coordinated with other ATs to perform more complicated roles. The AAANTS model introduces the concept of Behavioural Concentres (BC) [13] as the enabler for coordination among the ATs. The centres are created, adjusted and destroyed based on the reinforcements from the environment.

It is assumed that the innate repertoire of AAs should suffice the expected behaviour of an individual. However, absence of a particular behaviour in an individual does not imply that relevant AAs are

missing. Many of us possess the atomic actuations in the upper limbs to become an artist, though few of us are capable of such coordinated behaviour. Further, many of us have the innate AAs to play a violin, though few of us could. Therefore, the BCs and ATs are important in harnessing the capabilities of AAs. In most in-born talents such as art, music and athletics are mostly due to the inherited ATs. Hence, the assumption is that some types of special innate ATs are required to full-fill some higher-level complex behaviour. However, even with inherited ATs, without proper environmental adaptations to build up BCs could be called as a “waste of talent” by most of us.

SIMULATIONS AND EXPERIMENTS

The primary experiment was to develop an environment to simulate foraging activities of insects. The food collecting behaviour of insects called as foraging is a popular domain of experimentation among the researchers of collective intelligence [12]. Further, the experiments related to a grid-world, where agents are supposed to transit through states with the objective of finding the optimum path in reaching a defined goal have been popular among the artificial intelligence community for years [19], [20]. The original grid-world problem was enhanced to include foraging related aspects to the simulation. Key control variables and their configurations for different experiments are listed in Table 1.

There are many flavours of reinforcement learning methods such as Monte Carlo (MC), Dynamic Programming (DP) and Temporal Difference (TD) [14]. Each of these methods has advantages and disadvantages based on the domain of application. It is considered that MC methods scale better with respect to state space size than standard, iterative techniques for solving systems of linear equations [15]. Further, an MC method does not require explicit knowledge of the transition matrix of the problem domain [15]. Hence, MC method was selected as the reinforcement learning algorithm for the experiments of this research due to the above stated uniqueness and also due to the similarity in concept to other similar reinforcement learning methods. Therefore, the fundamental learning algorithm of the AAANTS learning model was based on the MC method.

In all the experiments, the exploration probability was kept constant. The initial exploration probability was kept at 0.99, which thereafter was linearly reduced after each episode. The reduction rate of exploration probability, hence, was kept at a constant across all the experiments. Further, a uniform reward distribution strategy was adhered across all experiments except in the grid-world experiment 1 scenario 1. The reward distribution was performed episodically while keeping state values to ascend from home to destination, hence encouraging the agents to follow a path of ascending state values similar to the effect of pheromones in ants.

The Simulated Grid-World Environment

A grid-world is an area with a restricted boundary as depicted in Figure 4. At a given instance there could be one or many participants within the grid that may perform state transitions either to reach the destination Food Source (FS) which is the goal state or else to return to the nest with the already captured food elements after reaching the goal state. Each participating agent is analogous to an ant in a colony.

A grid-world could be experimented along several dimensions such as spatial, temporal and functional. In terms of spatial aspects, the total grid is divided into small squares called as cells. Most of the discussed experiments are based on a 10 x 5 grid, but the same experiments were performed on 20 x 30 and 30 x 40 grid environments to assess the scalability. The movements within the grid are done on temporal clock cycles and the main functions of agents are searching and transporting food. The grid and obstacle layouts are totally configurable using the grid-world simulator front-end application.

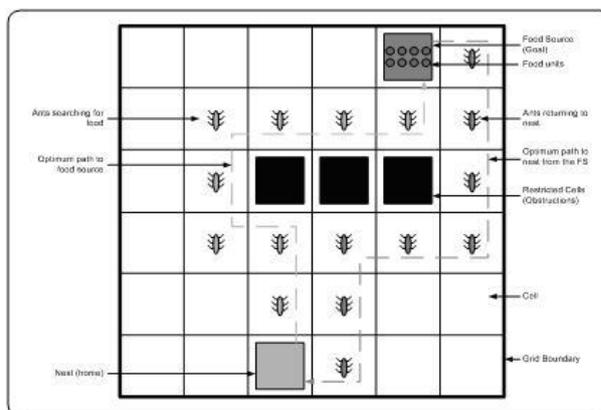


Figure 4: Grid-world model for the ant foraging simulation

The participants could travel from one cell to another in a horizontal or vertical direction, but restricted in travel diagonally. A single participant could inhabit a cell at a time during the search stage, though several may travel together while transporting a food unit collectively. However, there could be some cells that are obstructed and impassable by the agent to make the foraging task more realistic. A detailed discussion of the design of this experiment would be found in the Ph.D. dissertation of R.A.C Ranasinghe [18].

Several experiments were conducted using the grid-world simulation to evaluate the AAANTS coordination model.

- Grid-world experiment 1 – Single agent foraging.
 1. Scenario 1: One Step Look-Ahead Policy using Monte Carlo (MC) Method with Proportionate Reward Distribution.
 2. Scenario 2: Disproportionate Reward Distribution among the Participating States

Table 1: Control variable summary across all grid-world experiments

Variable Descripti	Ex 1 - Sc 1	Ex 1 - Sc 2	Ex 2 - Sc 1	Ex 2 - Sc 2	Ex 3 - Sc 1	Ex 3 - Sc 2
Grid Size	10 x 5	10 x 5	10 x 5	10 x 5	10 x 5	10 x 5
Obstacle arrangement	Constant	Constant	Constant	Constant	Constant	Constant
Characterises of agents	Constant	Constant	Constant	Constant	Constant	Constant
Learning algorithm	MC	MC	MC	MC	AAANTS	AAANTS
Number of agents	1	1	2	4	4	4
Number of search threads	1	1	2	4	1	1
Reward distribution	Equal	Dispropor- tionate	Dispropor- tionate	Dispropor- tionate	Dispropor- tionate	Dispropor- tionate
Look-ahead	1 Step	1 Step	1 Step	1 Step	2 Step	2 Step
Shared memory context	No	No	Yes	Yes	Yes	Yes
Implicit communication	No	No	Yes	Yes	Yes	Yes
Use of action templates	No	No	No	No	Yes	Yes
Knowledge Representation	Individual	Individual	Shared	Shared	Shared	Shared
Initial state initialisation	Random	Random	Random	Random	Random	Random
Exploration probability and rate of reduction	Constant	Constant	Constant	Constant	Constant	Constant

- Grid Experiment 2 – Cooperative Foraging Using Monte-Carlo Method.
 1. Scenario 1: 2-Agent Cooperative
 2. Scenario 2: 4-Agent Cooperative
- Grid Experiment 3 – Collective Foraging Based on the AAANTS Model
 1. Scenario 1: 4-Agents Using an Action Template of 2 Actions with 6 ATs
 2. Scenario 2: 4-Agents Using an Action Template of 2 Actions with 8 ATs

Observations of the Grid-World Experiments

The following observations of the grid-world experiments were captured as important to assess the hypothesis and objectives of this research.

1. When comparing results of experiment 1, scenarios 1 & 2, it is evident that disproportionate distribution of rewards among state values results in better convergence to the optimum path (Figure 5). The disproportionate distribution is analogous to pheromone distribution of insects where the concentration is maintained in an ascending rate when reaching the goal state. Even after changing the location of the goals and obstacles in scenario 2, the algorithm was able to readjust the state values to converge to the new path within a reasonable number of episodes.
2. The objective of the experiment 2 is to evaluate the effectiveness of implicit coordination methods using shared contexts on general learning algorithms such as Monte-Carlo. Both scenarios of experiment 2

demonstrated improvements when compared to the results of experiment 1, in which, the latter is void of any form of coordination. However, several more experiments were carried out with increased agent counts from one to ten. It was noticed that initial gradual improvements fade away after reaching an optimum threshold of agents, which was however, variable based on the grid sizes.

3. Among all the Monte-Carlo based experiments (experiments 1 & 2), the 4-agent cooperative method produced the best outcome (Figure 5). This was a modification done to the original Monte-Carlo method to include the cooperative aspects with the objective that it could be compared in similar grounds with the AAANTS model.
4. Experiment 3, introduces the full scale features of the AAANTS model. It introduces the capabilities of emergence, innateness and implicit communication. In experiment 3, a key difference when compared to experiments 1 & 2, is that though there are multiple agents, there exists only one search thread at a time. The multiple agents coordinate different elementary actions of the AT to navigate a single search node from source to destination. An AT is executed based on inputs from the environment and each elementary action is contributed by a single agent.
5. The results of the experiment 3 out perform that of experiments 1 & 2, and further demonstrate that capabilities improve when the innate layer contribute several ATs to survive in the environment. Most suitable AT needs to be selected based on the sensations from the environment.

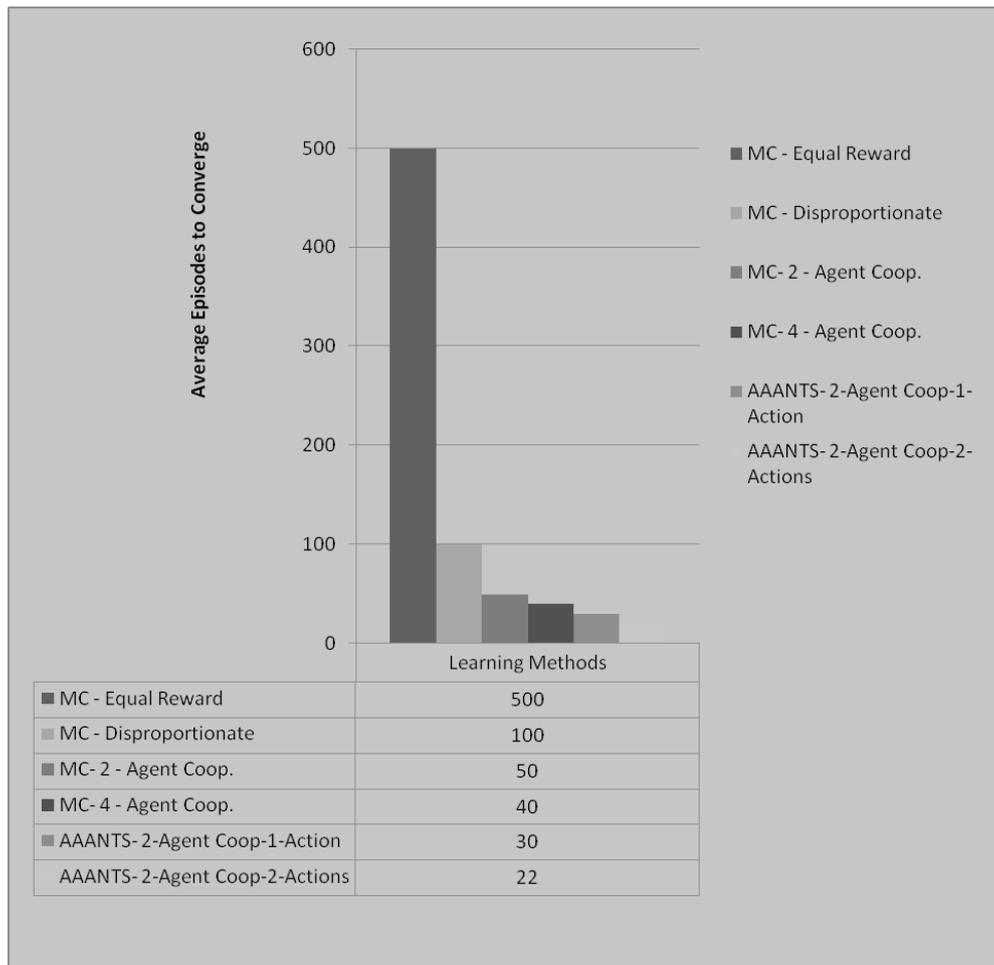
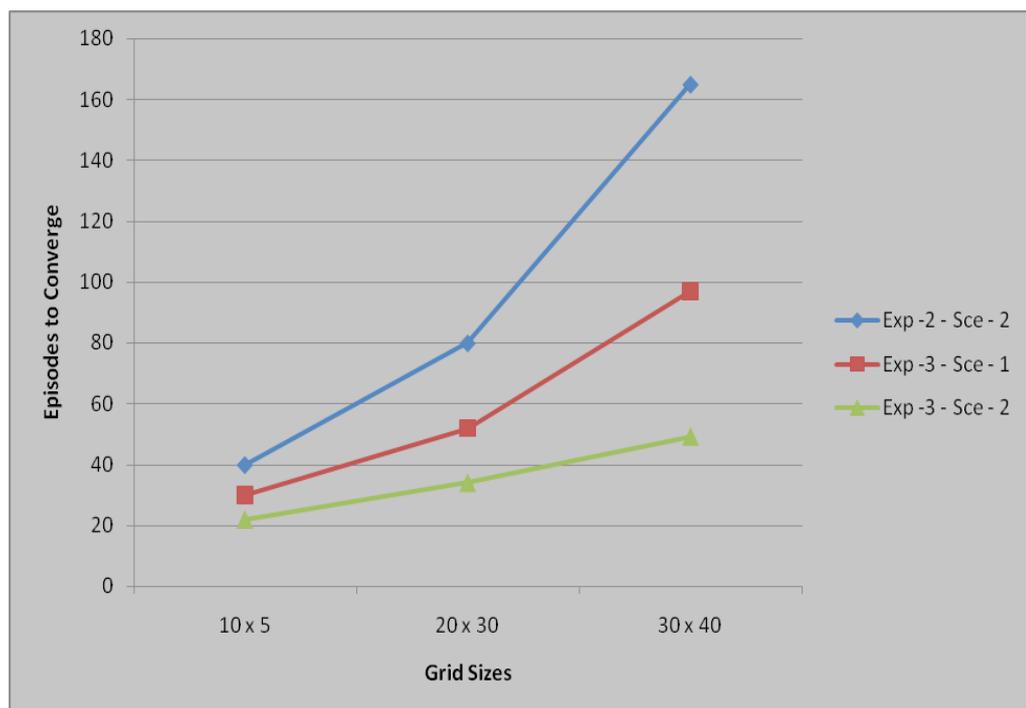


Figure 5: Comparison of average episodes taken to converge to the optimum path using different learning strategies

6. Further, it was noticed that when the amount of obstacles were increased within the grid-world, the AAANTS method converges considerably faster (within less number of episodes) than the Monte-Carlo methods. This was due to the fact that AAANTS uses obstacle characteristics as navigation markers during the initial exploration process. These obstacles were described as local-optima and specifically within the AAANTS model referred to as Hubs – special states that bridges regions of cells. For example, when there is a pattern of receiving high reward for moving forward when a certain type of obstacle is in the neighbourhood, the agents detects these situations as Hubs and adapts to executing the appropriate AT whenever such situations were faced.
7. The summary of the experimental outcomes of all the experiments of the grid-world domain is tabulated in Table 2. It could be stated that the number of episodes to converge and states to reach the goal state considerably reduces in the AAANTS domain. The final outcome is very stable in the AAANTS model when compared to the rest of the control experiments.
8. Figure 6, depicts the results of experiments conducted on extended search spaces of 20 x 30 and 30 x 40 grid sizes. The experiment 2-4 agent scenario was taken to represent the MC learning method, which is actually the best performing out of all the MC experiments. The MC method does show convergence to an optimal path, however, the overall number of episodes increases considerably when compared to the AAANTS learning model. Both experiments related to the AAANTS learning model show superiority in comparison to MC method. Out of the two AAANTS model based experiments, the method which contained higher number of ATs seems to converge with a relative lower number of episodes and further, the ratio of increase is lower. It could be concluded that the AAANTS learning model scales better in complex environments when compared to the MC method. The experiments conducted on the same grid sizes with increased number of obstacles demonstrated even better results in favour of the AAANTS model in comparison to the MC method.

Table 2: Observation summary of the grid - world experments

Observation/ Experiments	Ext 1 - Sc 1	Ext 1 - Sc 2	Ext 2 - Sc 1	Ext 2 - Sc 2	Ext 3 - Sc 1	Ext 3 - Sc 2
Average number of states of the optimum path from source to destination	15	10	9	9	8	8
Presence of local optima	Yes	Yes (relatively-low)	Yes	Yes	Yes	Yes
Stability after converging to the optimal path	No	Mostly	Mostly	Mostly	Yes	Yes
Ability to reach the optimal path	No	Mostly	Mostly	Mostly	Yes	Yes
Minimum number of episodes to converge to optimum path	> 100	50 - 100	30 - 50	27 - 40	25 - 30	20 - 22
Ability to converge after adjusting the location of the goal subsequent to reaching convergence	No	Mostly	Mostly	Mostly	Yes	Yes
Ability to converge after adjusting obstacle arrangement subsequent to reaching convergence	No	Mostly	Mostly	Mostly	Yes	Yes

**Figure 6:** Comparison of overall average episodes to converge in extended grid search spaces

CONCLUSION

The essence of emergence is that any of the contributors to the emergent behaviour is not aware of the master plan. The grid-world experiments 1 and 2 is void of any form of emergence, however, shows gradual improvements (within the 4 scenarios of the experiments 1 and 2) related to the use of shared contexts and implicit communication among the participants. However,

the grid-world experiment 3, focuses on the emergent nature of behaviour with the introduction of the full capabilities of the AAANTS model. The AAANTS model demonstrates considerable improvement over the standard Monte Carlo technique and specially performs exceptionally better in larger grid sizes. Further, it is concluded that dynamic changes in the environments (goal and obstacle location changes) are gracefully handled by the AAANTS model in comparison to the Monte-Carlo learning model. These observations

confirm the achievement of congruent behaviour in dynamic environments using the concept of the AAANTS model.

The grid-world experiments confirm that the behavioural acts built based on innate action templates provide better convergence to the optimum behaviour than using a pure adaptation strategy void of innate behaviour, which thereby confirm the respective objective set forth in the introduction. The purely adaptive experiments especially in the grid-world simulation, demonstrates that the tests conducted void of Action Templates takes relatively more episodes to converge to the optimum path and further, intermittently settle down on local optima.

References

1. Parunak V, Sauter J. and Clark S. (1997). Toward the Specification and Design of Industrial Synthetic Ecosystems, *Fourth International Workshop on Agent Theories, Architectures, and Languages (ATAL'97)*
2. Resnick M. (1994). Turtles, Termites, and Traffic Jams, *Explorations in Massively Parallel Micro-worlds*, ISBN: 0-262-18162-2
3. Holldobler B and Wilson E.O. (1994), Journey to the Ants: a story of scientific exploration, ISBN: 0-674-48526-2
4. Babaoglu O., Meling H., Montresor A., (2001). Anthill: A Framework for the development agent-based peer-to-peer systems, University of Bologna, Italy, Norwegian University of Science and Technology, Norway
5. García C.G., (2001). Artificial Societies of Intelligent agents
6. Minsky M., (1986), The Society of Mind, *A Touchstone book*, Simon and Schuster Publishers, New York, ISBN 0-671-65713-5
7. Salazar-Ciudad I., Garcia-Fernandez J., and Sole R. V., (2000). Gene networks capable of pattern formation: From induction to reaction-diffusion, Complex Systems Research Group, Department of Physics, FEN-UPC Campus Nord, Barcelona Spain
8. Elman J.L., Bates E.A., Johnson M. H., Karmiloff-Smith A., Parisi D., Plunkett K. (1999). Rethinking Innateness, A connectionist Perspective on Development, ISBN: 0-262-05052-8
9. Decker K. and Lesser V., (1995). Designing a family of Coordination Algorithms, Department of Computer Science, University of Massachusetts, Amherst
10. Colombetti M., Dorigo M., (1993). Training agents to perform sequential behaviour
11. Balch T., (1997). Learning roles: Behavioural Diversity in Robot Teams, Mobile Robot Laboratory, College of Computing, Georgia Institute of technology, Atlanta Georgia
12. Holldobler B. and Wilson E.O., (1990). The Ants, The Belknap Press of Harvard University Press Cambridge, Massachusetts. ISBN: 0-674-04075-9
13. Ranasinghe R.A.C., Madurapperuma A.P. (2005). Learning Coordinated Actions by Recognising State Patterns with Hubs, *Eighth International Conference on Human and Computers, University of Aizu, Japan*
14. Sutton R.S. and Barto A.G., (1998). Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, A Bradford Book
15. Barto A., Duff M., (1994). Monte Carlo Matrix Inversion and Reinforcement Learning, Computer Science Department, University of Massachusetts, USA
16. Pfeifer R., Scheier C, (1999). Understanding Intelligence, ISBN-0-262-66125-X
17. Nolfi S., Parisi D., (1999). Exploiting the power of sensory-motor coordination
18. Ranasinghe R. A. C. (2008). Emergence of Congruent Behaviour by Implicit Coordination of Innate and Adaptive Layers of Software Agents, *Ph.D. Dissertation, University of Colombo, Sri Lanka*
19. Arai S., Sycara K., (2000). Multi-agent reinforcement learning and conflict resolution in a dynamic domain, The Robotics Institute, Carnegie Mellon University, USA
20. Abbeel P., Andrew Y. Ng, (2004). Apprenticeship Learning via Inverse Reinforcement Learning, Department of Computer Science, Stanford University, Stanford, USA