Where there is no law, but every man does what is right
in his own eyes, there is the least of real liberty.

– Henry M. Robert III

University of Alberta

Collective Decision-Making in Decentralized Multiple-Robot Systems: A Biologically Inspired Approach to Making Up All of Your Minds

by

Christopher A. C. Parker

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Computing Science

Fall 2009
Edmonton, Alberta

Examining Committee

Hong Zhang

Renee Elio

Michael Bowling

Thomas Hillen

Chris Melhuish

To my wife, Elaine. I am forever grateful for all of your loving support and encouragement.

Abstract

Decision-making is an important operation for any autonomous system. Robots in particular must observe their environment and compute appropriate responses. For solitary robots and centralized multiple-robot systems, decision-making is a relatively straightforward operation, since only a single agent (either the solitary robot or the single central controller) is solely responsible for the operation. The problem is much more complex in a decentralized system, to the point where optimal decision-making is intractable in the general case. Decentralized multiple-robot systems (dec-MRS) are robotic teams in which no robot is in authority over any others. The globally observed behaviour of dec-MRS emerges out of the individual robots' local interactions with each other. This makes system-level decision-making, an operation in which an entire dec-MRS cooperatively makes a decision, a difficult problem. Social insects have long been a source of inspiration for dec-MRS research, and their example is followed in this work. Honeybees and *Temnothorax* ants must make group decisions in order to choose a new nest site whenever they relocate their colonies. Like the simple robots that compose typical dec-MRS, the insects utilize local, peer-to-peer behaviours to make good, cooperative decisions. This thesis examines their decision-making strategies in detail and proposes a three-phase framework for system-level decision-making by dec-MRS. Two different styles of decision are described, and experiments in both simulation and with real robots were carried out and presented here to demonstrate the framework's decision-making ability. Using only local, anonymous communication and emergent behaviour, the proposed collective decision-making framework is able to make good decisions reliably, even in the presence of noisy individual sensing. Social cues such as consensus and quorum testing enables the robots to predicate their behaviour during the decision-making process on the global state of their system. Furthermore, because the operations carried out by the individual robots are so simple, and because their complexity to the individual robots is independent of the population size of a dec-MRS, the proposed decision-making framework will scale well to very large population sizes.

Acknowledgements

No one completes a thesis on their own. Without the support and advice of many people, I would not have been able to complete this work. First, I would like to thank my supervisor, Dr. Hong Zhang. Over the course of my graduate studies, Dr. Zhang has always given me substantial freedom in choosing the ideas that I pursued, but he has also given me a lot of valuable advice and taught me a great deal about academic research and writing. I also would like to thank the other members of my supervisory committee, all of whom listened to my ideas and provided useful suggestions for this research. Several staff members in the Department of Computing Science at the University of Alberta also have been very helpful, booking facilities for my experiments, etc. In particular, I would like to thank Fran Moore, Edith Drummond, Karen Berg, and Sharon Gannon. Also, thank you to Dr. Stephen Pratt (Arizona State University) and Dr. Thomas Seeley (Cornell University), who responded to my endless emails inquiring about their research into the natural decision-making behaviours of ants and honeybees. Their input has been greatly appreciated. Finally, this research has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC, Government of Canada), the Alberta Informatics Circle of Research Excellence (iCORE, Government of Alberta), and the Killam Trust.

Table of Contents

List of Figures

List of Symbols

Symbol	Description
N	The number of individuals that compose a multiple-robot system.
N_a	The number of robots that favour a particular alternative in common.
N_i	The number of robots that favour some particular alternative i.
N_o	The number of robots that do not favour any alternative.
C_a	Apparent consensus: the proportion of a robot's teammates that favour the same alternative as it when its own opinion is not included. $C_a = \frac{N_a - 1}{N - 1}$.
$\tilde{C_a}$	Apparent consensus as estimated by an individual robot.
C_t	True consensus: the proportion of the robots in a system that favour some alternative in common. $C_t = \frac{N_a}{N}$.
Q	Quorum threshold $\in [0, 1]$: the number to which a robot compares its estimate of apparent consensus to determine whether or not its favoured alternative has satisfied quorum.
n	The number of teammate opinions used to compute an estimate of apparent consensus using the digital approach.
τ	The time constant that sets the rate of decay of an exponentially decaying variable, *e.g.* $e^{-t/\tau}$. Used in the analog approach to apparent consensus estimation.
T_0	The expected time between an individual robot's encounters with its teammates.
T_a	The expected time between an individual robot's encounters with teammates that favour the same alternative as it.
T_c	The commitment timeout: the period of time that a committed robot will remain in the committed state without receiving an acknowledgment to a commit-message before exiting a decision.
T_l	Recruiter lifetime: the length of time that a robot remains in the recruited state before returning to the idle state. Used only in honeybee-style iterative recruitment.
T_r T_{r_i}	The inter-recruitment period; the period of time between a robot's attempts to recruit the teammates that it encounters. T_{r_i} is T_r for some particular alternative i.
β_o, β_i	T_{r_o} and T_{r_i} expressed as rates rather than periods of time.
$q(t)$	The quorum index: an exponentially decaying variable that is incremented by a constant amount, Δ, every time that a teammate that favours the same alternative in a decision as it is encountered.
$k(t)$	The kin index: an exponentially decaying variable that is incremented by a constant amount, Δ, every time that a teammate is encountered.
Δ	The constant by which the quorum and kin indices are incremented when an appropriate teammate is encountered.
E_r	The expected number of teammates that a robot will recruit during its lifetime. Used only in honeybee-style iterative recruitment.

Symbol	Description
$vote_i$	The i^{th} most recently received vote-message received by a particular robot.
$\gamma(vote)$	A function that quantifies a vote-message to either 1 or 0.
r	The radius of a robot.
d	Distance
t	Time

Introduction

In the last twenty years, there has been a steady increase in the effort to build intelligent systems composed of many robots that cooperate with each other in order to achieve shared goals. The development of these *multiple-robot systems* (MRS) presents a significant challenge. Not only must the individual robots be mechanically, electrically, and computationally reliable, but their social interactions must be robust so as to bring about reliable collective behaviours, too. Despite these challenges, MRS are attractive for many reasons. A single robot cannot be in two places at once, nor can it be both large and small at the same time, but a MRS can [3]. Furthermore, a multiple-robot solution will tend to be more reliable, since several of the individual robots could fail and yet the MRS would continue to function.

Not all MRS are the same. Although two different systems could be compared along several different axes [24, 25], their organizational structure is particularly important. At one end of the spectrum are centralized MRS, in which a single centralized agent (a specialized robot, human operator, etc.) is in command of every other member of the system. The other end of the spectrum is occupied by decentralized MRS (dec-MRS). In a dec-MRS, none of the robots are in control of any of their teammates. Dec-MRS are particularly interesting, and it is these systems that are the focus of this thesis.

Robots by their very nature are decision-making machines. Supplied with a goal and the physical means to achieve it, the operation of a robot consists of an endless loop of sensing its environment, computing an appropriate response, and then carrying it out. For a solitary robot, deciding what to do is relatively simple to understand, since solitary robots have no teammates with which they must coordinate their actions. Similarly, decision-making by a centralized MRS is straightforward, because the centralized agent needs only to collect information from the robots that it controls, make a decision, and then dictate its decision back to them. Decision-making by a dec-MRS, however, is much less obvious [63]. Because their collective behaviours are bottom-up, emerging out of the myriad of individual robot-to-robot interactions, there is no central point from which a collective decision could emerge.

If dec-MRS are to be deployed as autonomous intelligent entities, rather than systems that continually must be monitored and have decisions made for them by external agents, then system-level decision-making is a problem that must be solved. It is important to distinguish *system-level* decision-making as a special operation. Many dec-MRS have been described in which decisions are made, but these tend to be local decisions made by individual robots that only affect them or their nearby teammates, rather than decisions that are made collectively by the entire population. The ability to make system-level decisions provides the illusion of centralized control, allowing a dec-MRS to be viewed and thus programmed as a cohesive intelligent entity, a *superorganism*. Optimal decision-making in a decentralized system in a realistic environment is intractable [7, 63], unless free, instantaneous communication is available[1]. Therefore, real-world collective decision-making in a dec-MRS demands a heuristic approach, especially when a dec-MRS contains many robots as it is assumed in this work.

When the survival of an organism depends on its ability to make good decisions, evolution will find a way for those decisions to be made, or the organism will be out-competed and become extinct. Social insects such as ants and bees are excellent natural analogs for large dec-MRS, and their examples have been followed many times with success by roboticists [8]. A particularly important system-level decision that a colony of social insects might have to make is to select a new site for its nest. A poor choice would penalize a colony long after the decision had been made, and so there is a strong evolutionary pressure to develop efficient collective decision-making behaviours. Both honeybees (*Apis mellifera*) and certain *Temnothorax* ants (*T. albipennis* and *T. curvispinosus*) often must make precisely this decision [78, 50, 71]. Because the complexity of the individual insects is similar to that of small mobile robots, and because their colonies are organized as decentralized systems, their approach to collective decision-making is particularly attractive for application to dec-MRS.

This thesis describes the adaptation of these social insects' decision-making strategy for modern dec-MRS. The result is a three-phase decision-making framework: search, deliberate, and commit. The nature of the approach developed by this work is intentionally general, permitting a wide variety of collective decisions to be made. Furthermore, like the behaviours of the social insects that inspired it, the decision-making framework relies only on local interaction and simple, short-range broadcast communication. This means that almost any dec-MRS could take advantage of this work. Two social behaviours central to the framework are developed in detail, and these could be applied to many problems beyond the focus of this thesis. These behaviours are iterative recruitment and consensus estimation/quorum testing. The first one enables a decentralized system to compare a list of alternatives and identify the best one. Especially important is that the precision and accuracy of the comparison accomplished by iterative recruitment *increases* as the population of a dec-MRS increases in size. The second collective behaviour, consensus estimation and quorum testing, allows the individual robots to predicate their own behaviours on the collective state of their system, providing a powerful social cue. The cost to the individual robots of these behaviours is independent of the population size of their system, so scaling a dec-MRS that employs them up to large population sizes

[1] Even if this impossibility somehow was overcome, optimal decision-making still would be PSPACE-hard.

will be economical. Experiments were conducted in simulation and with physical robots (designed and built especially for this work) to demonstrate the performance of the proposed decision-making framework in practice.

The main contribution of this thesis, to be reiterated in greater detail in Chapter 6 is as follows. It is shown that a MRS can make collective decisions without any sort of centralized control, using only local and anonymous communication. The layout of this document is as follows. In Chapter 1, some of the research related to the focus of this thesis is described. Next, in Chapter 2, the nest site selection behaviours of honeybees and *Temnothorax* ants are presented. The three-phase decision-making framework inspired by the insects' behaviours is presented in Chapter 3. Decisions can take many forms. Chapters 4 and 5 describe implementations of the framework to tackle two types of group decision. In Chapter 4, decisions to accept or reject a single proposed alternative to the *status quo* are described, and the results from experiments using both simulations and real robots are presented. Not all decisions can be represented by the accept-or-reject model, called a *unary decision* in this work. A more general-purpose approach is the *best-of-N* decision [79], in which the single best of N candidate alternatives must be selected. Best-of-N decision-making is described in Chapter 5 along with a series of robotic experiments. The thesis closes with Chapter, 6 in which the significance of this work is discussed, and the next research steps are outlined.

Chapter 1

Related Work

The focus of this thesis is system-level decision-making in decentralized multiple-robot systems, referred to in this thesis as *dec-MRS*. Although there has been relatively little work studying this problem in dec-MRS with large populations of relatively simple robots, this area is related to several others. In this chapter, several of these fields are discussed, including their relation to the goals of this thesis. Natural decentralized systems are very relevant to large dec-MRS research, so their discussion is given an entire chapter of its own, immediately following this one.

1.1 Decentralized Multiple-Robot Systems

The origins of multiple-robot systems research can be traced to the pioneering efforts of Grey Walter in the late 1940s [89, 39], but interest in this area has increased in recent years [26, 14]. The published studies describe systems too varied and numerous to be summarized here, but there has been relatively little work investigating system-level cognitive operations in large decentralized multiple-robot systems (dec-MRS) and, in particular, collective decision-making.

A dec-MRS is characterized by the complete absence of specialized agents that make plans and decisions on behalf of the rest of the system's members [14, 24, 26, 25][1]. Dec-MRS that contained relatively large numbers of robots were somewhat common in the early MRS literature. These systems were able to complete collective tasks such as sorting [19, 52, 38], foraging [84], cooperative load transport [47], construction [86, 91], and the stick pulling experiment [42]. The manner in which collective sorting was implemented is particularly illustrative of the general approach to decentralized control of a MRS common to many of these systems. Initially, the robots' environment would contain many objects of two or more colours, randomly scattered about. The robots collectively sorted these into piles based on object colour through a series of stochastic pick and place operations. As a robot wandered about, it would find different objects. If a robot encountered an isolated object, or one near to a number of objects of a different colour, the robot would be more

[1] Note that there is some disagreement in the literature as to whether the term "distributed" or "decentralized" should be used to describe the systems intended by the term *decentralized* in this work. Readers should keep the meaning intended by the use of this term in this thesis in mind when reading related literature.

likely to pick it up. While carrying an object, a robot would be more likely to put it down when it encountered a cluster of like-coloured objects already on the ground. In this way, the environment itself directed the robots' sorting operation. As the robots worked, the manner in which they modified the environment through their actions further stimulated their behaviour. This *stimulation by the progress that had been achieved* is known as *stigmergy* [6], and is still a common feature of dec-MRS control.

It is important to point out that explicit cooperation was unnecessary for this sorting behaviour to work. A solitary robot could have sorted the objects just as well as a MRS, except that it would take longer to do so. The only time that two robots ever had to deal with each other was to avoid interference and collisions. Some tasks, such as cooperative transport [47] and the stick pulling experiment [42] did require the collective efforts of multiple robots in order to be completed, but the individual robots still were ignorant of each other's existence. Instead, success *emerged* out of their interactions. The problem with stigmergy, the most common form of emergent behaviour in dec-MRS is that it critically depends on the dynamics of the interaction between the robots and their environment [6]. A stigmergic dec-MRS and its environment together can be thought of as an elaborate Rube Goldberg machine. Changing one small aspect of the system, perhaps the interaction of a robotic bulldozer's plow and the fill that must be moved, could have a significant (and potentially disastrous) effect on the likelihood of successfully achieving the overall goal [64]. The dec-MRS operations at the heart of the decision-making framework proposed by this thesis also are emergent, taking advantage of the bottom-up nature of dec-MRS. However, instead of modifying the environment, the robots directly modify the states of their teammates, reducing or eliminating the impact of the robots' environment on their decision-making behaviour.

Recently, more sophisticated dec-MRS have begun to appear, composed of very large numbers of robots, on the order of 100 or more [51, 46]. However, the majority of the research conducted with them concentrates on either strategies to allow an external operator to centrally control a dec-MRS (*e.g.* teleoperation), or the development of local behaviours to realize desired global emergent behaviours. For example, [51] discusses swarm teleoperation and mentions several emergent behaviours that could be controlled with the proposed interface. In some cases, although the specific dec-MRS might contain many robots, the proposed algorithms utilize only a few of them at any given moment. The environment mapping operation in [46] uses "a small number (1-5) of robots working completely autonomously, often out of contact with the base station". In a sense, this does not describe a very large robotic team, but rather a very small one with many spare robots standing by. Nonetheless, these systems do not command themselves. The decision-making upon which this thesis focuses would permit a dec-MRS containing many robots to collectively monitor and respond to itself and its environment.

1.1.1 Collective Decision-Making In Dec-MRS

There have been several different approaches to the collective decision-making problem that have been described for dec-MRS. However, it appears that only one of these actually is practical for systems composed of many simple individuals.

Competitive team sports is a domain in which multi-agent systems must make system-level decisions relatively often. RoboCup [75] is an international competition in which MRS developed by research institutions from all over the world are pitted against each other in a robotic version of soccer. The competition is divided into several different leagues, each emphasizing different technical challenges. The Middle-Sized League (MSL) is the most relevant to dec-MRS research, since teams in the MSL often are organized as decentralized systems, and the individual robots are not permitted to take advantage of hardware beyond the robots themselves, nor do they have complete knowledge of the state of the game. Soccer also is a particularly interesting domain, since it is not episodic (*i.e.* the two opposing teams do not take turns as they would in a game like chess), and so decision-making speed and accuracy must be balanced appropriately.

The most obvious kind of system-level decision that a robotic soccer team would need to make is cooperative play selection. Given the current state of the game and the skills of the opponent, the robots must decide on the most effective play to run. Because each robot might have a different opinion of a game's state, these decisions must be coordinated so that all of the robots will agree about the correct play and each robot's role in it. Kok *et al.* have described several play selection strategies intended for dec-MRS soccer teams in the MSL [45, 88]. One particular approach presented in [88] is called an *anytime algorithm*, since it begins with a feasible play that is continually refined. Thus at any time, the decision-making process could be terminated and a runnable play still would be produced. Soccer teams typically are composed of a small number of relatively sophisticated players. Therefore their decision-making strategies are not generally applicable to dec-MRS containing a large number of simple robots.

Free market strategies and, in particular, auctions [33] also have been suggested to guide the decision-making of dec-MRS. The task allocation problem [34] is an area identified as well-suited to the auction strategy. The general idea of auction-based task allocation works as follows. Individual robots identify tasks, and put these up for auction. Other robots then can submit bids based on their cost of executing them. The auctioneer rewards the robot that submitted the lowest bid with the task. Robots can bid on and win multiple different tasks, in turn auctioning off the most expensive ones in the hopes that another robot might be able to complete them even more economically. Over time, tasks will be allocated to robots so that the total cost of completing all of the tasks will be minimized.

A conflict arises when a dec-MRS composed of many robots uses auctions to allocate tasks amongst its members. If an auction is to be worthwhile, it must attract a sufficiently large audience of potential bidders that the auctioneer is likely to "sell" its task. This in turn suggests that the auctioneer should increase the range of its transmissions. However, since many auctions are supposed to be run in parallel, long-range communication would increase the likelihood of different auctioneers (and their bidders) interfering with each other. It seems as though auctioneers should heed the advice of [36] and reduce the ranges of their transmissions as much as possible, leading to many short-range auctions, or each robot should take its turn and conduct its auction(s) with long-range communication. The former would lead to many ineffective auctions, whereas the latter implicitly requires a centralized controller to implement the necessary time-division multiplexing of the shared communication channel.

In [53], another collective decision-making system was described. A simple, emergent decision-making strategy was presented to enable a dec-MRS to form small groups of robots that would depart in convoys from a rendezvous point. The proposed mechanism at the heart of the robots' behaviour was *chorusing*, inspired by the natural abilities of frogs, fireflies, and crickets to synchronize their emission of signals. By monitoring the strength of the collective chorus, a kind of social cue, the individual robots were able to estimate the total number of their teammates that had assembled. In this way, a subset of a dec-MRS was able to make a kind of collective decision to depart together. However, the collective departure was brought about using a simple open-loop mechanism. As individual robots in the assembling convoy came to believe that the group had become sufficiently large, they would set internal countdown timers. Once a robot's timer had reached zero, the robot would emit a special signal, most easily understood as a message to those other robots assembled: "it is time for us to depart". The duration of the countdown is important, and is difficult to tune without empirical data. Furthermore, although the departure of the convoy was a collective behaviour, the decision to depart was made by an individual: the first robot to believe that a sufficiently large number of robots had gathered. As the desired population size of the convoy is increased, so to would the probability of at least one robot overestimating the actual size of the assembled population, which would result in a premature collective departure.

One of the few studies to date that has described a system-level decision-making strategy that truly is well-suited to large-population dec-MRS is that of Wessnitzer and Melhuish, in [93]. In that work, a large dec-MRS was tasked with pursuing and immobilizing two "prey" in a series of simulated experiments. The robots, each of which possessed minimal sensing capability and short communication ranges, used majority voting and a hormone-inspired approach to cooperatively decide which prey to follow. Initially, the system would collectively decide to follow one of the prey based on which of the two was the closest to the majority of the system's robots. Once it had been immobilized, and the robots agreed that this was the case, their focus would switch to the other prey. Here, *system-level* decisions were made. *We agree* that that prey-A should be pursued. *We agree* that the first prey has been immobilized, and thus the other one should be pursued. The decision-making framework proposed by this thesis is similar in some ways to the approach presented in [93]. However, the strategy presented by this thesis is very loosely coupled to the specific decision at hand by design. It therefore will be applicable to a wider variety of decision-making scenarios. Furthermore, the addition of a commitment phase synchronizes a dec-MRS's individual robots' exits from a given decision.

1.1.2 Gossip

In a dec-MRS, there is no central agent to coordinate inter-robot communication. In many ways, the robots' interactions constrained similarly to nodes in a sensor or ad-hoc network. In such a system, the data throughput is maximized when each node or robot communicates with the minimum range possible [36]. Unlike the nodes of a typical sensor network, however, the individual robots of a dec-MRS continuously move about. To minimize the likelihood of different peer-to-peer robot conversations interfering with each other, it is advocated in this work that robot communication ranges be made as short as possible, and that the robots move about in order to find teammates with which to communicate. For systems composed of very small robots, physical movement is less expensive than communication or computation [13], so this strategy, in addition to reducing interference, will also be energy-efficient for swarms of nano- or micro-robots.

Stochastic peer-to-peer communication, known more commonly as gossip [10] is intended to operate in a network environment very similar to that described above. Gossip algorithms for decentralized averaging [10], message routing, spanning tree computation [44], and resource location [43] have been presented, amongst others. One advantage that a dec-MRS has over a sensor network is that the individual members of the dec-MRS are mobile, whereas the nodes in a sensor network typically are stationary [1]. While the continual making and breaking of network connections due to the constant movement of the robots might be a weakness from a sensor networks perspective, it can be viewed as an advantage in the context of this research. If an individual robot randomly wanders about and communicates sufficiently infrequently with the robots that it encounters, the communication partners that it has will be completely random and uniformly distributed over all of its teammates. This means that the necessary conditions for *uniform gossip* will be satisfied. It has been shown that information can be spread to all n nodes of a system with a given reliability in $\mathcal{O}(log(n))$ time steps [43], where the length of each time step is the time between a robot's communicative encounters with teammates. These conditions apply in the proposed decision-making framework of this thesis during the period of iterative recruitment known as the *deliberation phase* (see Section 3.3).

In the later *commitment phase* of a collective decision (Section 3.5), the robots communicate with every teammate that they encounter as they wander about, so the communication partners of each robot will be less uniformly distributed, and the resulting pattern will resemble something in between that of uniform gossip and *spatial gossip* [43]. In both cases, attrition of knowledge is involved, since the individual robots described in this thesis forget what they previously knew when they are given new information. At the time of writing, attrition in gossip algorithms does not appear to have attracted significant attention.

1.2 Markov Decision Processes

A Markov decision process is a mathematical framework that allows an agent to reason about its actions in some environment. They are common in the field of artificial intelligence, as they can be combined with machine learning techniques to enable an agent to learn plans of action, called policies. Markov decision processes (MDP) come in many different varieties, only three of which are discussed here, briefly. They are not generally useful for decision-making in dec-MRS that are composed of may simple agents. However, they are included in this work, as a thesis on decision-making would be lacking were it to omit mentioning them altogether.

The simplest of these processes, denoted MDP, describes a single decision-making agent in a completely observable environment. It is represented by a tuple: $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \beta \rangle$. The world is modeled as a finite set of states $s \epsilon \mathcal{S}$, a finite set of actions $a \epsilon \mathcal{A}$ that the agent could take, and a transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \prod(\mathcal{S})$ that provides a probability distribution over $\mathcal{S}$ for each action in each state. That is, if an agent takes action a in state s, $\mathcal{T}$ specifies the probability of the agent ending up in each of the possible states $s' \in \mathcal{S}$. $\mathcal{R} : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ is a reward function giving the expected immediate reward to the agent for taking action a in state s. Finally, $0 \geq \beta \geq 1$ is a discount factor that reduces the immediate value of the reward for future actions[2] [49]. Numerous methods have been developed over the years to derive optimal policies for an agent when an MDP applies. An optimal policy for a finite horizon MDP can be found in polynomial time [56].

The main weakness of the basic MDP is that it assumes that the complete state of the environment is available to the agent. That is, the agent knows precisely the current state $s \in \mathcal{S}$ of the world. This, however, is an unrealistic assumption. In reality, an agent only can know about the likelihood of being in a particular state. The *partially observable* MDP, or POMDP, incorporates this uncertainty about the state of the world. As Monahan put it, if an MDP models a frog hopping from one lily pad to the next on a pond on a clear day, then a POMDP does the same in foggy weather [55]. A POMDP also is represented by a tuple, but it contains additional elements: $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{Z}, \mathcal{O} \rangle$. $\mathcal{S}, \mathcal{A}, \mathcal{T}$, and $\mathcal{R}$ are the same in a POMDP as they are in a MDP. $\mathcal{Z}$ is a finite set of observations that the agent could make, and $\mathcal{O} : \mathcal{S} \times \mathcal{A} \rightarrow \prod(\mathcal{Z})$ is an observation function. $\mathcal{O}(s', a, z)$ is the probability of the agent making the observation z upon entering state s' after to taking action a [49]. Like a regular MDP, the solution to a POMDP is a policy for just one agent. However, the addition of the uncertainty regarding the state of the world considerably complicates policy derivation. The complexity of a finite horizon POMDP is PSPACE-hard, meaning that the optimal policy cannot be found in polynomial time with a deterministic algorithm [7].

There have been several extensions to the MDP and POMDP frameworks to model problems

[2] This is somewhat like inflation in a financial system. Getting paid $500 today is more valuable than getting paid $500 in the future, and β provides this discount on future rewards to the agent so that expected rewards can be compared on a level playing field.

in which multiple agents must be considered. The one most applicable to decision-making in a decentralized MRS is the decentralized POMDP, or DEC-POMDP [7]. Here, each world state $s \in \mathcal{S}$ includes the state of each individual agent. The transition probabilities for the world now take the form $P(s'|a^1, \ldots a^m)$, where a^i is the action taken by agent i. The reward function $\mathcal{R}$ now also depends on every agent's action, and each agent i is given its own observations, Ω^i (Bernstein *et al.* use the symbol Ω to indicate the observations that Littman denoted $\mathcal{Z}$). The observation probabilities $\mathcal{O}$ also become much more complex, taking the form $\mathcal{O}(o^1, \ldots, o^m|a^1, \ldots, o^m)$, since each agent's individual observations depend on the actions taken by it and every other agent. Clearly, this is a much more complicated problem than either the MDP or the POMDP. It has been shown to be NEXP-complete for all but the trivial single agent case (which is just a POMDP), which means that it *provably* cannot be solved in polynomial time. In the special case in which the union of all of the agents' observations completely specifies the state of the world, a DEC-MDP is produced (*i.e.* the world is collectively observable), and these are NEXP-complete for systems composed three or more agents [7].

The general DEC-(PO)MDP does not assume that the agents can communicate with each other. Its complexity arises from the fact that every agent must reason about every other agent in its system, combined with all of the related uncertainty. It has been shown that the addition of free instantaneous communication reduces a DEC-POMDP to a POMDP [73]. However, communication in the real world is not free, nor is it instantaneous. The inherent intractability of the DEC-POMDP has forced researchers to seek out heuristic solutions [16]. This has lead to a wealth of additional frameworks to capture the assumptions made by such studies. Pynadath and Tambe have developed a new model called the COMmunicative Multiagent Team Decision Problem, or COM-MTDP so that all of these frameworks can be compared equitably [74].

Clearly, a relatively sophisticated agent is required to develop policies using the DEC-POMDP framework or its derivatives, and the problem only becomes more difficult as the number of agents increases and their environment and potential actions become more fine-grained. Furthermore, even if such agents were practical in a dec-MRS, it is not at all clear how the collective decision-making problem could be modeled as a Markov decision process. Without some model of a collective decision as a set of states with potential transitions between them (*i.e.* what are $\mathcal{S}$ and $\mathcal{T}$ in a generic collective decision?), results from the MDP literature are unlikely to be of much use. It is for these reasons that these frameworks are not practical for the collective decision-making by the simple, resource constrained, real-world agents that are the robots of the large-population dec-MRS for which this work is intended.

1.3 Ant Colony Optimization

Many eusocial insects [9] such as ants use emergent recruitment behaviours to collectively compare different alternatives that are known to their colonies. Recruitment is central to the decision-making framework proposed by this thesis, but this work does not describe the first algorithm for an artificial system to take advantage of it.

Dorigo *et al.* have developed a novel optimization algorithm called Ant Colony Optimization, or ACO, that applies ant-inspired pheromone trail following to solve otherwise intractable optimization problems [21, 20]. ACO mimics the ability of ants that lay down chemical trails (such as *Lasius niger*) to find the most efficient paths between their nests and food sources. When a real trail laying ant finds a food source, it lays down a trail of a volatile chemical called pheromone as it travels back to its nest. The amount of pheromone that it deposits increases with its opinion of the food source. When other ants encounter this trail, they will be recruited to follow it with a probability that increases with the concentration of pheromone in it. These ants make their own evaluations of the food source, and reinforce the trail with their own pheromone as they head back to the nest. This increases the likelihood of the trail recruiting even more ants. Because each ant will tend to cut corners, the trail also will be refined over time as though it was a string between the nest and the food source being pulled taught. Eventually, the trail will converge to follow the lowest cost path between these two points.

ACO adapts this emergent social behaviour by representing complex optimization problems spatially in a simulated environment such that candidate solutions to them will take the form of paths through the solution space. *Artificial ants* search for solutions by wandering about, and then lay down an *artificial pheromone trail* to identify them. Other artificial ants are more likely to to follow paths in the solution space that are marked by stronger pheromone trails, and the corner-cutting behaviour of the individuals refines the identified solution(s) to be locally maximal. Over time, the behaviour of the entire artificial colony will tend to converge to a single solution to the problem [21], which then can be identified by an external user.

Because the real ants have evolved their behaviour to find the shortest paths in their environment, ACO is particularly well suited to problems such as the traveling salesman problem, often encountered in networking domains. ACO, however, is not an autonomous behaviour when considered at the system level. The artificial ants do not recognize the purpose of their behaviour, instead continually applying their local rules at the micro level[3]. A complex problem must be set up for ACO by some external operator, the process initiated, and the solution later identified after the system has converged. Thus, it is primarily an engineering design tool [21], somewhat similar to simulated annealing [15], an optimization strategy that applies the principles of thermodynamics to obtain good solutions to combinatorial problems.

[3]The individual ants could be thought of as specialized particles that are released into an engineered environment, in which their collective behaviour is ascribed special meaning. In this way, they are no different than the molecules of oil and air in a bubble-level, the collective behaviour of which permit an external observer to know when the device is horizontal.

1.4 Quorum and Quorum Testing

A quorum is defined as *the number of members of any society or assembly that must be present if the business done is to be legal or binding* [2]. In other words, a quorum is a threshold of participants below which action by a group or collective would be inappropriate. Quorums are common in the formal rules of order used by governments and other deliberative assemblies, such as Robert's Rules of Order [41], but they also are common in the natural world. Some species of ant and bee employ quorums when they decide on the location of a new nest site. By delaying their commitment to a particular site until a quorum of insects agree that it should be chosen, a system is able to make decisions much more accurately than an individual insect would be able to [31]. Even bacteria, some of the simplest organisms of all have been found to employ quorums so that their collective behaviour can be socially coordinated [90, 54]. This suggests that quorum is an effective, yet low cost social mechanism.

Social insects and other simple, socially interactive species have been the inspiration for numerous works in computing science and engineering. The manner in which they test quorum has been adapted for used in artificial systems, too. In [67], a mobile agent on a mobile ad-hoc network hopped from host to host, testing quorum to determine whether or not a sufficient proportion of the hosts agreed about some proposed action (*e.g.* revoking the key of a malfunctioning host). Only once the agent believed that quorum had been satisfied would the action be taken. The manner in which the mobile agent tested quorum is somewhat similar to the digital quorum test presented in Section 3.4.2, except that in this work, many robots test quorum simultaneously.

1.5 Summary

In this chapter, several areas of research related to decision-making in dec-MRS have been discussed. Much of the existing decision-making work has focused on smaller systems composed of relatively sophisticated individuals, whereas studies of large-population dec-MRS have tended to focus on emergent behaviours and their remote control by an external and centralized operator. One field of great importance to dec-MRS that has been neglected in this chapter, however, is that of biology. The next chapter is devoted to this area and, in particular, the behaviours evolved by some naturally occurring decentralized systems (which obviously lack remote teleoperators to make system-level decisions on their behalf) to make system-level decisions.

Chapter 2

Cooperative Decision-Making by Social Insects

Cooperative decision-making is common in animal societies, both in human civilization and in the natural world [17]. In this chapter, the decision-making behaviour of honeybees (*Apis mellifera*) and the ants *Temnothorax albipennis*[1] and *Temnothorax curvispinosus* is examined in detail. It is their ability to make unanimous group decisions through only local interaction and communication, despite the decentralized nature of their colonies, that is the inspiration for the decision-making algorithm at the core of this thesis.

2.1 The Nest Relocation Behaviours of Honeybees and *Temnothorax* Ants

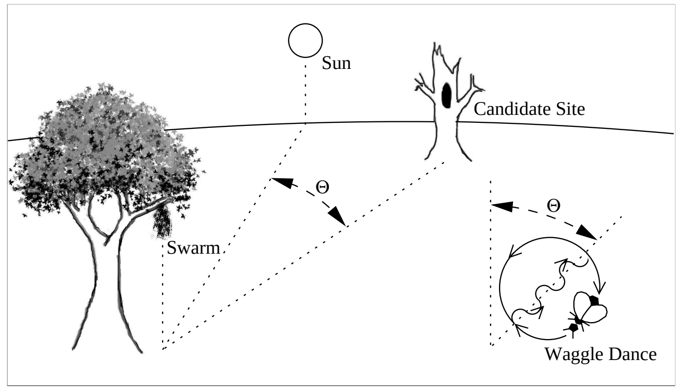
The decision-making algorithm proposed by this thesis for multiple-robot systems was inspired by the cooperative nest relocation behaviours of honeybees and *Temnothorax* ants. In this section, a description of the manner in which each species selects a new home is given in detail. Both appear to have evolved the same approach to decision-making independently, adapting pre-existing behaviours to a common framework.

2.1.1 Nest Relocation in Honeybees

Honeybees are a well-studied social insect due to their long-standing economic importance. Because the bees put so much of their effort into the production of honeycombs and the storage of food, the construction of a nest represents a substantial investment [31]. A colony that is successful will tend to increase the size of its population, and eventually it will outgrow the cavity (*e.g.* an old hollow tree trunk) in which it had built its nest [78]. If it is to continue to prosper and grow, a new nest site must be found.

[1] Close inspection of the literature will identify two other species of ant: *Leptothorax tuberointerruptus* and *Leptothorax albipennis*. These names are in fact former misclassification of *Temnothorax albipennis* used prior to its current name.

(a)



(b)

Figure 2.1: Honeybees collectively decide on their new home after they have left their old nest and formed a swarm on a tree branch or other structure, shown on the left. The individual bees search out candidate nest sites, and advertise them to other bees at the swarm using the waggle-dance shown on the right. Bees that have found better sites tend to perform more dances than those that favour poorer sites, so the best site will attract the majority of the decision-makers. Once a bee determines that its favoured site is sufficiently popular, it rouses the swarm and helps guide it to its new home. The photo in 2.1(a) is copyright Thomas D. Seeley and reproduced with his permission, 2.1(b) after Figure 1 in [31].

The relocation process begins with the division of a colony. The queen and half of the workers form a swarm and leave the nest, lighting upon a nearby tree branch or some other structure (see Figure 2.1(a)). The remaining workers and a newly reared queen are left at the old hive to continue on there [80]. The majority of the bees in the swarm cluster around the queen and become dormant, so as to protect her and conserve their energy reserves (they gorge themselves on honey before leaving their old home). Some of the swarm's members (about 5% by population) remain active and make the decision about where the swarm should build its new nest [31][2]. Several of these bees act as scouts and search the countryside for potential nest sites, often covering an area as large as 150 square kilometers [31].

The bees are able to determine the absolute quality of a potential site based on a variety of features [79]. Once a scout has found a site and measured its quality, it returns to the swarm. There, it advertises its site to other bees using the waggle-dance (Figure 2.1(b)), an elegant behaviour that communicates the location its site [78]. The number of individual dances that a bee will perform to advertise its site depends on its own opinion of the site's quality. Better sites will tend to illicit more dancing from the scouts that find them [11]. Each dance that a bee performs is a set of waggle-dances. Other bees that observe a scout's dances might fly to the specified site to evaluate it for themselves. In practice, multiple sites will tend to be found and advertised at the swarm. The probability of a bee flying to visit a site advertised by a dance that it has observed is fixed, and the bees observe dances at random, so the probability of a dance-observing bee flying to visit a particular

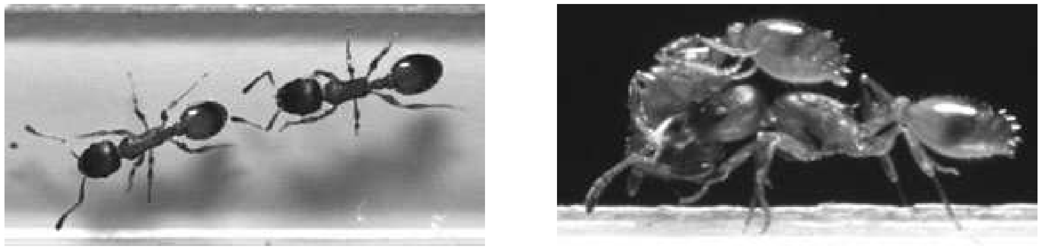
[2]It is not clear how this division of labour is accomplished. Perhaps, as it is the case with harvester ants [35], the older individuals take on the relatively dangerous scouting role, since these insects would be the most knowledgeable of the surrounding environment and are close enough to the ends of their lives that the swarm could afford to risk them.

site is directly proportional to the relative number of dances performed that advertise it [87]. For example if twice as many dances are performed for site A as for site B, then a dance-observing bee will be twice as likely to visit site A relative to site B. Bees that observe a dance and fly to its site are said to be recruited to that site.

After performing a series of dances, a scout will leave the swarm to visit its site, and then return to the swarm to dance again. Each time a scout returns to the swarm, it decreases the number of dances that it performs linearly, by approximately 15 dances each time. Eventually, a bee will cease dancing altogether, and joins the dance-observing bees on the swarm. Its previous experience as a dancer appears not to bias which dances it might choose to observe or follow [77]. Because the number of dances that a bee performs is determined by its opinion of its site's absolute quality, bees that favour better sites will tend to perform more dances before they stop dancing altogether, and thus will tend to induce more of their teammates to inspect their sites than bees favouring poorer sites. In this way, the bees' independent evaluations of absolute site quality, without any direct comparisons of sites, swell the populations of bees favouring the better sites more rapidly than those favouring poorer ones. If the bees favouring a particular site are unable to recruit any of their teammates before they stop dancing, then that site will be forgotten by the swarm, thus eliminating it from the list of candidates for their new home [81, 31].

Eventually, the positive feedback of the bees' dancing and dance-following behaviour, if allowed to continue long enough, would tend to eliminate all but one of the candidate sites from the swarm's collective memory (this demonstrated in greater detail in the next chapter), leading to complete consensus in favour of one site. However, it has been demonstrated that consensus is neither necessary nor sufficient for a swarm to complete its decision. Swarms can lift off while dances for several different sites still are being performed [11, 78, 82]. Instead, it is the detection of a quorum at one of the candidate sites that triggers the bees to commit to it [80, 82]. Because each site-favouring bee regularly visits the site that it favours, each site known to a swarm will have some number of bees visiting it at any given moment. The size of this visiting population is an indicator of its popularity, and in turn, the colony's net opinion of its quality. Bees are able to infer the size of the visiting population during their visits to a site, which they then compare to a threshold called the quorum.

Once a quorum of bees is observed at a candidate site, the frequency of *buzz-running* on the swarm increases. In this behaviour, the bees that have observed quorum burrow through the swarm cluster, buzzing their wings vigourously. It is believed that this stirs up the swarm, causing the inactive majority of the bees to warm their flight muscles in preparation for lift-off. An audible signal called *piping* also increases during this process. Buzz-running and piping appear to broadcast to the dormant members of the the swarm that a collective decision has been made. The swarm then lifts off and is guided to the new site by the scouts, who appear to herd the swarm by flying through it in the direction of its new home [11].



(a) (b)

Figure 2.2: When an ant first finds a candidate nest site, it leads other ants to it one at a time using tandem-runs, shown on the left. Each ant delays prior to leading its first tandem-run, and the length of this delay decreases the better an ant perceives its site's quality to be. Therefore, better sites will have ants lead to them more rapidly than poorer ones. While at a candidate site, an ant measures the number of other ants that also are visiting it. Once this exceeds a threshold called the quorum, it stops tandem-running and instead uses transports, shown on the right. Transports are three times faster than tandem-runs, so once quorum is satisfied, the colony quickly will be relocated to the new site. Because the best site is the most likely to satisfy quorum first, the colony will tend to choose the best one for its new home. The photos in this Figure are copyright Stephen C. Pratt and are reproduced here with his permission.

2.1.2 Nest Relocation in *Temnothorax* Ants

Certain ants of the genus *Temnothorax* also encounter the collective house-hunting problem [50]. These ants live in relatively small colonies containing only a few hundred workers[3] [30], and build their nests in natural rock fissures, the fragility of which likely necessitates frequent emigrations to new nest sites [70].

The selection of a new nest begins with individual scouts leaving the nest to search for candidate sites [50]. Like the bees, the individual ants are able to measure the absolute quality of sites that they find [70]. Because the ants' search covers a relatively small area (about 1 m^2 [31]), and because potential nest sites for them are fairly common, it is not unusual for an ant to find more than one site. The ants are able to make direct comparisons and identify the single best one of those that they happen to find [31]. Following the evaluation of a candidate site, a scout returns to the current nest to recruit other ants to inspect the site for themselves. Recruitment in *Temnothorax* ants is carried out using two distinct methods: tandem-runs and transports. In a tandem-run (see Figure 2.2(a)), an ant leads its recruit from the nest to its candidate site in a follow-the-leader fashion. When an ant transports a teammate, on the other hand, it picks up its recruit and carries it to the site [50]. The advantage of a transport is that it is fast - about three times faster than a tandem-run [70]. However, a tandem run *teaches* the recruited ant the route between the nest and a candidate site, as the recruit is able to observe landmarks along the way [32].

Initially, an ant will use tandem-runs to recruit other ants to inspect its favoured site. As is the case with honeybees, the greater the number of ants recruited to a site, the more ants will tend to be visiting it at any given moment. When it visits a candidate site, an ant measures the size of the

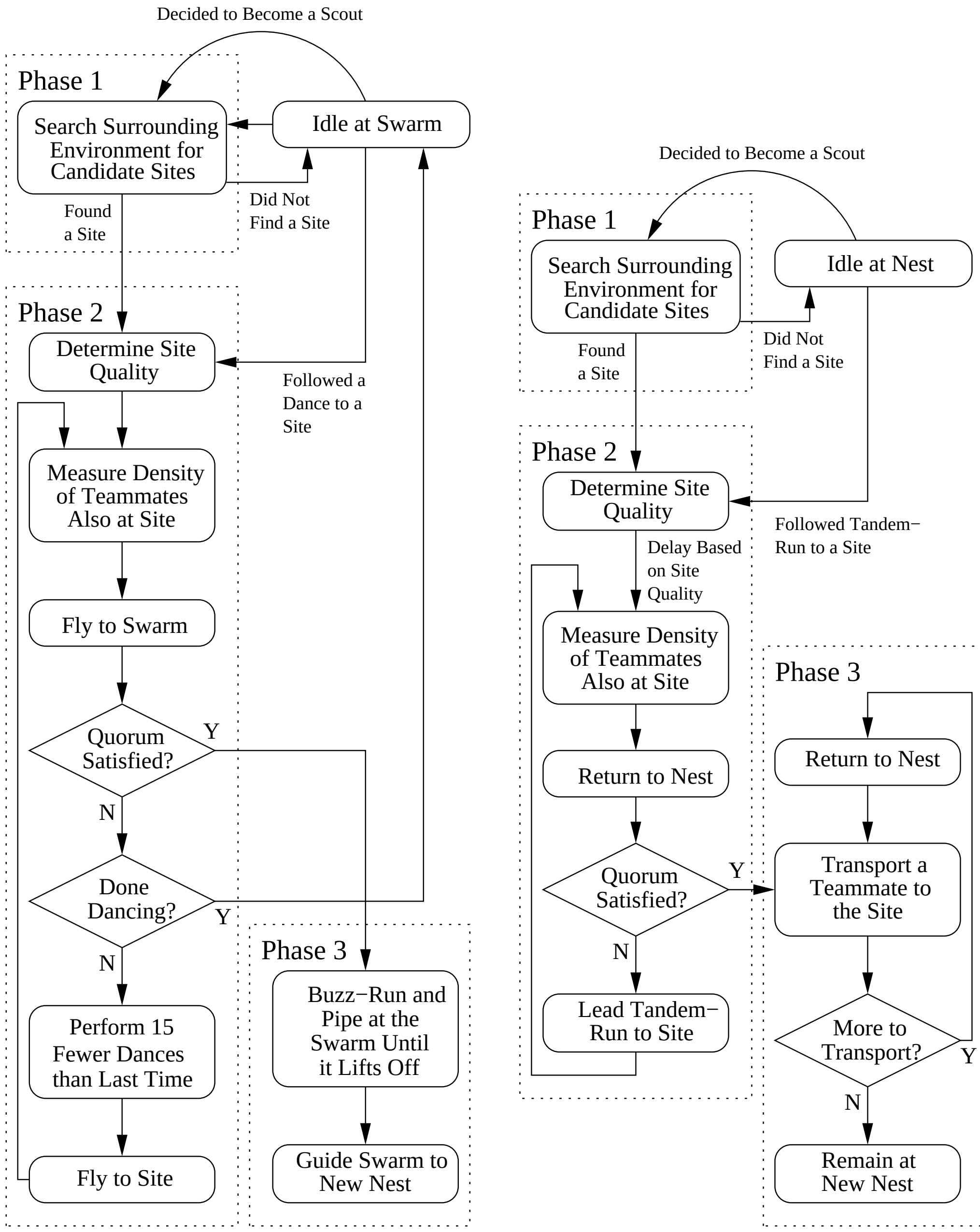
[3] In addition to workers, colonies contain a fertile queen and brood items (eggs and larval ants).

visiting population via the rate at which it encounters other ants as it wanders about there [69]. Once an ant has determined that the size of the population visiting its favoured site has reached a quorum, it changes its recruiting behaviour. Instead of using tandem-runs, it adopts the much faster transports, which rapidly increases the number of ants at its candidate site [70]. This induces additional ants to observe quorum, and they switch from tandem-runs to transports as well. The quorum used by the ants appears to be adaptive, since larger colonies tend to employ greater quorums than smaller ones [69].

The ants' recruiting behaviour is affected by their individual opinions of their favoured sites' qualities. Upon returning from a candidate site for the first time, an ant will delay for site-quality dependent period of time before leading a tandem-run. The better an ant perceives its site to be, the shorter this delay will be. This delay provides better sites with a head-start in the recruitment process relative to poorer ones. Ants brought by tandem-run to a site assess its quality for themselves, and delay their own first recruitments accordingly [31]. Subsequent tandem-runs are not preceded by a quality-dependent delay, and thus occur at a rate independent of site quality [70]. Individual ants do not drop out of the recruitment process, and so the collective behaviour of their recruiting resembles a race between the different sites to determine which will satisfy quorum first. Ultimately, the better a site is perceived to be by the individual ants, the more rapidly it will tend to increase the size of its recruited population, and thus the more likely it will be to be the first to satisfy quorum and induce the ants that favour it to switch from tandem-runs to transports.

Ants are that are brought to a site by a transport do not themselves return to the original nest to lead tandem-runs or transports. However, ants are observed leading tandem-runs in the opposite direction (reverse tandem-runs), from the site back to the original nest. This reallocates the idle ants were they are needed most: back at the original nest site so that they can help transport additional ants from there to the new nest site [70].

The tandem-running (forwards) can be thought of a probationary phase of a decision, in which the ants have made individual decisions about the candidate sites, but are waiting for a sufficient population of ants to agree with them before they commit. The observation of quorum signals to an ant that a sufficiently large agreeing population has been recruited, and so it can commit confidently to its chosen site as the colony's new home and switch to the more rapid transports in order to quickly wrap-up the decision-making process before another site satisfies quorum, too. The tandem-runs serve the additional purpose of educating a sufficient number of ants about the location of a candidate site so that, once quorum has been met, there will be enough ants able to participate in the transportation phase so that the relocation can be completed rapidly. In the event that multiple sites satisfy quorum, the reverse tandem-runs allow the quorum-satisfying sites to compete with each other via attrition to be the single new nest site ultimately chosen by a colony.



(a) Honeybee Decision-Making (b) *Temnothorax* Decision-Making

Figure 2.3: This figure summarizes the nest site selection behaviours of honeybees (left) and *Temnothorax* ants (right). Although there are differences between the two approaches, they are remarkably similar, both being organized into three distinct phases. First, individual insects search for candidate sites. Next, through a decentralized recruitment process, the known sites are ranked. The insects commit to the first site that becomes sufficiently popular (*i.e.* satisfies quorum) and it is adopted as their new home.

2.2 A Comparison and Contrasting of the Ants' and Bees' Behaviours

Both honeybees and *Temnothorax* ants demonstrate the ability to make consensus-based [17] collective decisions when they need to select a new site for a nest. The general structures of their decision-making strategies are remarkably similar (see Figure 2.3). Both are organized into three phases, beginning with a search by *individuals* for candidate nest sites in the surrounding environment. This search is followed by a period of recruitment, during which the individual insects' opinions of absolute site quality drives positive feedback, increasing the popularity of better sites more rapidly than that of poorer ones. The bees' positive feedback is driven by the tendency of idle bees at the swarm to observe and follow other bees' dances to candidate sites, and the dancing bees' ability to modulate how much they dance with their opinions of their favoured sites. The ants implement their positive feedback through the amount of time that they delay before leading their teammates one-by-one to the sites that they favour. Eventually, sufficient recruitment will take place so that one of the sites becomes so popular that the insects that favour it alter their behaviour. Both species appear to come to this conclusion based on the density of insects at a particular favoured site, and each individual that favours a site independently measures its favoured site's popularity. The ants switch from slow tandem-runs to the more rapid transports, whereas the bees begin to buzz-run at the swarm. The impact of both of these behaviours is to rapidly bring the collective decision to an end, unanimously selecting one of the alternatives that was found during the initial search.

This process of individual search, competitive recruitment and quorum-triggered commitment is a robust form of collective decision-making that requires only local communication. There are subtle differences between the ants' and bees' behaviours, however. The remainder of this section examines these differences, and discusses how they might guide the development of a collective decision-making algorithm for dec-MRS.

2.2.1 Initiating a Collective Decision

Although the bees and the ants both use their emergent decision-making behaviour to choose a new site for their nest, the manner in which they initiate their decisions differs. Honeybee decisions begin when the queen departs with half of the colony to form a swarm, and it is from the perch of the swarm that the decision is made. Although the process of forming the swarm and departing the nest might be an emergent one and thus decentralized at its root, honeybee decisions can be thought of as centrally begun. All of the bees know that a decision is necessary, and the formation of the swarm synchronizes its start. A *unary* decision, described in detail in Chapter 4, would allow a decentralized system to make such a decision, although the mechanism employed by the honeybees is not clear.

Temnothorax ants, on the other hand, are known to be opportunistic and will move from a current nest to a better site whenever one is found [23]. Their nests do not require a significant investment of energy for their construction, which take advantage of naturally occurring structures that can be used almost as-is [27, 28]. Therefore, unlike the bees, which must abandon their elaborate honeycombs when they emigrate, the ants only incur the cost of a move itself when they relocate. A single ant could initiate a decision to move to a new site if it happens to find a sufficiently good one while exploring the area around the nest. However, for a decision to be made, the nest mates that it would lead to the site that it found also would have to agree that the site was good enough to justify a colony-wide move. Eventually, an ant will give up on a candidate site if it detects a lack of progress in its recruiting, so a colony will not constantly be on the move if one of its members[4] keeps finding new nest sites in the surrounding environment.

2.2.2 Quorum

Quorum is central to the decision-making of both the ants and the bees. It is the mechanism by which an emergent site-ranking behaviour is linked to a mass-migration behaviour, the resulting synthesis being a collective decision. The use of a quorum makes a decision analogous to an election in which the candidates are the potential nest sites. Individual insects "vote" for a site that they favour by visiting it, and they also act as the pollsters by measuring the size of the population visiting a favoured site while they are there. The alternative of measuring the popularity of a site by polling individuals at the swarm or old nest, although more accurate (discussed in the next chapter), is more complicated, since it would require the individual decision-makers to explicitly communicate to each other the precise site that they favour. The spatial nature of site selection simplifies the quorum-testing problem by sorting the voters by their favoured site, so the insects need only visit their site to count and be counted.

Experimental evidence suggests that the ants tally votes via the rate at which they encounter other ants while visiting a site [69]. The ants use a quorum that is correlated to the population size of their colony, which makes large colonies less likely to make rash decisions by employing too low a quorum, and small colonies less likely to stagnate due to too high of a quorum [22]. They also vary the value of quorum depending on the urgency of the decision at hand. For example, when a colony moves to a new site simply because it has found a better one, a large quorum will be employed, since the ants have the luxury of making a careful decision. On the other hand, when the ants' nest has become damaged, a lower quorum is used so that a new nest site will be selected rapidly and thus provide the colony with shelter as soon as possible.

The bees also require a quorum to be satisfied before they take flight and head to the selected site, and it appears that they evaluate it at the candidate sites similarly to the ants, although the precise mechanism used to measure the popularity of a favoured site is unknown as of yet [82]. Unlike the ants, the bees relocate to a site as a cohesive group. Individual bees begin to buzz-run throughout their swarm once they have observed quorum at a favoured site, stirring up the mostly dormant bees

[4] There is considerable variation in the calibration of the ants in a given colony [72].

causing them to warm up their flight muscles in preparation for lift-off. A single bee appears to be unable to stir up its swarm on its own, and so lift-off requires several bees to observe quorum. A dormant swarm of bees has a sort of inertia (*i.e.* a swarm at rest tends to remain at rest), so a quorum of bees that have observed quorum (*i.e.* a *meta-quorum*) is required to rouse it. Whether this is an evolved advantage or a fortunate coincidence, it will tend to increase the reliability of the bees' decision-making.

2.2.3 Split Decisions, and Recovery From Them

A split decision occurs when a colony commits to more than one candidate site. Split decisions can be dangerous. At the very least, they will delay the completion of a group decision. At the worst, however, they could cleave a system into two or more pieces, each relocating to a different site. Ultimately, only the group that contained the queen would survive[5]. Furthermore, the portion of the swarm that included the queen would be greatly weakened by its reduction in overall numerical strength. Therefore, split decisions generally should be avoided, and both the ants and the bees have evolved behaviours to recover from them.

When honeybees decide on a new nest site, they already have left their old nest and are exposed to predators and the elements. When more than one site satisfies quorum, the honeybees favouring the different quorum-satisfying sites cooperatively induce the swarm to take flight, but when they attempt to herd the now airborne swarm, they find that they all are attempting to pull it in a different direction, each bee pulling towards the site the it believes has satisfied quorum. Evidently, a swarm is able to detect this mode of failure, and it reforms on a nearby tree or other structure to restart the recruitment process again. During this repeat of the collective deliberation, the bees seem to remember the sites that they had favoured immediately prior to the failed lift-off, as the bees very quickly begin to visit and dance in favour of them with no new search for candidate sites being necessary. This behaviour was described in [80].

The ants solve the problem of split decisions in a different manner. Unlike the bees, the ants do not relocate to a new nest as a cohesive team. Instead, as individual ants independently detect quorum at a favoured site, they each make the decision to switch from tandem-runs to transports. Therefore, their transition from Phase 2 to Phase 3 (Figure 2.3(b)) is gradual, made one ant at a time. In the event that more than one candidate site is believed to satisfy quorum, ants will begin transporting their teammates to each one. It is here that the reverse tandem-runs appear make sense [70]. These bring ants from sites that have satisfied quorum back to the site of the original nest, so that they might help transport additional ants to the site from which they were led. Ultimately, if a tie is to be broken, it will be broken in the manner of attrition: whichever site can deploy the most ants to transport ants from the other sites to it will "win" in the end. Of course, because the ants

[5] In a honeybee colony, as is the case in most social insect societies, the queen is the only fertile individual. All of the workers are sterile sisters, and males are only produced during mating season. Therefore, if a colony were to split into two or more groups, only one of them would be able to reproduce, while the others would tend to slowly die off and not be replaced.

migrate over such short distances, it seems likely that stragglers eventually will find their colony's new home, even if they are not brought there by a teammate. The ants' apparently rash decision-making allows them to capitalize on several sites, and then sort out the resulting mess later on via transporter attrition. Their adaptive quorum [29] would allow them to choose between methodical monolithic decisions or rapid decisions with clean-up as circumstances dictated.

2.3 Summary

The survival of both honeybees and *Temnothorax* ants is tied directly to their abilities to select good sites for their colonies' nests. A colony that cannot find a suitable site, or one that cannot choose a good site over a poor one will be at a competitive disadvantage and will be less likely to pass its genes on to the next generation. Thus, despite the subtle differences in their respective strategies, the behaviours of both species should be held up as examples of robust group decision-making using only local interaction. Their general strategy of search, recruitment-driven solution ranking, and quorum-based commitment is a powerful framework that enables collectives of simple individuals to make intelligent decisions without any form of centralized control.

It is interesting to note that the ants and bees have adapted behaviours that serve other roles in day-to-day colony life to their decision-making behaviour. Tandem-running is common in *Temnothorax* ants to lead nest mates to profitable sources of food, whereas honeybees employ their waggle-dance to this end [31]. When viewed side-by-side, their respective decision-making algorithms are remarkably similar; different implementations of the same three-phase framework: search, deliberate, commit. That both implementations appear to have evolved independently of each other is a further endorsement of the framework's utility. In the next chapter, these insects' decision-making behaviour is adapted for use by dec-MRS, and the behaviours responsible for the different decision-making phases are described in greater detail.

Chapter 3

Collective Decision-Making from Social Insect Behaviour

The nest relocation strategy of honeybees and *Temnothorax* ants described in the last chapter is an elegant collective decision-making behaviour. In this chapter, a general-purpose, three-phase decision-making framework for use by dec-MRS based on the insects' behaviour is presented, and the internal details of the key mechanisms are analyzed.

3.1 A General Purpose Decision-Making Framework

This section outlines the decision-making framework upon which the rest of this thesis focuses. The framework is organized into three basic phases: *searching*, *deliberation*, and *commitment*. Figure 3.1 illustrates how these are assembled to create a cooperative decentralized decision-making behaviour.

The process begins with the recognition of the need for a group decision, which is followed by a search for alternatives. During this searching phase, the individual robots identify potential solutions to the problem that necessitated a cooperative decision. Searching is discussed in Section 3.2.

In the next phase of a decision, the robots deliberate over the alternatives that were identified by the initial search. The deliberation phase of a decision uses positive feedback to compare the different alternatives based on their quality as perceived by the individual robots. The fundamental operation at the core of this decentralized comparison is called recruitment. Robots recruit each other amongst the different alternatives, and the best one becomes apparent as the one favoured by the largest number of robots. This process is detailed in Section 3.3.

During the deliberation phase, as the robots recruit each other, they also estimate the popularity of the solutions that they favour. Using only local, anonymous communication, they are able to compute the consensus in favour of the alternatives, and this is compared to a threshold called the *quorum*. Once the consensus in favour of one of the solutions reaches this threshold, quorum is said to be satisfied, which initiates the final phase of a decision. Decentralized consensus estimation and quorum testing are described in Section 3.4.

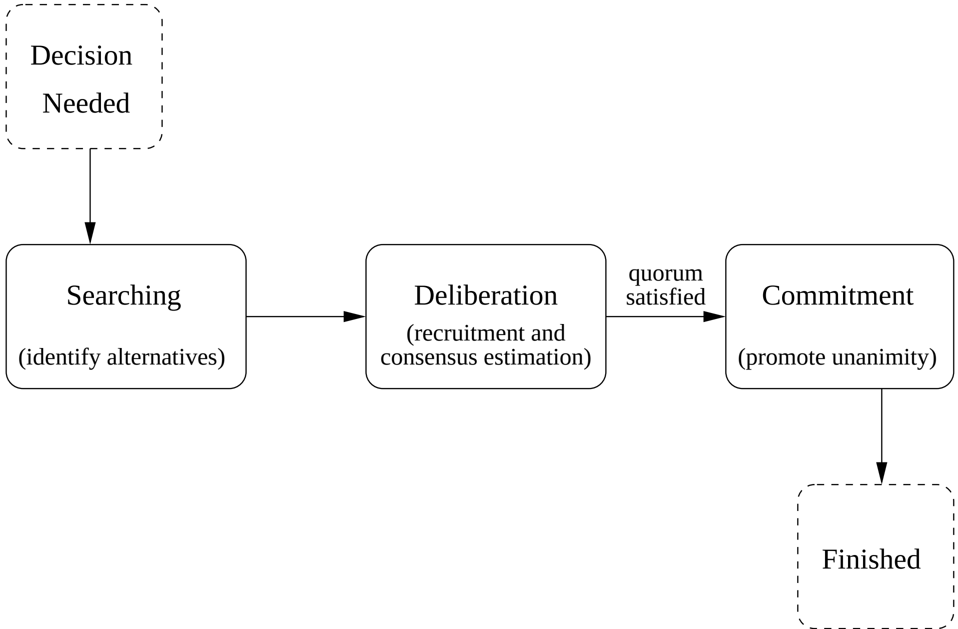


Figure 3.1: The contribution of this thesis is a general-purpose decision-making framework inspired by the nest-site selection behaviour of honeybees and *Temnothorax* ants. When a decision is required in a dec-MRS, its members begin by conducting a decentralized search for alternatives. The search is followed by a period of deliberation in which the best of these alternatives is identified. Through consensus estimation and quorum testing, the robots determine when sufficient deliberation has occurred, after which the final commitment phase promotes the unanimous adoption of the alternative that was identified as best. This chapter describes the mechanisms at the cores of these three phases.

Once alternatives have been found, and the best of these has been identified, it is time to complete the decision unanimously. In a decentralized system, global communication is impractical, and so a strategy that uses local communication to spread the message that a particular alternative should be adopted by every robot is required. This is accomplished in the final phase of a decision: commitment. The commitment phase is described in Section 3.5.

3.2 Searching for Alternatives

An autonomous group decision must be preceded by the recognition of the need for a decision. This recognition could occur in many ways, but in this thesis, it is assumed that the robots all agree initially that a decision is required[1]. The first stage in a decision is to identify candidate solutions to the problem necessitating a group decision. These solutions are referred to as the *alternatives*. The nature of the alternatives sought by the robots depends on the specific decision being made. For example, in the collective relocation problem [57], the problem that the ants and bees tackle, the alternatives are candidate sites in the surrounding environment.

On the other hand, if some intractable optimization problem was the focus of a decision, such as

[1] In the final chapter of this work, decentralized initiation of group decisions is discussed.

the traveling salesman problem, each robot could propose a route using some heuristic method. A variety of routes would be generated if each robot randomly seeded a common heuristic, or perhaps each robot's unique history would result in different routes being proposed. In this way, the group decision could be viewed as a kind of Ant Colony Optimization [21], except that the proposed decision-making framework is better-suited to real robots than the ACO algorithm.

In certain domains, the alternatives might be known *a priori*. Consider a dec-MRS designed for building security. Such a system might be programmed to carry out a fixed set of operations: patrol, follow intruder, respond to alarm, etc. Group decisions made by this system would allow it to dynamically decide which operation best-suited the circumstances at hand.

The deliberation and commitment phases do not depend on the nature of the problem being solved, only the ability of the individual robots to understand it. Therefore, to keep the discussion in this chapter as generally applicable as possible, the term "alternative" will be used unless a specific kind of decision is implied.

3.2.1 The Transition from Searching to Deliberation

The searching phase is followed by the deliberation phase, in which the alternatives that have been found are compared and the best one is identified. However, it is very important to keep in mind that, unless steps are taken to prevent it, the individual robots of a dec-MRS will enter the deliberation phase asynchronously. Each searching robot will begin the process of recruitment once it has found an alternative and evaluated its quality. Those alternatives that are identified earlier will get a head start in the process.

In practice, this head start can be beneficial. The real world is not episodic, and a dynamic environment cannot be "paused" while the robots search for alternatives. Therefore, the value of a solution often *should* take into account the time that was required to identify it. The optimal alternative is of little value if it takes too long to find. In some situations, however, circumstances will permit a longer searching phase, and in such cases the robots should continue to search for for alternatives until some predetermined time. Delaying the onset of deliberation until an agreed time would prevent quickly identified but poor alternatives from acquiring an unwarranted lead in the deliberation phase.

The implementation of this delay could be as simple as "do not begin to deliberate until three o'clock". An alternative found after this time would still be considered by the collective deliberation, but it would have to be sufficiently high in quality so as to overcome its late introduction into the process if it is to be selected in the end.

3.3 Deliberation and Recruitment

Recruitment is a common social behaviour in natural decentralized systems, including social insect colonies. Therefore, it should come as no surprise that it is used in their decision-making, and in turn in the biologically inspired decision-making framework proposed by this thesis. Recruitment enables a decentralized system to compare several alternatives using simple peer-to-peer operations to identify the best one. It does this by amplifying [12] the differences in alternative qualities *autocatalytically*, meaning that it employs a kind of positive feedback. At its root is the emission of a signal by recruiting individuals, and the tendency of those that receive the signal to participate in its emission as well, thus amplifying it[2]. The individuals that emit the recruitment signal are called *recruiters*, and they modulate the strengths of their emissions based on their own opinion of the particular alternative at the focus of each individual's recruitment. The probability of an individual being recruited to a particular alternative increases as the combined strength of the recruitment signal emitted by those that favour it increases relative to the strengths of those emitted by individuals favouring other alternatives. The alternative that tends to be held in the highest opinion will tend to attract additional recruits more rapidly than the other ones.

3.3.1 Indirect Recruitment

The laying of pheromone trails by ants (such as *Lasius niger* [5, 4]) between their nest and a source of food is a particularly well-known recruitment mechanism. Here, the recruiters are foragers that have visited some food source and believe that it is sufficiently high in quality that other ants should be recruited to help exploit it. The recruitment signal is a trail of volatile chemical, called pheromone, laid down by the recruiting ants as they travel between the food source and the nest. The recruiters are able to modulate the strength of the recruitment signal by depositing drops of pheromone more or less frequently as they travel between the nest and the food source. The recruiting strength of a trail is an increasing function of the number of recruiting ants that continually reinforce it and their opinions of the quality of the food source at its end (since this determines how frequently each recruiter deposits pheromone along the trail). As the number of recruiters increases, so does the likelihood of their collectively laid pheromone trail recruiting additional ants, which in turn further increases the strength of the trail via their own pheromone deposits. Because the quality of the food source to which the trail leads also contributes positively to the trail's ability to recruit, the number of ants visiting better food sources will tend to increase more rapidly than the populations visiting poorer ones. Over time, the best food source will attract the majority of the ants. The positive feedback of the recruitment process allows a decentralized system of simple individuals to identify and exploit the best alternative that collectively is known to them.

Pseudo-pheromone trails have been employed with success by artificial decentralized systems in virtual environments. The best example of their use is Ant Colony Optimization [21], which has been used to find shortest paths through telecommunication networks. Chemical trails, however, are poorly-suited to artificial systems operating in real-world environments for the simple reason that current robots are unable to manufacture the necessary chemicals themselves. Furthermore, the

[2]The nature of the recruitment signal and the channel over which it is transmitted must be such that the emissions of multiple individuals will constructively interfere.

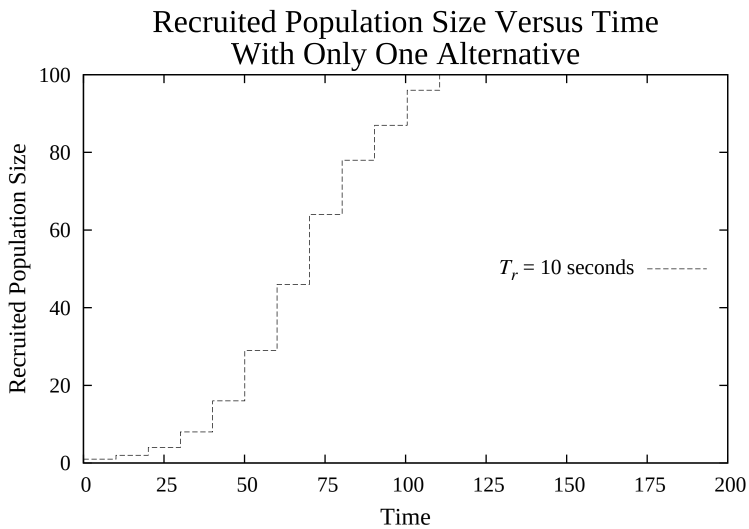


Figure 3.2: This figure illustrates how the positive feedback of direct recruitment increases the size of a population favouring some lone alternative. Initially, all 100 of the robots are idle, except for one that favours the alternative. Every $T_r = 10$ seconds, the robots that favour the alternative each send a randomly selected teammate a recruit-message. Idle robots that receive these messages are recruited, and join the ranks of those that favour the alternative. Initially, the growth of the recruited population is exponential, but as the proportion of robots that are recruited grows, more and more of the recruit-messages are sent to robots that already are recruited, and thus have no effect. This causes the population growth to slow. Because the only stable state for a robot is favouring the alternative, every robot eventually is recruited.

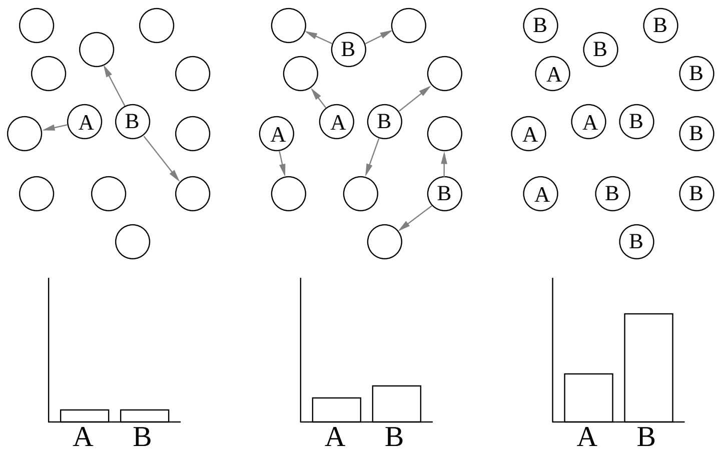
specific chemicals used to lay down trails must be tailored to the specific environment of operation. Finally, trail-following is limited to the comparison of alternatives that are distributed spatially.

3.3.2 Direct Recruitment

There is a much better way to implement autocatalytic recruitment in a a real dec-MRS, though, and the ants and bees described in the last chapter provide it. Instead of indirectly recruiting teammates by the strength of some intermediate signal, the robots could recruit each other directly, by explicitly sending an encountered teammate some message equivalent to “I have recruited you to this specific alternative.” The bees’ waggle dances and the ants’ tandem-runs perform precisely this function. Here, the probability of a recruitment signal successfully recruiting those that receive it would be constant. In other words, the amplitude of each recruitment signal would be fixed. In order to vary the rate of recruitment with an alternative’s perceived quality, the recruiters simply vary the frequency with which they attempt to recruit their teammates. More frequently attempting to recruit teammates would increase the expected rate at which teammates were successfully recruited. In other words, the robots *themselves* are the pheromone trail.

Let us assume that each robot in a dec-MRS will ascribe the same quality to some alternative once it has been recruited to favour it[3], and that only one robot favours the alternative initially.

[3]To be clear on the terminology used in this work, some robots are *recruitable*, and these robots can be *recruited*. Once



(a) (b) (c)

Figure 3.3: This figure illustrates the basic concept of direct recruitment when more than one alternative is known. Individuals recruit teammates to their favoured alternative at a rated that depends on the alternative's quality. Over time, the relative qualities of the known alternatives becomes apparent via the number of robots that favour each one. In this particular example, there are two alternatives, A and B. Robots are represented by circles, and the letters in the circles denote the alternative that they favour. Alternative A is half as good as B, and so the robots that favour A recruit half as frequently as B-favouring robots. Initially (a) each alternative is favoured by a single robot. As recruitment progresses (b), the B recruits more quickly than A. After all of the robots have been recruited (c), B's superiority is clear, since more robots favour it than A.

Each recruiter varies the rate at which it attempts to recruit teammates via the parameter T_r, the inter-recruitment period. T_r is a positive, non-zero, decreasing function of a favoured alternative's quality as it is perceived by the robot that favours it. Consider a dec-MRS in which a single robot favours an alternative at time $t = 0$ and all other robots favour no alternative at all. At time $t = T_r$, the alternative-favouring robot will attempt to recruit one of its teammates, selected at random (assume that every attempt to recruit a teammate will be successful). Because all of its teammates are recruitable, the attempt will succeed and the population favouring the alternative will double to two robots. At time $t = 2T_r$, both robots will select teammates at random and send them recruit-messages. If we assume that the dec-MRS contains many robots, the two robots will be unlikely to randomly select the same teammate or each other, so the recruited population likely will double again to four robots. This exponential growth will continue, but its rate will slow as more and more of the robots are recruited, since there will be fewer robots left to recruit, and so more and more of the recruit-messages will be sent to teammates that already have been recruited to favour the alternative. A simulation of this sort of bounded population growth is illustrated for a 100-robot system by Figure 3.2, in which $T_r = 10$ seconds.

When multiple alternatives are known, their differing qualities as perceived by the individual robots will cause their respective recruiters to compute different values for T_r. The better an alterna-

they have been recruited, they will *favour* the alternative to which they were recruited. A short-hand for this is to say that a robot is *recruited to favour the alternative*.

tive is, the more frequently those robots that favour it will send recruit-messages, and thus the more rapidly will robots be recruited to it. Figure 3.3 depicts this principle graphically. In this example, there are two known alternatives: A and B. B is twice as good as A, and so the robots that favour B recruit twice as often as those that favour A ($T_{r_B} = \frac{1}{2}T_{r_A}$). As a result, the size of the B-favouring population, N_B, will tend to grow at twice the rate of the A-favouring population, N_A. By the third step, B's superiority to A is clearly evident via the greater number of robots recruited to favour it.

Figure 3.3 presents a rather naive illustration of a collective comparison through direct recruitment, since only those robots that had not already been recruited were sent recruit-messages. The size of the unrecruited population is denoted N_o. In the remainder of this section, several direct recruitment strategies are presented under two general classifications: *immutable* and *iterative* recruitment. They differ in how the robots that favour an alternative (*i.e.* those not in N_o) behave when they receive recruit-messages. This simple difference results in markedly different collective behaviours.

Immutable Recruitment

The simplest implementation of direct recruitment is *immutable recruitment*. Under this model, the only recruitable robots are those that do not already favour an alternative. Once a robot has been recruited to favour some alternative, it will continue to favour that alternative indefinitely. This means that the populations that favour the known alternatives never decrease; they either increase or remain constant.

In basic immutable recruitment, each alternative-favouring robot's attempts to recruit a teammate are separated by an interval of time called the *inter-recruitment period*, denoted T_r. T_r is a decreasing function of a robot's opinion of its favoured alternative's quality. The better a robot believes its favoured alternative to be, the lower a value for T_r it will compute, leading to more frequent attempts to recruit by robots that favour better alternatives. A simulation of basic immutable recruitment in which a pair of alternatives of different quality initially are known, each favoured by a single robot, and in which $T_r = \frac{10}{\text{alternative quality}}$, is plotted in Figure 3.4.

Eventually, every robot is recruited to favour one of the two alternatives, and the number of robots recruited to alternative B is greater, indicating that it is the better one. Mathematically, the behaviour of basic immutable recruitment can be represented Equations 3.1-3.3.

$$\dot{N_o} = -\beta_A N_o N_A - \beta_B N_o N_B \quad (3.1)$$

$$\dot{N_A} = \beta_A N_o N_A \quad (3.2)$$

$$\dot{N_B} = \beta_B N_o N_B \quad (3.3)$$

Here, β_A and β_B are positive constants inversely proportional to T_{r_A} and T_{r_B}, respectively. Similarly, N_A and N_B are the total number of robots that favour alternatives A and B. It is clear

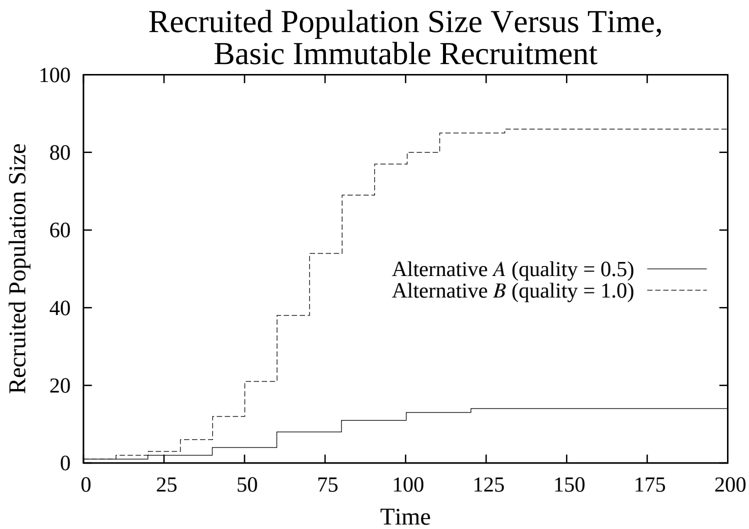


Figure 3.4: In this figure, the growth of two competing populations of recruiters is plotted. Initially, one of the robots favours alternative A, and another favours alternative B. The period of time between a robot's attempts to recruit randomly selected teammates, denoted T_r is inversely proportional to the quality of the alternative that it favours. In this case, robots that favour B recruit twice as often as those that favour A, since B is twice as good as A. Once a robot is recruited to favour a particular solution, it will never change its mind. This is called *immutable recruitment*. Eventually, all of the robots are recruited to favour one of the two alternatives, and the better one can be identified by the greater size of its recruited population.

from inspection that all of the robots eventually will be recruited to one alternative or the other, so the set of system states corresponding to $N_o = 0$ are equilibria. The only non-zero eigenvalue of this system's Jacobian when $N_o = 0$ is equal to $-\beta_A N_A - \beta_B N_B$. Because none of the eigenvalues have positive real parts, a system employing basic immutable recruitment is stable once all of the robots have been recruited to favour an alternative, regardless of the distribution of robots amongst the alternatives.

In the simulation used to generate this Figure 3.4, it was assumed that each robot was able to measure the quality of its favoured alternative precisely. However, if a single robot happened to make an error and overestimate the quality of its favoured alternative, it would recruit teammates more frequently than it should, biasing the outcome of the collective comparison in favour of its alternative. The earlier that such an error were to occur, the more it would bias the outcome of the collective comparison towards the alternative favoured by the erroneous robot. *Temnothorax* ants modify basic immutable recruitment in a manner that addresses this problem. Only an ant's *first* attempt to recruit a teammate is preceded by a delay that is a function of its favoured alternative's quality [70]. Every attempt of a robot to recruit after its first attempt is preceded by a different delay, one that is independent of the quality ascribed by the robot to its favoured alternative. This quality-independent delay, denoted T_{r_o}, is common to all robots. With this simple modification, if an error is made, it will only impact a single attempt to recruit. The error would only be amplified if the robots

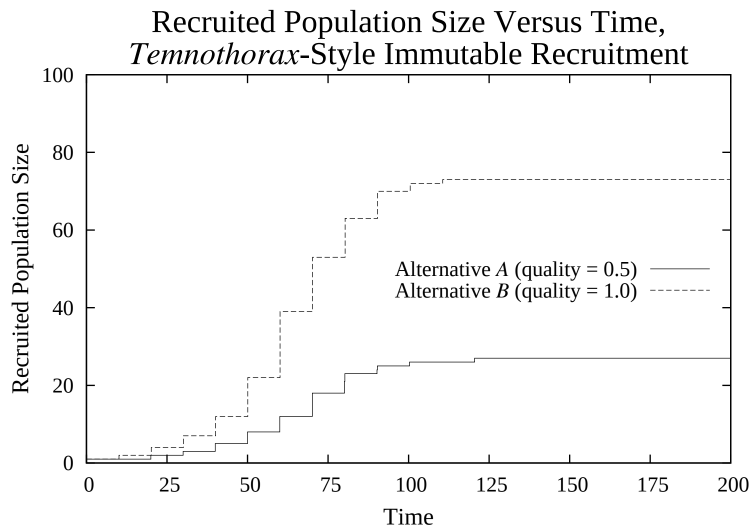


Figure 3.5: *Temnothorax* ants employ a slight variation on the basic immutable recruitment behaviour in their decision-making. Only the delay prior to an individual's very first attempt to recruit a teammate is influenced by its perception of the quality of its alternative. Every subsequent attempt to recruit is preceded by a quality-independent delay, T_{r_o}, which here is 10 seconds. The better alternative still clearly is able to recruit more robots in the end. The advantage *Temnothorax*-style immutable recruitment is that errors made by individual robots when measuring alternative quality will have less of an impact on the overall recruitment behaviour, making it more robust to noisy sensors.

that it recruited also made similarly erroneous evaluations of the alternative's quality. Despite this reduction in alternative-dependent positive feedback, Figure 3.5 illustrates that the better alternative still tends to recruit the more robots than the poorer one by the comparison's end. This system also can be modeled as a set of rate equations, given by Equations 3.4-3.8. Here, β_o is a constant inversely proportional to T_{r_o}. The population of robots that favour each alternative is subdivided into those that have recently been recruited and are delaying prior to their first attempt to recruit, and those recruiting regularly every T_{r_o} seconds. These two groups are denoted by the superscripts W and R (*e.g.* N_A^W and N_A^R), respectively.

$$\dot{N_o} = -\beta_o N_o N_A^R - \beta_o N_o N_B^R \tag{3.4}$$

$$\dot{N_A^W} = \beta_o N_o N_A^R - \beta_A N_A^W \tag{3.5}$$

$$\dot{N_B^W} = \beta_o N_o N_B^R - \beta_B N_B^W \tag{3.6}$$

$$\dot{N_A^R} = \beta_A N_A^W \tag{3.7}$$

$$\dot{N_B^R} = \beta_B N_B^W \tag{3.8}$$

As was the case for basic immutable recruitment, it is clear from an inspection of Equations 3.4-3.8 that all of the robots eventually will be recruited to favour one of the two known alternatives. The non-zero eigenvalues of the Jacobian of this system are $-\beta_o(N_A^R + N_B^R)$, $-\beta_A$, and $-\beta_B$.

Therefore, regardless of how the robots eventually are distributed amongst the known alternatives, the system will be stable once all of the robots have been recruited.

The attraction of immutable recruitment is its simplicity. Its fundamental drawback, though, is that it is immutable: robots cannot be re-recruited. Alternatives that are found early enjoy substantial head-starts in immutable recruitment and, even in the *Temnothorax*-inspired variant, it still tends to be sensitive to individual errors. Furthermore, if many alternatives are found, the final sizes of the populations recruited to each one might all be too small to trigger commitment, leading to stagnation in the deliberation phase of a group decision.

Iterative Recruitment

The second variation on direct recruitment is called *iterative recruitment*. Iterative recruitment is identical to immutable recruitment, except that the individual robots can be recruited more than once. Iterative recruitment was inspired by the recruiting behaviour of honeybees, so a strategy patterned after their behaviour will be presented first. Recall that the honeybees, like the ants, are recruited from an idle state. Unlike the ants, however, the recruited bees eventually return to the idle state. It could be said that the recruiters are "born" out of the idle population, and then "die" by returning to it some time later. It is during their lifetimes as recruiters that they attempt to recruit other individuals. The number of times that a recruiter will randomly select a teammate and send it a recruit-message over its lifetime is equal to the product of its lifetime and the frequency with which it sends recruit messages. Therefore, robots that favour better alternatives will tend to send more recruit-messages than those that favour poorer ones.

This behaviour can be formalized as follows. Only robots that do not favour an alternative are recruitable, and once recruited, robots attempt to recruit their teammates every T_r seconds, which is inversely proportional to alternative quality as before. After a robot has favoured an alternative for T_l seconds, it returns to the idle state, from which it might be recruited again. Mathematically, this is represented by Equations 3.9-3.11, in which $\beta_l \propto \frac{1}{T_l}$ is the rate at which alternative-favouring robots rejoin the unrecruited population. A simulation of honeybee-style iterative recruitment using the same relationship between alternative quality and T_r as was used in the examples of immutable recruitment, with $T_l = 50$ seconds is plotted in Figure 3.6.

$$\dot{N_o} = \beta_l N_A + \beta_l N_B - \beta_A N_o N_A - \beta_B N_o N_B \tag{3.9}$$

$$\dot{N_A} = \beta_A N_o N_A - \beta_l N_A \tag{3.10}$$

$$\dot{N_B} = \beta_B N_o N_B - \beta_l N_B \tag{3.11}$$

Not only does the better alternative attract the greatest recruited population, but the population favouring the lesser alternative is completely eliminated. A steady-state population of robots favouring the remaining alternative also reached that is less than 100% of the robots, which can be explained as follows. Over the course of a recruiter's lifetime, it will will attempt to recruit T_l/T_r times. Each of a recruiter's recruit-messages will be sent to a robot that does not favour an alternative with a probability equal to $\frac{N_o}{N-1}$. Therefore, the expected number of robots that a recruiter will recruit in its lifetime, E_r is given by Equation 3.12.

$$E_r = \frac{T_l}{T_r} \cdot \frac{N_o}{N-1} \tag{3.12}$$

A population of robots favouring a particular alternative will grow in number if $E_r > 1$, and will decrease if $E_r < 1$. When E_r is unity, the population will remain constant at a dynamic equilibrium. Each alternative-favouring population is coupled to every other one through N_o, the number of recruitable robots, since the different alternative-favouring populations compete with each other to recruit these individuals. Equation 3.13 is obtained by rearranging Equation 3.12 and solving for N_o.

$$N_o = \frac{E_i T_r (N-1)}{T_l} \tag{3.13}$$

By setting $E_r = 1$, N_o is the minimum number of idle robots that must be available to an alternative-favouring population to allow it to at least maintain its size. If some of these idle robots are recruited, then E_r will drop below unity and the population will decrease in size. This is precisely how the population of robots that favours the best alternative drives the others to extinction. They are able to swell their numbers even after sizes of the populations favouring lesser solutions have plateaued. When they recruit additional robots, N_o decreases and so E_r for the poorer alternative-favouring populations will reduce to less than unity, in turn decreasing those populations and eventually eliminating them altogether.

Given sufficient time, iterative recruitment is characterized by its ability to completely discard all but one alternative, and the remaining one will tend to be the best of those that was found. There exist a pair of subtle problems with the honeybee-inspired approach. First, N_o will always be greater than zero, so at equilibrium, there will always be some idle robots. The size of this unrecruited population depends directly on the remaining alternative's absolute quality, so if the best known alternative is relatively poor, the steady-state population favouring it (equal to $N - N_o$) will be relatively small. If this happened to be too small to satisfy quorum, stagnation would follow[4]. This means that T_l and the mapping of alternative quality to T_r must be tuned to specific decisions, including the expected qualities of the alternatives.

Second, honeybee iterative recruitment is poor at comparing multiple alternatives of equal quality. Consider the situation in which two alternatives are equally good. The idle population at equilibrium for both alternatives will be the same (since both use the same value of T_r), so neither one will be able to push the other to extinction by recruiting a few additional robots and reducing the

[4]It might seem odd that both recruitment mechanisms used by natural systems permit stagnation. Both the ants and the bees employ low quorums, have tuned their behaviours to their specific environments, and they have evolved additional behaviours to break deadlocks.

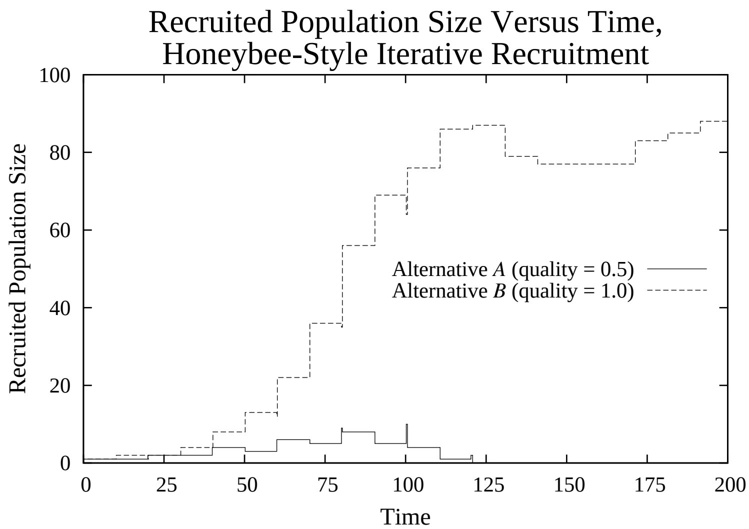


Figure 3.6: Iterative recruitment differs from immutable recruitment (Figures 3.4 and 3.5) in that the robots can be recruited to more than one alternative throughout the process. This is the strategy used by the honeybees. When a robot is recruited, it favours the alternative of its recruiter, and attempts to recruit others, but eventually it will return to the idle state, from which it might get recruited again. Therefore, the alternatives compete against each other for the idle robots. Robots are more likely to be recruited to the better alternatives, so the populations favouring the lesser alternatives eventually are wiped out. Ultimately, only one alternative will remain. However, because robots favouring an alternative reenter the idle state at a finite, non-zero rate, some constant proportion of a system's robots always will be in the idle state, preventing 100% unanimity from being achieved.

other's E_r to less than unity. Only through stochasticity might one of the populations extinguish the other, and this will become increasingly unlikely as the overall population size of the dec-MRS increases. Identical alternatives might seem improbable, but keep in mind that two alternatives will be *perceived* by the robots to be equally good if they differ only in some manner undetectable by them. When two alternatives are equally good, $\beta_A = \beta_B = \beta$ in Equations 3.9-3.11, and at steady-state, $N_o = \frac{\beta_l}{\beta}$. The only non-zero eigenvalue of the honeybee Jacobian at this equilibrium is $-\beta(N_a + N_B)$. Since it is less than zero, the system is stable, so neither alternative will be able to extinguish the other through any means other than random chance.

The problem of a steady-state recruited population less than 100% of the robots is easily be overcome by eliminating the tendency of recruited robots to return to the idle state (*i.e.* make T_l infinite), and enabling robots to recruit alternative-favouring teammates directly. However, the tie-breaking problem remains. This recruitment strategy, called *basic iterative recruitment* is described by Equations 3.14-3.16.

$$\dot{N_o} = -\beta_A NoN_A - \beta_B NoN_B \tag{3.14}$$

$$\dot{N_A} = \beta_A N_A(N_o + N_B) - \beta_B N_A N_B \tag{3.15}$$

$$\dot{N_B} = \beta_B N_B(N_o + N_B) - \beta_A N_B N_A \tag{3.16}$$

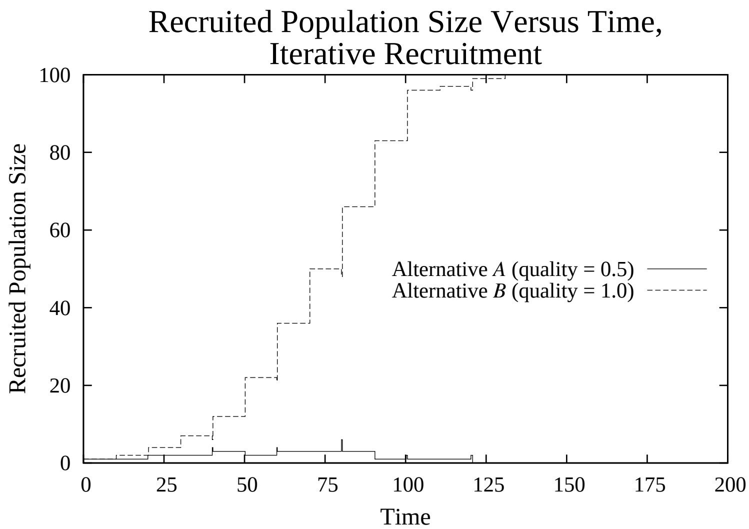


Figure 3.7: The main problem with the honeybee approach to iterative recruitment is that the steady-state size of the population favouring the best alternative directly depends on the alternative's absolute quality as perceived by the robots. This is because the only stable state in that approach is the idle state. Instead of recruiting robots only from the idle state, regular iterative recruitment allows robots to be recruited directly from favouring one alternative to favour another. This means that favouring an alternative is a stable state, since robots will continue to favour a alternative until they receive a recruit-message from a teammate favouring a different alternative. Similarly to honeybee iterative recruitment, all but the best alternative will be forgotten. Unlike honeybee iterative recruitment, however, the steady-state size of the population favouring the remaining alternative will always be 100% of the robots.

Clearly, the number of unrecruited robots eventually will decrease to zero, since $\dot{N}_o$ is always negative when at least one alternative is known, but the only potentially non-zero eigenvalue when $N_o = 0$ is $(N_B - N_A)(\beta_A - \beta_B)$, which is zero when $\beta_A = \beta_B$. Again, unanimity will only be achieved through random perturbations in the recruitment process.

The tie-breaking problem can be solved by tying the value β_i not only to an alternative's quality, but it's popularity, too. That is, β_i in the basic iterative recruitment strategy is replaced by $\beta_i \gamma N_i$ here, where γ is some positive constant. This new strategy, referred to in this work simply as *iterative recruitment* is described by Equations 3.17-3.19.

$$\dot{N}_o = -\beta_A \gamma No N_A^2 - \beta_B \gamma No N_B^2 \tag{3.17}$$

$$\dot{N}_A = \beta_A \gamma N_A^2 (N_o + N_B) - \beta_B \gamma N_A N_B^2 \tag{3.18}$$

$$\dot{N}_B = \beta_B \gamma N_B^2 (N_o + N_B) - \beta_A \gamma N_A^2 N_B \tag{3.19}$$

A simulation of iterative recruitment is shown in Figure 3.7. $\beta_A = \beta_B = \beta$, $N_A = N_B$,$N_o = 0$ is the equilibrium point in this system when two equal alternatives are being compared. The eigenvalues of the Jacobian of this system at this tie are $-\frac{1}{2}\beta\gamma N^2$, 0, and $\frac{1}{2}\beta\gamma N^2$. Because one of the eigenvalues is positive, the equilibrium unstable. Randomness in the rate at which the robots

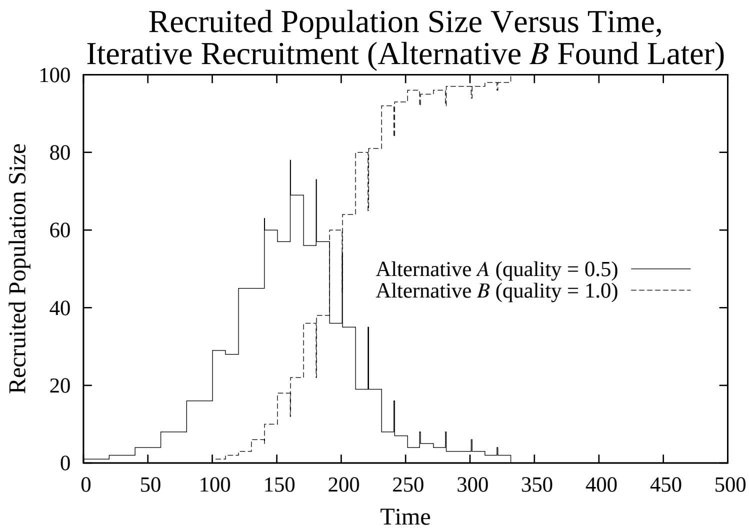


Figure 3.8: Another great advantage of iterative recruitment over immutable recruitment is that a better alternative found later on in the process can still obtain the unanimous support of a dec-MRS if it is sufficiently good. In this figure, alternative A is found at $t = 0$, whereas alternative B is found at time $t = 100$. Alternative B is twice as good as alternative A, so robots that favour it to recruit twice as frequently as those that favour A. The population favouring B grows rapidly, eventually recruiting every robot in the system.

actually encounter each other will push the system off of the equilibrium, at which point the rate of recruitment for the more popular alternative will increase and that of the less popular one will decrease, thereby pushing the system away from the equilibrium point and towards the unanimous adoption of a single alternative.

Although the robots do not actually know how many of their teammates agree with them and thus cannot implement the system described by Equations 3.17-3.19 directly, a simple behaviour-based strategy permits it to be implemented in a more emergent manner as follows. Robots respond to recruit-messages, telling the recruiter whether or not the recipient of the message already favoured the specified alternative *a priori*. When a robot is informed that attempted to recruit a teammate that already agreed with it, it will become *frustrated*. Instead of scheduling its next attempt to recruit a teammate T_r seconds into the future, a frustrated robot will attempt to recruit the very robot that it encounters. It will remain frustrated until it successfully recruits a teammate, at which point it will cease to be frustrated and instead attempt to recruit teammates at a rate governed by T_r.

If it assumed that the T_r is greater than the expected inter-robot encounter period, T_0, then becoming frustrated increases up the rate at which a robot attempts to recruit teammates. The average rate of recruitment by the population of robots that favours a particular alternative increases with the proportion of them that are frustrated. This in turn increases with the popularity of their favoured alternative, since robots that favour more popular alternatives are more likely to encounter agreeing teammates and become frustrated than robots that favour less popular alternatives.

Not only does iterative recruitment lead to the unanimous favouring of a single alternative, but a sufficiently good alternative discovered after deliberation already has begun might still emerge as the one identified as best by the robots. This property is illustrated by Figure 3.8. The later that an alternative is found, the better it would have to be to overcome the head-start enjoyed by those alternatives that were found earlier.

3.4 Quorum Testing and Consensus Estimation

In a collective decision, the quorum is the minimum support that an alternative must attract in order for it to induce commitment and complete the decision. Quorum is said to be *satisfied* by an alternative once the consensus in favour of it meets or exceeds quorum. *Quorum testing* is the process of measuring consensus and comparing it to the quorum. Thus, at its core, quorum testing is a parameter estimation and thresholding operation.

During the deliberation phase of a collective decision, the number of robots that favour each alternative will evolve. Better alternatives tend to increase their recruiting corps whereas poorer alternatives tend to lose support. It is quorum testing that ends the deliberation phase, signaling that the best alternative has been identified. The satisfaction of quorum triggers the commitment phase, in which the quorum-satisfying alternative is unanimously adopted by the entire dec-MRS. The ideal quorum test would be instantaneous and accurate. Speed and accuracy are competing interests, however, and so a real-world quorum test must strike an appropriate balance between them. Recall also that the individual robots will know of at most a single alternative: the one that they favour. Therefore, an individual robot can test quorum only for the alternative that it favours. As a result, the total number of alternatives being collectively compared does not increase the computational complexity of consensus estimation and quorum testing for the individual robots. It also reduces the likelihood of a robot making an error and concluding that a less popular alternative has satisfied quorum, since fewer robots test quorum for the less popular alternatives, simply because fewer robots favour them.

3.4.1 Off-Swarm Consensus Estimation

In certain domains, the nature of the decision being made by a dec-MRS can be exploited to simplify quorum testing. When alternatives are spatially distributed, as they are in the collective relocation problem, quorum could be tested for an alternative by comparing the number of robots that visit its location to a threshold.

For this strategy to succeed, the robots that favour an alternative would have to visit its site regularly to assess the visiting population. Each robot that favoured a particular alternative therefore would spend some proportion of its time at its favoured alternative's location. As an alternative became more popular, the size of the average visiting population would increase. Quorum in this case would be an absolute number of robots bounded by the population size of the dec-MRS.

Because only some of each robot's time is spent visiting its alternative, the average visiting population size is unlikely to exceed a fraction of the total number of robots in a system, even if every robot favoured the same alternative. Therefore, quorum would have to be relatively small in order to have a realistic chance of one of the robots observing a quorum of its teammates when it visited the site of its alternative. A low quorum would increase the probability of a robot that favoured a poorer alternative observing quorum first. This might occur if all of the robots that favoured a particular poor alternative happened to visit its location at the same time.

The advantage of this *off-swarm* approach is that consensus estimation and quorum testing do not require explicit communication, since only passive observation of teammates is required. Furthermore, the synonym problem is avoided. Alternatives are represented by locations in a shared environment, so two robots that visit the same alternative's location know that they both favour the same alternative, regardless of the label given to the alternative by them. To paraphrase Shakespeare, *a particular alternative by any other name should smell just as sweet* [83].

3.4.2 Consensus Estimation by Explicit Opinion Sampling On-Swarm

Off-swarm consensus estimation exploits a specific decision-making domain and environment. It assumes that the candidate alternatives have unambiguous mappings to distinct geographic locations. The advantage of the off-swarm approach is that the communicative capabilities of the individual robots in a dec-MRS could be very limited, yet they still could take advantage of it. However, the articulative aspects of inter-robot communication can be considered a solved problem in many ways. The problem that confronts a robot is not *how* to transmit a message, but *what* a message should contain. In this section, a scalable approach to consensus estimation is presented that takes advantage of explicit inter-robot communication, and two different implementations are provided.

Consider a dec-MRS that is composed of N robots, N_a of which favour some particular alternative in common, and the remaining $N - N_a$ favour different alternatives (or none at all). If one of the N_a robots randomly selects a teammate and asks it whether or not it also favours the same alternative, the probability of the queried robot responding affirmatively would be $\frac{N_a-1}{N-1}$, since $N_a - 1$ of the querying robot's $N - 1$ teammates also favour its alternative. This fraction is called the *apparent consensus*, denoted C_a. It is the consensus apparent to a robot when it does not include its own opinion. Without knowing the number of teammates in its dec-MRS, a robot cannot include its own opinion, since it would be unable to compute an appropriate weight so that it equitably could be included in is measurement of consensus (*i.e.* it would tend to over- or under-value its own opinion). Note that apparent consensus is strictly less than *true consensus*, $\frac{N_a}{N}$, denoted C_t. However, this difference becomes insignificant as N increases.

The response of a queried robot is called a $vote\text{-}message$, and a vote-message is either "yes" or "no", indicating whether or not the queried robot favours the same alternative as the querying robot[5].

[5]It does not matter if the asking robot asks "Do you favour alternative X?", to which the queried robot would respond

Over time, robots favouring an alternative will receive a sequence of vote-messages as they query the teammates that they encounter. Let $vote_i$ denote the i^{th} vote-message received by a particular robot, and the function $\gamma(vote_i)$ quantify this vote according to Equation 3.20.

$$\gamma(vote_i) = \begin{cases} 1 & : \quad vote_i = \text{“yes”} \\ 0 & : \quad \text{otherwise} \end{cases} \tag{3.20}$$

If a robot receives receives n vote-messages, obtained from randomly selected teammates, it can estimate C_a for its favoured alternative from them using Equation 3.21. Just like when people's opinions are gathered for a survey of public opinion, it is important that the vote-messages that each robot uses to estimate C_a are collected from randomly selected teammates. Since it is assumed that the individual robots are unable to identify the specific teammates that they encounter, it falls the the stochastic nature of their interactions to generate a random sample of teammate opinions. The term *well-stirred* [62] is used to characterize such interaction. In a well-stirred system, the identity of the robot encountered by an individual is equally likely to be any of its teammates, and each encounter is a statistically independent event. In a densely populated dec-MRS, the independence of each encounter might not be a valid assumption, but the well-stirred assumption could be made to hold by counting only every i^{th} encounter with a teammate, or only those separated by a sufficiently long period of time. In this way, additional opportunity would be given for the system to stir itself between the querying of teammates.

$$\tilde{C_a} = \frac{1}{n}\sum_{i=1}^{n}\gamma(vote_i) \tag{3.21}$$

The symbol $\tilde{C_a}$ indicates that the quantity on the left side of Equation 3.21 is an estimate of C_a, not the precise apparent consensus itself. Throughout this discussion and derivation, it must be kept in mind that the individuals do not have global knowledge of dec-MRS state. Furthermore, each robot computes $\tilde{C_a}$ only for the alternatives that it favours. Therefore, a specific value of $\tilde{C_a}$ corresponds only to the alternative favoured by the robot that computed it.

Digital Consensus Estimation

A simple way to implement consensus estimation would be to provide each robot with an n-element queue, into which quantized teammate opinions would be inserted. With every insertion of a new teammate opinion, the $(n+1)^{th}$ opinion would be discarded, and $\tilde{C_a}$ at any given moment would be the mean value of the opinions in the queue. This concept is referred to in this work as *digital consensus estimation*, and is illustrated by Figure 3.9.

How n should be selected? The greater n is made, the more opinions will be included in the computation of $\tilde{C_a}$, and thus the more accurate the estimate will tend to be. However, as n is increased, $\tilde{C_a}$ depends more upon older opinions, and thus tracks changes in apparent consensus

"yes" or "no", or if the asking robot asks "What alternative do you favour?", and then interprets the queried robot's response as a "yes" or a "no".

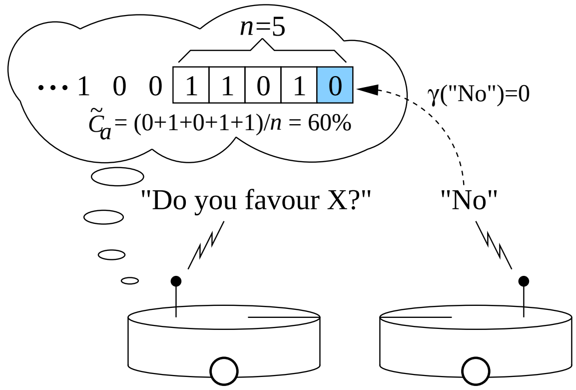


Figure 3.9: This illustration depicts digital consensus estimation. The robot on the left is computing an estimate of apparent consensus for some alternative X that it favours. Upon encountering a teammate, it asks it if it also favours X. The teammate does not, and its response "No" is converted to the numerical value 0 by the quantifying function $\gamma(vote_i)$. Some earlier collected opinions were "Yes", and these were assigned a value of 1. In this example, $n = 5$, so only the five most recently received quantified opinions are averaged to compute $\tilde{C}_a$, which in this case is 60%. The previous three values for $\tilde{C}_a$ by the robot on the left were 60%, 40% and 60%. By increasing number of opinions used to compute $\tilde{C}_a$, the accuracy of each estimate increases but requires more teammate opinions and thus more time to compute.

more slowly. Keep in mind that each estimate of apparent consensus is compared to a threshold, and that a behavioural change will occur only if $\tilde{C}_a$ is greater than or equal to the threshold. A compromise must be found between the speed and accuracy of quorum testing.

Because vote-messages are assigned a value of either 1 or 0, each one is a Bernoulli random variable in which $P(vote_i = 1)$ is equal to C_a. Let the threshold to which a robot compares $\tilde{C}_a$ be denoted by Q, where $Q \in [0, 1]$. A robot concludes that quorum has been satisfied once $\tilde{C}_a \geq Q$, which will occur when at least $\lceil nQ \rceil$ of the n most recently received vote-messages are affirmative. The probability of receiving i affirmative votes in a sequence of n follows the binomial distribution. Thus the overall probability of a particular set of n vote-messages suggesting that quorum has been satisfied is the sum of the binomial distribution over the range $i \in [\lceil nQ \rceil, n]$, which is given by Equation 3.22

$$P(\tilde{C}_a \geq Q) = \sum_{i=\lceil nQ \rceil}^{n} \binom{n}{i} (C_a)^i (1 - C_a)^{n-i} \tag{3.22}$$

A *false positive* quorum test is said to occur when $\tilde{C}_a \geq Q$ and $C_a <$ quorum. That is, when a robot erroneously believes that the apparent consensus satisfies quorum when in fact it does not. Equation 3.23 provides this mathematically.

$$P(\tilde{C}_a \geq Q \wedge C_a < \text{quorum}) = \sum_{C_a \in [0, quorum)} P(\tilde{C}_a \geq Q) \tag{3.23}$$

Figure 3.10 plots Equation 3.22 for different values of n and Q. Equation 3.23 can be interpreted graphically as the area under one of these curves to the left of $C_a =$ quorum. Note from Figure

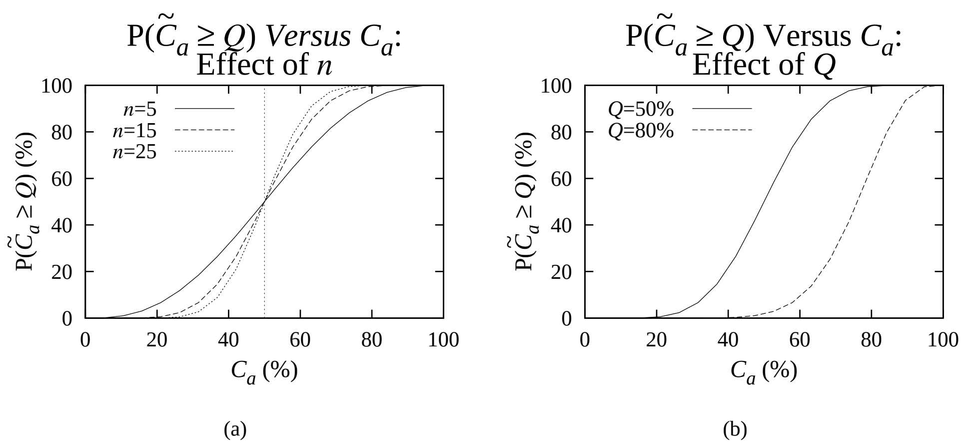


Figure 3.10: The graphs in this figure illustrate how the parameters n and Q (the number of teammate opinions used to compute $\tilde{C}_a$ and the threshold to which $\tilde{C}_a$ is compared to in order to test quorum) affect the probability of a robot believing that quorum is satisfied versus the actual value of apparent consensus, C_a. Increasing n makes the curve more step-like, decreasing the likelihood of a robot prematurely committing. Increasing Q does not significantly change the shape of the curve, instead shifting it to the right.

3.10(a) that this area decreases as n is increased, as the curve approaches the shape of a step at $C_a = Q$. However, the area to the left of $C_a =$ quorum will never be nonzero, and a very large value of n would be necessary to reduce it substantially. This would slow down consensus estimation significantly.

Changing the value of Q while keeping n constant does not significantly alter the shape of the curves, instead shifting them to the right as Q is increased. This is shown by Figure 3.10(b). Quorum and Q, however, need not be the same. Since the goal is to prevent a robot from making false positive errors, the desired quorum should be specified, and then the parameters n and Q selected strike an acceptable balance between the desire for speed and the likelihood of a false positive test occurring. For example, $n = 15, Q = 80\%$ appears to be a good choice of parameters to test a quorum of 50%.

The commitment phase of a decision begins once one of the robots believes that quorum has been satisfied. The probability of commitment occurring is equal to one minus the probability of none of the N_a robots testing quorum believing that quorum has been satisfied:

$$\begin{aligned} P(commitment) &= 1 - (1 - P(\tilde{C}_a \geq Q))^{N_a} \\ &= 1 - \left[\sum_{i=\lceil nQ \rceil}^{n} \binom{n}{i} (C_a)^i (1 - C_a)^{n-i} \right]^{(\lfloor C_a(N-1) \rfloor + 1)} \end{aligned} \tag{3.24}$$

The concern is that, whereas the probability of a particular robot concluding that quorum is satisfied is independent of the number of robots that compose a dec-MRS, the probability of at least one of the robots in a system committing is not, indicated by the exponent in the second line of Equation 3.24. Although the probability of at least one robot prematurely believing that quorum has been satisfied does increase with N, the behaviour of a given configuration is quite stable and

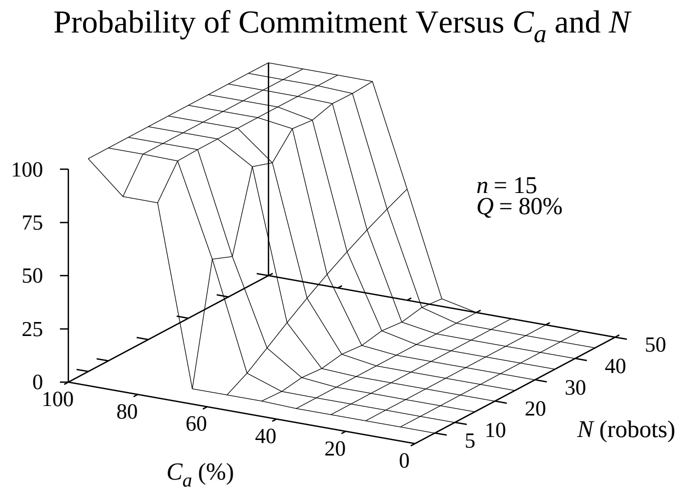


Figure 3.11: The commitment phase begins once one of the robots estimating consensus believes that quorum has been satisfied. As C_a increases, there by definition will be more robots estimating apparent consensus, and so the chance of one of them making an error by overestimating C_a and prematurely triggering commitment phase also will increase. This graph plots the probability of at least one robot believing that quorum has been satisfied as a function of C_a and the population size, N. Although the behaviour of quorum testing does depend somewhat on the population size, the decrease in reliability as N increases is minimal.

predictable, as illustrated by Figure 3.24. In particular, the data plotted in this figure shows that the configuration $n = 25$, $Q = 80\%$ is a good test for a quorum of 50% when the population is 15 robots or less, and remains a good test for a quorum of 40% for populations up to at least 50 robots.

Analog Consensus Estimation

A robot also could estimate the apparent consensus amongst its teammates using an analog implementation. At first, an analog solution to a computational problem might seem quaint or out of date, conjuring images of Grey Walter's tortoises [89]. However, a single-purpose analog circuit often is more compact and more efficient than its digital equivalent. A concise analog implementation of a quorum test would be very useful in micro- or nano-scale robots. Just as bacteria test quorum to select their individual behaviors [54, 90], so could the members of a microscopic dec-MRS. In a system of such simple robots, unable to make accurate observations of their environment on their own, the pooling of individual opinions and quorum testing could be of critical importance. The contributions of this section build upon those of [62].

The analog quorum test uses a pair of exponentially decaying variables to estimate apparent consensus. These are called the *quorum index* and the *kin index*, denoted by $q(t)$ and $k(t)$, respectively. When such a variable is incremented by a constant amount at a regular interval, a sawtoothed wave is formed, illustrated by Figure 3.12. The peak value of this wave will increase over time, reaching an equilibrium value that is determined by the time constant of exponential decay (τ), The amount

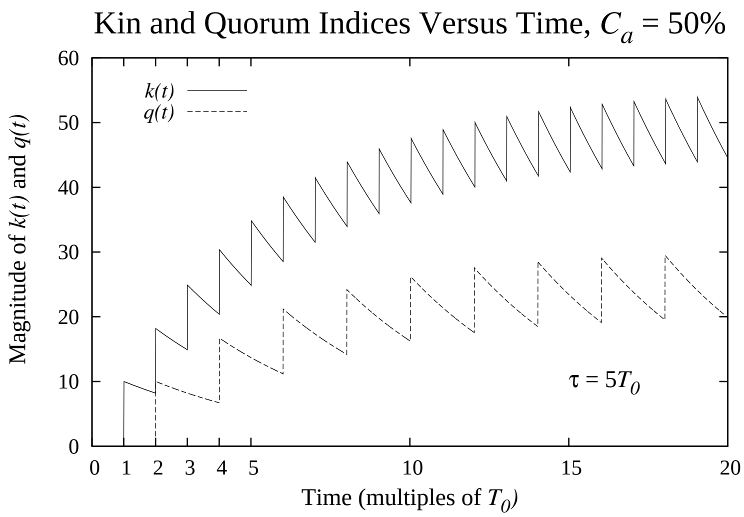


Figure 3.12: At the heart of analog consensus estimation are a pair of exponentially decaying indices. These are called the kin and quorum indices, denoted $q(t)$ and $k(t)$, respectively. Periodically, these are incremented by a constant amount. $k(t)$ is incremented whenever any teammate is encountered, whereas $q(t)$ is incremented only when an agreeing teammate is encountered. Both curves adopt a sawtoothed shape, reaching equilibria determined by the frequency with which they are incremented. In this figure, the quorum index is incremented only half as often as the kin index, and its peak equilibrium is half that of the kin index. In general, the peak equilibrium value of the quorum index will be equal to that of the kin index scaled by the apparent consensus, and so their ratio computes $\tilde{C}_a$.

by which the variable is incremented, and the period of time between the regular increments. The quorum index is incremented every time that an agreeing teammate is encountered, whereas the kin index is incremented with every teammate encounter. Both indices are incremented by the same amount, Δ. It will now be shown that $\frac{q(t)}{k(t)} \approx Ca$.

Assume that, in a well-stirred dec-MRS, a robot will encounter one of its teammates every T_0 seconds. Because C_a of a robot's teammates agree with it, the time between encounters with agreeing teammates $T_a = T_0/C_a$. Therefore, $q(t_0 + T_a) = q(t_0)e^{-T_a/\tau}$ immediately before Δ is added to the index. At equilibrium, $q(t_0) = q(t_0 + T_a) + \Delta$, and $q(t_0)$ will be the peak equilibrium value of the quorum index, denoted by q_{equ}. This quantity can be expressed as a function of apparent consensus as shown by Equation 3.25.

$$
\begin{aligned}
q_{equ}(C_a) &= \frac{\Delta}{1-e^{\frac{-T_a}{\tau}}} \\
&= \frac{\Delta}{1-e^{\frac{-T_0}{\tau C_a}}} \qquad (3.25) \\
&\approx \Delta \frac{\tau C_a}{T_0}
\end{aligned}
$$

The final approximation in Equation 3.25 is obtained with the following limit: $\lim_{f(x)\to 0} e^{f(x)} = 1 + f(x)$. Therefore, when $\frac{-T_0}{\tau C_a}$ is small, $e^{\frac{-T_0}{\tau C_a}}$ is well-approximated by $1 - \frac{T_0}{\tau C_a}$. Substituting this

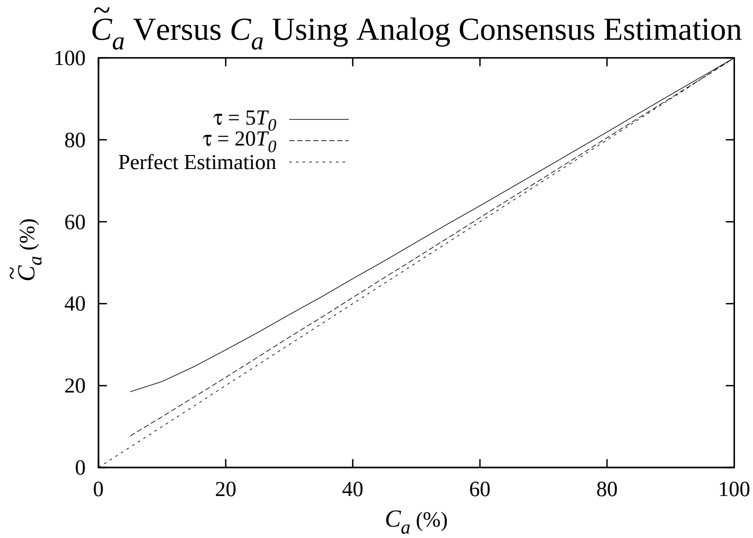


Figure 3.13: The ratio of the peak equilibrium values of the quorum and kin indices closely approximates the apparent consensus. This figure plots this ratio, using two different values for τ, which specifies the rate at which the indices decay. Increasing τ increases the accuracy of the estimate, but also will increase the time required to make the estimate.

into the second line of Equation 3.25 yields the third.

The kin index, $k(t)$ also reaches a peak equilibrium value, k_{equ}, but this is independent of the apparent consensus. By following the same steps as in the case of the quorum index, the result of Equation 3.26 is obtained.

$$\begin{aligned} k_{equ} &= \frac{\Delta}{1-e^{\frac{-T_0}{\tau}}} \\ &\approx \Delta\frac{\tau}{T_0} \end{aligned} \tag{3.26}$$

It should now be clear that the ratio the two exponentially decaying indices estimates the apparent consensus, illustrated by Equation 3.27.

$$\begin{aligned} \frac{q_{p_{equ}}(C_a)}{k_{p_{equ}}} &\approx \frac{\Delta\frac{\tau C_a}{T_0}}{\Delta\frac{\tau}{T_0}} \\ &\approx C_a \\ &= \tilde{C}_a \end{aligned} \tag{3.27}$$

All that remains is to choose the time constant τ, which determines the rate at which the kin and quorum indices decay. As was the case when digital consensus estimation was examined, there is a trade-off between the accuracy of the individual robots' estimates, $\tilde{C}_a$, and the time required to compute them. It is the value of τ relative to T_0 that determines what kind of balance is struck between these competing concerns.

Recall that the linear approximation of the peak equilibrium values of $k_{p_{equ}}$ and $q_{p_{equ}}$ assumed that the exponent to which the base of the natural logarithm was raised was small. This is achieved when $\tau \gg T_0$. The greater τ is, the better the approximation will hold. Because of this, increasing

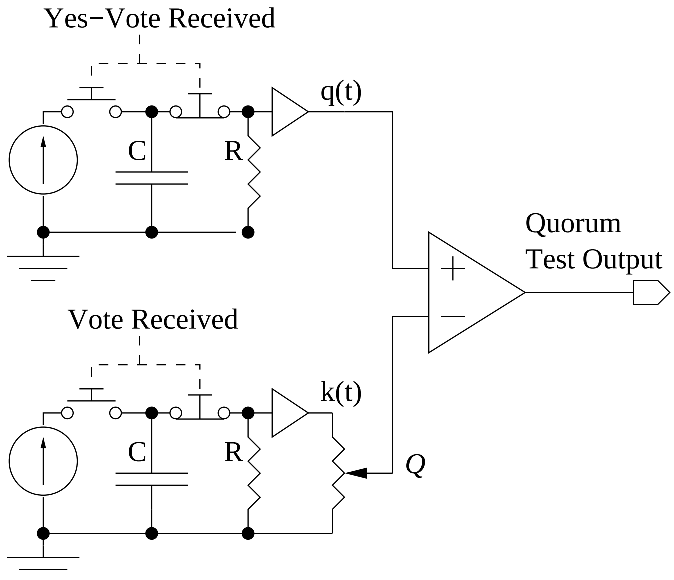


Figure 3.14: This analog circuit permits even the smallest and simplest of robots to estimate apparent consensus and use it to test quorum. The upper and lower RC-circuits produce sawtoothed waves denoted q(t) and k(t), the DC peak values of which are proportional to $N_a - 1$ and $N - 1$, respectively. Quorum is tested by comparing $q(t)$ to $Qk(t)$ via the comparator on the right. When $q(t) \geq Qk(t)$, the comparator switches on, signaling that quorum is satisfied.

τ improves the accuracy of consensus estimation. The more slowly the kin and quorum indices decay, however, the longer it will take for them to reach their equilibrium values, increasing the time required to compute $\tilde{C}_a$. Refer to Figure 3.13. The curve corresponding to the higher value of τ approximates the actual value of C_a better that the one using a lower τ. Furthermore, note that the errors in consensus estimation are greater when C_a is lower, and that $\tilde{C}_a > C_a$.

Figure 3.14 presents a simple implementation of an analog quorum test, illustrating how easily this powerful concept could be integrated into even the simplest of robots. In this derivation, it was assumed that the robots of a well-stirred dec-MRS will encounter their teammates at some regular interval T_0. It is important to note that T_0 is *not* an exact value, but a random variable with a distribution. T_0 should be chosen to be as low as possible, so that a robot would be unlikely experience two teammate encounters separated by less than T_0. Another option would be to add a timer to the circuit of 3.14 that prevented the kin and quorum indices from being incremented more frequently than T_0, ignoring any encounters that occurred before this encounter-timer elapsed.

The time-dependence of analog consensus estimation can introduce errors in $\tilde{C}_a$ not present when it is estimated digitally. For example, if a robot were to get lost and stop encountering its teammates altogether, the kin index would decrease (measured absolutely) more rapidly than the quorum index, since the former always is greater than the latter and exponential decay is proportional to instantaneous magnitude. This in turn would increase the ratio $q(t)/k(t)$, resulting in an overestimate of C_a. Eventually, a robot would believe that 100% consensus existed. If this hypothet-

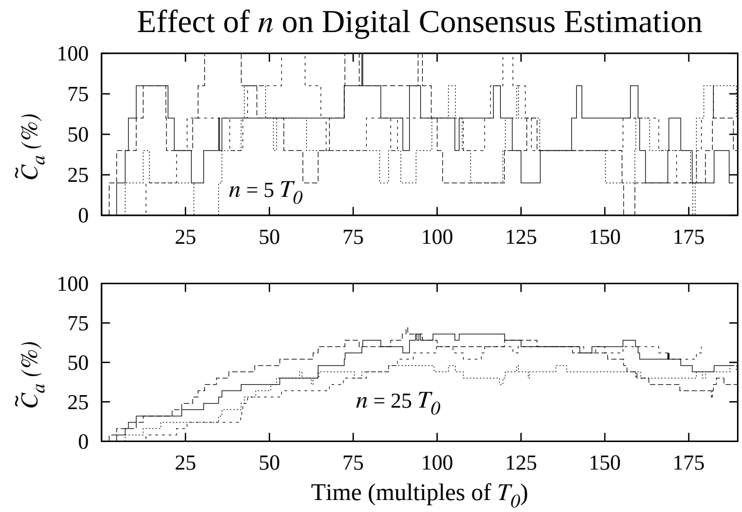


Figure 3.15: This figure plots four robots' estimates of apparent consensus computed with two different values of n in a system where $C_a = 50\%$. Both graphs use the same sequences of robot interactions; only n differs. When n is small, $\tilde{C}_a$ can change rapidly, but there is a substantial amount of noise in the measurements. Increasing n greatly increases the accuracy of the measurements, but they take longer to reach a steady-state value.

ically lost robot were to rejoin its teammates, and the first one that it encountered was an agreeing teammate, the robot again would believe that $C_a = 100\%$, since both indices would have decayed to near zero while it was lost and both would have Δ added to them by this encounter, producing a unity ratio once again. To eliminate these two sources of error, robots should only consider an analog estimate of apparent consensus valid only immediately after a teammate has been encountered and only if $k(t)$ has reached equilibrium. This is a significant contrast to digital consensus estimation, the accuracy of which is independent of the time that elapses between each reception of a teammate opinion.

3.4.3 Real-World Performance of Anonymous Digital and Analog Consensus Estimation

Thus far the digital and analog strategies to estimate apparent consensus have been described, and predictions have been made regarding their performance. In this section, simulated results are presented that verify these claims.

In order to collect data, the TeamBots simulator [85] was used to implement a 15-robot MRS in which the robots randomly wandered about a circular environment, reorienting to random headings when an obstacle (either a teammate or the outer wall) was encountered. Eight of the robots agreed with each other, and the remaining seven robots agreed with nobody. This configuration meant that C_a was 50% for the first eight robots and 0% for the remaining ones. The same generic log of robot interactions was used to generate the data for each of the following plots[6], so each data point

[6]The details of this technique are presented in greater detail in Chapter 4

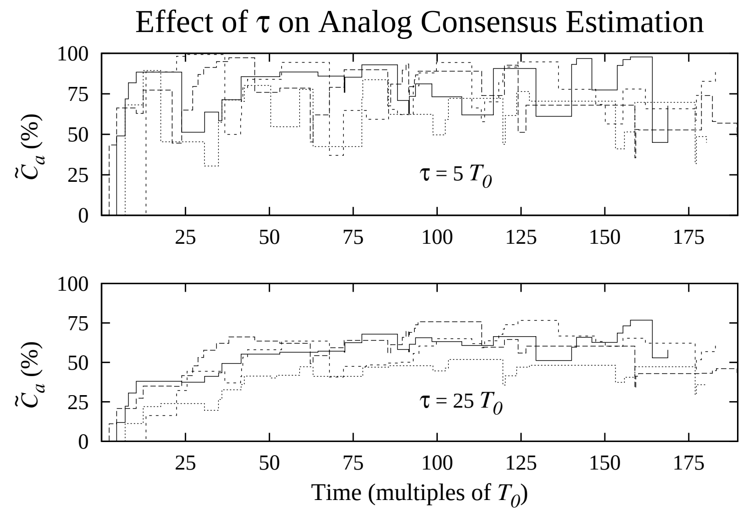


Figure 3.16: In this figure, the same sequence of robot interactions as were used to generate Figure 3.15 are used along with analog consensus estimation to compute $\tilde{C}_a$. $\tau = 5T_0$ in the upper graph, and it is increased to $25T_0$ in the lower graph. Low values of τ allow $\tilde{C}_a$ to vary rapidly, whereas increasing this parameter promotes more accurate measurement. Note also that when $\tau = 5T_0$, the robots tend to overestimate apparent consensus, a phenomenon predicted by the data in Figure 3.13.

corresponds to a fixed series of robot interactions, reinterpreted given the different parameterizations of apparent consensus estimation.

Time in the following figures is measured in terms of T_0, where T_0 units of time can be thought of as the time between a robot's to random teammate encounters. As was discussed in the last section, T_0 in practice is not a constant, but a random variable. After the simulations were run, the distribution of inter-robot-encounter times was computed, and its first quartile (10.55 seconds) was used as the value for T_0.

Figure 3.15 plots four of the eight agreeing robots' digital estimates of apparent consensus over time for two different values of n. In the upper graph of this figure, $n = 5$. The estimates are able to rise very rapidly, but they also are very inaccurate, often over- or underestimating C_a. When n is increased to 25, the robots' opinions are much slower to rise from the initial estimate $\tilde{C}_a = 0\%$ to mirror the actual apparent consensus, but their steady-state estimates are much less noisy.

Figure 3.16 plots the very same robots' opinions of C_a as in Figure 3.15, but in these graphs, the robots used the analog consensus estimation strategy. In the upper graph $\tau = 5T_0$, whereas it is increased to $\tau = 25T_0$ in the lower one. Increasing τ in the analog strategy has the same effect here as increasing n did when $\tilde{C}_a$ was computed digitally. As τ increases, the time required to measure the apparent consensus increases, but the error in measurements decreases. Notice that C_a tends to be overestimated when τ is low. This behaviour was predicted by the $\tau = 5T_0$ curve in Figure 3.13.

Figure 3.17 summarizes the relationship between speed and accuracy for both consensus estimation strategies. Measurement accuracy is plotted versus measurement time for several different

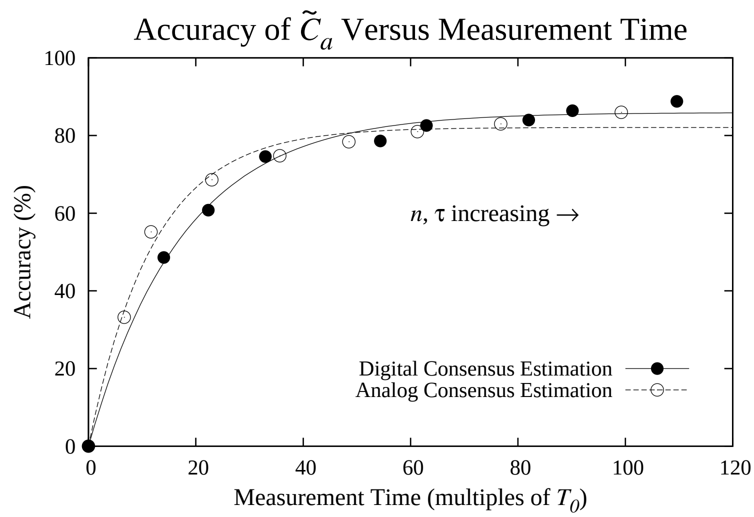


Figure 3.17: In Figures 3.15 and 3.16, it could be seen that increasing n or τ increased the accuracy of apparent consensus measurement, but also increased the time necessary for a measurement to be made. This figure summarizes these results, plotting the accuracy of robots' estimates of C_a against the measurement time. When n or τ is low, measurement accuracy increases rapidly with increased measurement time, but diminishing returns are encountered. Note the great similarity between the performance of digital and analog consensus estimation.

parameterizations, with $C_a = 50\%$. These two quantities for a given configuration were calculated as follows (Refer to Figure 3.15 or 3.16). The time required to make a measurement of apparent consensus was measured as the mean time required for a robot's $\tilde{C}_a$ to reach 50%, which is a measurement of the mean rise-time of the estimates of C_a. The accuracy of a configuration was calculated as 100% minus the mean relative deviation of a robot's $\tilde{C}_a$ from $C_a = 50\%$ in the time after $\tilde{C}_a$ reached 50%. The estimates of apparent consensus during the initial rise-time are not included in the calculation of measurement accuracy. The data plotted were computed from 40 simulated trials, each lasting long enough for approximately 200 robot interactions. The relationship between measurement accuracy and measurement time is remarkably similar for the digital and analog strategies. In both cases, when τ or n is small, the slope of the two accuracy-versus-speed curve is steep, meaning that a small increase n or τ would deliver a substantial increase in measurement accuracy for a small decrease in speed. As n or τ is increased, though, diminishing returns are encountered, and the time-cost of making more accurate measurements increases rapidly. The two approaches to computing $\tilde{C}_a$ are so similar in their observed performance that they can be considered identical for the purposes of quorum testing.

3.4.4 Compound Quorum Testing

Premature commitment due to a false-positive quorum test is a general problem encountered in decentralized consensus estimation and quorum testing. This is because each robot computes $\tilde{C}_a$ with some finite error. The commitment phase is initiated by the first robot that believes that quorum is satisfied for its favoured alternative, and so it is likely that commitment will be initiated by the first robot to experience a false-positive quorum test.

Although premature commitment cannot be eliminated entirely, the problem could be reduced somewhat with the addition of a second quorum test. Once a robot believed that its favoured alternative had satisfied quorum, it would begin to measure the proportion of its teammates that also believed that its favoured alternative had satisfied quorum. These robots would continue to estimate apparent consensus, and if this were to fall below Q, they would reset the second quorum test. In other words, apparent consensus would be measured for two populations, and two quorums tested. The second quorum would be tested only as long as the first was believed to be satisfied. Clearly, the addition of this meta-quorum would increase the reliability of the overall quorum test. Whether or not its inclusion is worth the extra complexity (albeit minor) would depend on the nature of the decision at hand.

This concept of *compound quorum testing* is demonstrated by the structure of honeybee decision-making. Because the bees relocate their colony over long distances, it is not practical to move the individuals one at a time as do the ants. Instead, the swarm, the majority of which remains somewhat dormant during the decision-making process, must be roused by the bees once they believe that quorum has been satisfied. This task is too great for a single bee to accomplish, and so the swarm relocates only once a sufficient number of bees have observed quorum so that their combined effort is sufficiently great to induce the swarm to lift off. Once it lifts off, the airborne swarm is guided by the committed bees to its new home. In their case, the second quorum is implemented via the dormant swarm's inertia and resistance to lift-off, but the net effect is the same.

3.4.5 Population Size and the Resolution of Apparent Consensus

It is important to briefly touch upon the relationship between the number of robots that compose a MRS and the resolution range of apparent consensus (or true consensus, for that matter). To illustrate this relationship, consider the extreme example of a system that contains just two robots. If the two robots favoured different alternatives, then both would observe an apparent consensus of 0%. However, if one of them were to change its opinion to match that of its teammate (perhaps because one robot recruited the other), both would observe an apparent consensus of 100%. In this small system, only these two values of apparent consensus are possible. On the other hand, in a 100-robot MRS, as the population of robots that favoured a particular alternative increased or decreased by one robot, those remaining in the population would observe a change in apparent consensus of $\pm\frac{1}{99} \times 100\%$.

Both hypothetical systems, the 2-robot MRS and the 100-robot MRS experienced the same change - one robot changed its opinion - yet to the individual robots in these systems, the magnitude of the change in the social cue observable by them was significantly different. This in turn limits the number of different quorum thresholds that could be employed in practice. In the two robot-system, any $Q > 0$ would be equivalent to $Q = 100\%$. In general, the apparent consensus in a MRS can only be one of N discrete values, and it will vary over the range $[\frac{0}{N-1}, \frac{N-1}{N-1}]$ in increments of $\frac{1}{N-1}$.

In this thesis, consensus estimation drives a thresholding operation: the quorum test. However, in other applications, a robot might vary its actions continuously with its estimate of apparent consensus, and so the smoothness of its response would depend on the smoothness of apparent consensus[7]. All other factors being equal, as the number of robots that compose a MRS increases, the better the system will be able to take advantage of consensus estimation and direct recruitment. This is not a serious problem, of course, since a system composed of only a few robots likely would be better off employing a more traditional centralized or hierarchical control structure to coordinate the actions of its individual members.

3.5 Commitment

The final phase of the proposed decision-making strategy is the commitment phase. The initial searching phase of a decision identifies the individual alternatives over which the robots will deliberate. Once these have been found, the robots use recruitment to compare them and identify the best one. However, both of these processes are completely decentralized. None of the robots are aware of how many alternatives might have been found, nor their relative qualities. Through consensus estimation, though, an individual is able to estimate the popularity of the particular alternative that it favours. Once a robot believes that the popularity of its favoured solution has reached the threshold of quorum, it concludes that its alternative has become sufficiently popular that the group decision should be completed, with its alternative adopted by the entire system. It is the task of the final phase of the decision-making framework, commitment, to accomplish this goal.

3.5.1 Individual Commitment

The simplest approach to commitment is to have the individual robots adopt their favoured alternative only once they independently determined that their favoured alternative had satisfied quorum. That is, decisions would not involve a commitment phase *per se*; robots simply would exit a decision once they had independently determined that quorum had been satisfied. Unanimous decisions would only be possible under this approach if robots that had exited the decision-making process continued to both encounter and respond to the vote-queries of their still-deliberating teammates. Individual commitment is somewhat similar to the strategy employed in [92].

This commitment strategy does not guarantee that all of the robots will exit a group decision. Some of robots might stagnate in the deliberation phase. For example, lone robots favouring unique alternatives might not be recruited before all of their teammates detected quorum and exited the

[7] When C_a is estimated using the digital approach, the resolution of $\tilde{C_a}$ will be determined by $min(n, N-1)$, since $\tilde{C_a}$ will vary from 0% to 100% in increments of $\frac{1}{n}$. It would depend on the system and application at hand whether n or N was the limiting factor.

decision. Even though the lone robots would still receive vote-messages from their teammates, C_a for them would be zero, and so quorum would never be satisfied.

In some ways, *Temnothorax* commitment behaviour is a kind of individual commitment, although it involves a kind of stigmergy that makes it more robust. An ant will commit to its favoured site once it detects a quorum of teammates while visiting it. Once the ant has committed, it appears to treat its favoured site as the colony's new home [70], quickly transporting teammates that it encounters to it. Other ants commit to their favoured sites only once they too have determined that quorum has been satisfied. However, because the population of a site that has induced commitment will begin to rise quickly following commitment (due to the ants transported there by committed ants), other ants that favour a site that has induced commitment will become more likely to commit to it as well, triggering a kind of chain reaction of commitment. Note that this stigmergic feedback in the ants' commitment behaviour only exists because the process of relocating a colony to a particular site (which follows commitment) directly interacts with the manner in which quorum is tested.

Nonetheless, commitment to multiple sites is possible, and individual commitment does not specifically address this particular fault. When two or more alternatives induce commitment, a dec-MRS will fragment amongst them. In order to increase the likelihood of bringing about the unanimous adoption of a single alternative by the end of a collective decision, some additional mechanism is required.

Individual commitment can augmented slightly to bring about a more social behaviour. In [53], the robots assembling into convoys estimated number of robots that had gathered using a biologically inspired feedback mechanism called chorusing. Once one of the robots determined that a sufficiently large group had gathered (analogous in this discussion to a robot having detected the satisfaction of quorum), it would set a timer. As soon as the timer elapsed, the robot would broadcast to its teammates that it was time to collectively depart, and the assembled robots would depart with the reception of the first such message. Because the chorusing mechanism is somewhat noisy, the complete formation of a convoy tends to lag the chorusing signal reaching its preset threshold. By calibrating the timer to account for this lag, a more reliable collective departure would become more likely. However, this is an open-loop mechanism which is very domain specific. In the next section, feedback from the collective state is used to guide the commitment process.

3.5.2 Gossip-Based Commitment

A better approach to the commitment phase explicitly would take into account the goal of decision unanimity. A robot that believes that quorum has been satisfied for its favoured alternative should not selfishly exit the decision-making process on its own. Rather, it has a responsibility to its teammates to ensure that all of them commit to its favoured alternative, too. This is particularly important when a decision with some geographic component is being made, since robots will tend to move away from the location at which the decision was made as they exit the decision, likely putting them out of range of their teammates still in the decision-making process. A collective decision also could be made to adopt some new behaviour that should not begin until all work on the current task has halted. This concept of mutually exclusive behaviours is discussed in greater detail in the next chapter.

If the individual robots possessed global broadcast communication capabilities, unanimous commitment would be a trivial problem. Once a robot entered the committed state, it simply would broadcast the instruction to do so to its teammates, and unanimity would be assured. However, single-hop global broadcast is impractical at best in a dec-MRS [62].

Gossiping [10, 43] is an elegant stochastic communication technique that allows robots with very short range communication capabilities to share information system-wide. When a robot wishes to broadcast some message system-wide via gossiping, the robot sends the message to randomly selected teammates. In a well-stirred dec-MRS, a robot can achieve this simply by transmitting the information to the teammates that it encounters as it wanders about. The robots that receive the information adopt the same behaviour, and the information will flood throughout the entire system.

Clearly, if the robots continue to gossip long enough, all of them eventually will receive the information disseminated by the initiating robot. How should the process be terminated, then? Demers et al. identified a simple approach (although they use the term *rumour spreading* instead of gossiping). As each member of the system transmits the information to a randomly selected teammate, it identifies whether or not the teammate already knew what was sent. If the transmission increased the size of the informed population, then the sender is unaffected. On the other hand, if the recipient already knows the information, the sender exits the gossiping behaviour with probability $\frac{1}{k}$. The parameter k controls the thoroughness of the collective behaviour. As k is increased, the probability of all of the robots receiving the information by the time the last knowledgeable one exits increases. In another variation, k is a counter, and each robot counts the number of times that it sends a redundant message. Again, increasing k increases the probability of all of the robots becoming informed [18]. Because gossip is a stochastic process, it is possible that some robots might never receive the information before the process terminates. Although increasing k increases the probability of the information reaching every member of a dec-MRS, the probability of collective success will never be 100%. This is a general property of decentralized, stochastic algorithms.

By organizing commitment as a gossip-style algorithm, unanimity becomes much more likely. The "information" that is shared in the case of a group decision is the belief that a particular alternative has satisfied quorum. Once a robot believes that its favoured alternative has satisfied quorum, it begins gossiping this belief, and the rest of its system rapidly ($O(ln(N))$ [43]) quickly will commit to it as well. The basic concept is easily modified to accommodate multiple alternatives inducing commitment. Simply put, a robot will commit to whatever alternative was specified in its most recently received commit-message.

Note the difference between the commitment and deliberation phases of a decision. During deliberation, the quality of an alternative directly influences the likelihood of a robot being recruited to it. Once an alternative has induced commitment, its quality no longer plays any role. The purpose of commitment is not to select the best alternative, its purpose is to promote unanimous commitment to the single alternative that was identified as the best one by the preceding deliberation. The commitment phase thus can be thought of as a war of attrition between robots committed to different alternatives. The likelihood of a robot receiving a commit-message referring to a particular alternative is directly proportional to the proportion of its teammates that are committed to it. Therefore, the alternative to which the most robots have committed will tend to be the one selected unanimously by the commitment phase's end. Because the number of robots committed to a given alternative will rapidly increase via positive feedback following the initial commitment to it, the first alternative that is believed to satisfy quorum and thus induce commitment will be the one most likely to be unanimously adopted by the decision's end.

Instead of using a counter to terminate the process, a simpler approach employing a timer is adequate. Here, when a robot receives a commit-message specifying an alternative to which it is not committed (*e.g.* a commit-message that is not redundant), it commits to the alternative and responds with an acknowledgment. If a robot receives a commit-message specifying an alternative to which it already is committed, it does not respond at all. Committed robots reset an internal timer whenever they receive an acknowledgment, or whenever the alternative to which they are committed changes. As more and more of the robots commit to the same alternative, the probability of an individual robot receiving an acknowledgment will decrease, and so the robots' commitment timers will be less likely to be reset. A committed robot will exit the group decision once its timer reaches a preset limit, called the *commitment timeout*. This time limit is somewhat analogous to the counter k in the work of Demers et al. in [18]. Increasing the commitment timeout increases the probability of the commitment phase achieving unanimity by increasing the amount of time a committed robot will remain in the committed state without receiving an acknowledgment. Note the similarity between this timer-based approach and analog consensus estimation. Instead of a dedicated timer, a the quorum and kin indices could be reused, with the quorum index (now a sort of committed index) being incremented with the reception of every acknowledgment. Individual committed robots would exit a decision once the estimate of consensus fell *below* some preset threshold. Similarly, if a count of received acknowledgments to commit-messages would be preferable, the digital consensus estimation hardware could be reused. Again, a committed robot would exit a decision once it believed that the proportion of its teammates that had not committed to its alternative dropped below some preset threshold.

3.6 Summary

In this chapter, a decentralized collective decision-making framework was described, inspired by behaviours of ants and honeybees. This three-phase approach, composed of searching, deliberation and commitment, utilizes several decentralized behaviours, all within reach of very simple robots. These are just the sort of robots that are expected to compose large-population dec-MRS, and so the framework should be widely applicable to decision-making problems that such systems might encounter. Only the initial searching phase of a decision is significantly coupled to a specific decision.

In the later deliberation and commitment phases, decentralized computation is carried out independent of the environment in which the MRS operates, as the robots modify each others' states directly, rather than through an intermediate environmental channel. The deliberation phase singles out the best of the alternatives that the individual robots were able to find during their initial search using the positive feedback of recruitment. Because each robot favours only one alternative at a time, the computational complexity of this distributed comparison operation to the individual robots is constant, regardless of the number of alternatives known collectively.

Concurrent with their recruiting activities, the robots also estimate the popularity of the particular alternatives that they favour. Once again, this operation is computationally simple, and its complexity to the individual robots is constant. Two different approaches to consensus estimation were described: one digital, and one analog. Although the digital approach is likely to be the more widely-used of these, the simple fact that a small analog implementation is practical demonstrates that consensus estimation, an operation likely to be valuable in many other dec-MRS problems, could be taken advantage of by systems composed of the most elementary robots.

Decentralized deliberation continues until one of the robots determines that its favoured alternative has become sufficiently popular to end deliberation so that its alternative might be selected unanimously. This decision is made by an individual robot through a quorum test, in which it compares its estimate of its favoured alternative's popularity to a threshold. Consensus estimation and quorum testing can be tuned to emphasize speed or accuracy, or to strike a balance between the two.

Finally, in the commitment phase, the entire dec-MRS coalesces around a single alternative, the one that satisfied quorum. Using a gossip-style approach, the knowledge that one particular alternative is to be adopted will flood the system, causing it to be chosen unanimously. In this way, a decentralized system composed of simple, locally communicative individuals can behave as a cohesive whole, an *artificial superorganism*, and make intelligent decisions as though they were a single intelligent entity.

Chapter 4

Unary Collective Decision-Making: The Cooperative Task Transition Problem

The focus of this thesis is collective decision-making in decentralized multiple-robot systems. In the general case, a collective decision can be viewed as a best-of-N operation in which the best of N *a priori* unknown alternatives is selected unanimously. In this chapter, a simpler operation, called the unary decision, is made by modifying the deliberation phase of the decision-making framework by eliminating recruitment. Unary decision-making is applied in this chapter to a problem that has limited the development of more advanced dec-MRS: cohesively stepping through a sequence of subtasks. The unary decision-making strategy is described in detail, and experiments with simulated and physical robots demonstrate its performance in the context of a collective construction task.

4.1 Introduction

In general, most decisions can be represented as best-of-N decisions in which the decision-makers must identify and evaluate a set of candidate alternatives from which one unanimously is selected. A subset of these decisions are those in which the candidate alternatives are known *a priori,* and of these, the situation in which only one alternative exists arises with surprising frequency. In this case, the question facing the decision-makers is whether or not the alternative should be adopted in place of the *status quo*.

Complex missions often are decomposed into a series of simpler subtasks. The overall mission is achieved by completing each of the subtasks in order. At each transition from one subtask to the next, a decision is required: "is the current subtask done?". Either it is, and the focus of the team should shift to the next subtask, or it is not, and the *status quo* of working on the current subtask should be retained. Because the decision is to accept or reject a single proposed alternative, these decisions are referred to as *unary decisions* in this thesis. The focus of this chapter is the application of unary group decision-making to the synchronization of dec-MRS subtask transitions

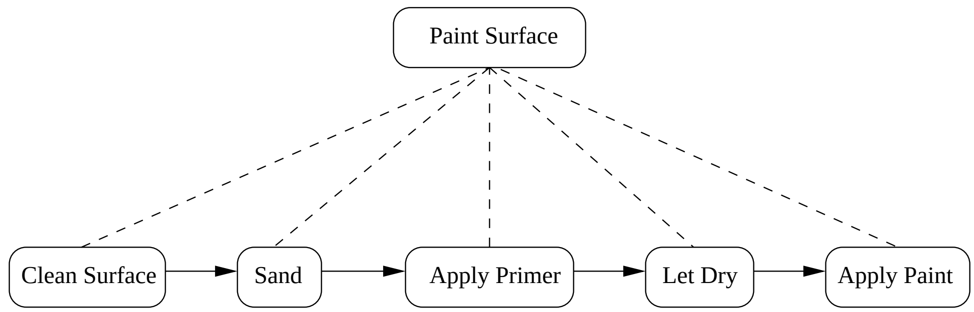


Figure 4.1: This figure illustrates the decomposition of a painting task into a sequence of simpler subtasks. In order to complete the overall mission a system must complete each subtask in order. A group decision is required at each subtask transition to ensure that all of the robots make the transition at the same time, otherwise robots in adjacent tasks might interfere with each other, resulting in a failure of the overall mission. At the same time, a transition must not occur until the current subtask has been completed. These two concerns are addressed by a cooperative unary decision.

in a complex mission, but the reader should keep in mind the generic nature of the unary collective decision-making problem.

Refer to the surface-painting task depicted in Figure 4.1 as an example of a dec-MRS task that requires unary decision-making. This mission can be decomposed into five subtasks: clean the surface, sand it, apply primer, allow the primer to dry, and apply the paint. These subtasks are *mutually exclusive*; each one must be completed and then all work on it halted before any work can commence on the next one [62][1]. Thus there is a need to synchronize the transition from one subtask to the next so that all of the robots make the transition at the same time. Consider the transition from the sanding subtask to priming. If the robots did not all make the transition at the same time, the overall painting mission would fail because the two mutually exclusive subtasks would have robots working on them simultaneously, leading to interference (e.g. the dust from sanding would prevent the primer from being applied correctly). When this occurs, it is said that *mutual exclusivity* is *violated*.

On its own, synchronization is not a difficult problem, as a single global broadcast operation could achieve it. However, it also is important to ensure that each subtask truly is complete before the transition to the next one is initiated. Consider the transition from the Let Dry subtask to the Apply Paint subtask. Clearly, if the robots were to begin the Apply Paint subtask prematurely, even if they did this in a synchronized manner, the primer would not have been given sufficient time to dry. Mutual exclusivity would not have been violated, but the paint would be applied to a wet surface and thus would not adhere properly, once again resulting in overall mission failure.

If we assume that a mission and its decomposition into subtasks are known *a priori* by the robots, a unary decision could govern each subtask transition. For a solitary robot, the problem is as simple as accurately determining when each subtask has been completed. Since no other robots would be

[1] It could be argued that the individual subtasks in this example could be designed so that some of them could be carried out in parallel. However, for the purpose of this discussion, assume that this is not the case, and that any two robots working on different subtasks simultaneously would destructively interfere with each other, resulting in overall mission failure.

present, the interference problem is non-existent. Multiple robots would be able to complete each subtask more rapidly than a single robot, so the ability to deploy a MRS to complete a sequence of subtasks is desirable. The solution for a centralized MRS is not much more complicated than that of its solitary counterpart. The central controller would make the decision to begin the next subtask based on data collected by the other robots and transmitted to it, and its decision then would be dictated to the entire MRS.

Not only would a MRS be able to complete each subtask faster than a solitary robot, it also would be able to measure the state of each subtask more rapidly. This is because many independent measurements could be made in simultaneously. Furthermore, these measurements would be made with many different sensors, since each robot would report measurements taken with its own sensors, so individual sensor calibration would be less of a concern. This would tend to render the deliberations of the MRS more precise.

The synchronization of subtask transition demanded by mutual exclusivity is provided in a unary decision by the commitment phase, just as it is in the general decision-making framework. The question of *when* the subtask transition occurs is governed by the deliberation phase, but recruitment is absent. In the case of a best-of-N decision, the robots are deciding *which* alternative to adopt, and it is assumed that the *status quo* must be replaced by one of them. The question posed by a unary decision is: "Should the *status quo* be replaced by the proposed alternative?" Deliberators in a unary decision have made the individual decision that the proposed alternative should be adopted and are waiting until a sufficient proportion of their teammates independently have come to the same conclusion. Once a deliberator believes that the size of the deliberating population has satisfied quorum, it concludes that there is sufficient consensus in favour of the proposed alternative, so it commits to it. This initiates the commitment phase's gossip-style broadcast operation, which leads to the proposed alternative's unanimous adoption.

Quorum in a unary decision serves a subtly different purpose than it does in a best-of-N decision. In a best-of-N decision, quorum testing delays commitment until iterative recruitment has identified one of the candidate alternatives as better than the others. Its role in a unary decision is to prevent commitment until a sufficient proportion of the robots independently have decided that the proposed alternative should replace the *status quo*. If the robots of a dec-MRS had perfect sensing, a quorum of zero would be satisfactory, since a single robot's decision that the alternative should be adopted would be sufficiently reliable to justify commitment. Of course, perfect sensing does not exist in the real world. As such, quorum should be set so that, once it is satisfied, it will be unlikely that a collective mistake will have been made. Given a quorum, the number of independent decisions necessary to satisfy it increases with population size (since quorum is expressed as a relative, rather than absolute population size). In this way, a unary decision allows a dec-MRS to leverage its redundancy, rather than fall victim to it.

In this chapter, the decision-making framework of this thesis is applied to the decentralized

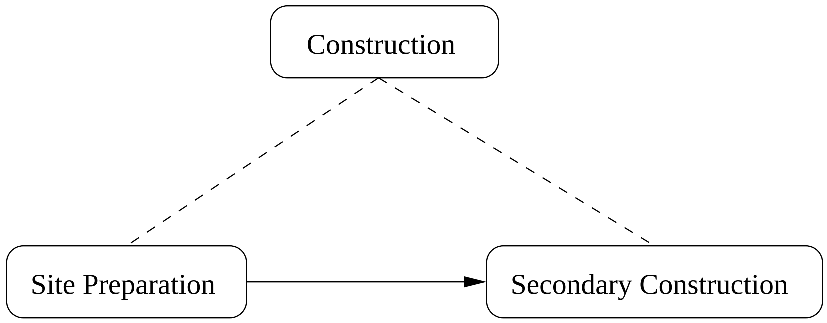


Figure 4.2: A construction task can be decomposed into an initial site preparation subtask followed by secondary construction. The purpose of site preparation is to remove debris from the construction site so that the more advanced secondary construction can proceed. These two subtasks are mutually exclusive, so a unary group decision is required to coordinate the transition between them.

unary decision-making problem for task-transition. In the next section, a collective construction task is presented to provide context for this application, and the details of the unary decision-making behaviour are explained. Experiments that were conducted in simulation are then presented, followed by a series of experiments using real robots. The chapter then closes with some conclusions on the performance of decentralized unary decision-making.

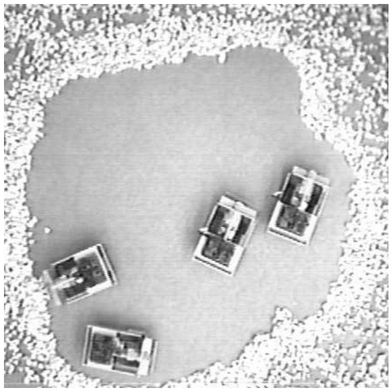
4.2 Collective Construction and Decision-Making Behaviour

In this section, the blind bulldozing collective construction behaviour is introduced. The goal of the robots is to decide unanimously that the construction task has been completed, thus allowing some subsequent group task to commence without interference due to robots continuing to work on blind bulldozing. Next, the details of the individuals' behaviours that produce a unary group decision are described. In the next section, a series of simulated experiments are described, followed by a section detailing similar experiments with real robots.

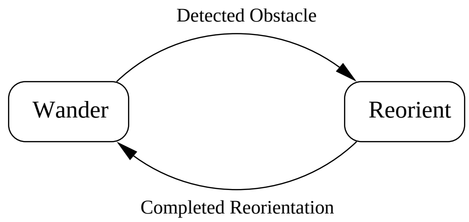
4.2.1 Collective Construction

Collective construction is a good example of a task that can be decomposed into a sequence of mutually exclusive subtasks. For example, building materials have to be located, brought to the construction site and then assembled [60]. If assembly were to begin before foraging for materials had halted, the foragers might retrieve materials to the stockpile from those that had already been assembled into the desired structure. Were assembly to begin prematurely, the quantity of construction material gathered might be insufficient to complete the assembly process.

Often, advanced construction must be preceded by preparatory work. For example, before a structure can be erected, the site upon which it is to be built must be cleared of debris. This clearing process is called *site preparation*, and it has been identified as a critical task for MRS missions to other planets [40, 66]. The more advanced construction that follows site preparation is referred



(a)



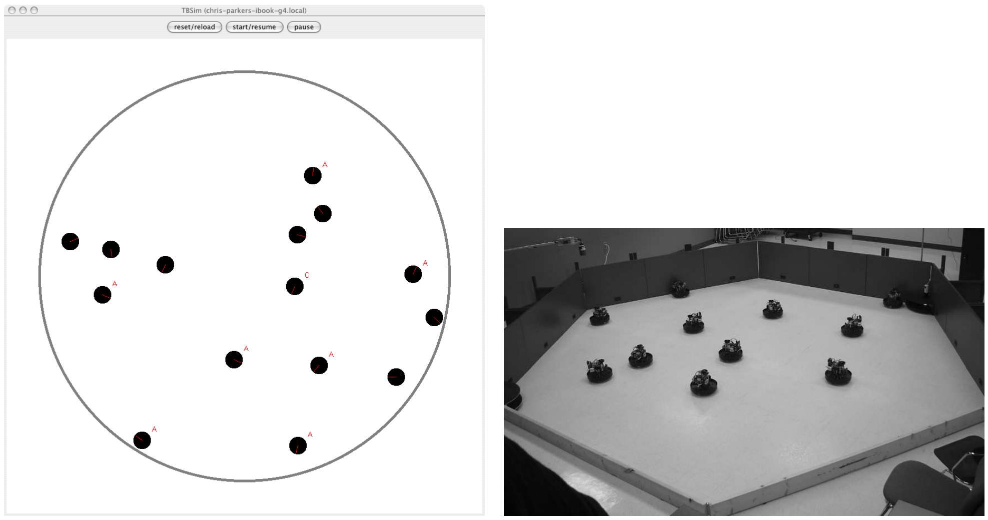
(b)

Figure 4.3: The image on the left of this figure depicts four robots engaged in the blind bulldozing site preparation task. Their goal is to expand the initial clearing in the debris-field to permit more advanced construction to take place. Their individual behaviours are controlled by the simple state-machine given on the right. The robots clear debris by plowing in straight lines in the wander state, and then randomly reorient once the debris has been pushed into the site's wall or whenever a teammate is encountered [65, 64, 60].

to as *secondary construction*. Site preparation is a somewhat coarse task, involving earth moving equipment, whereas secondary construction will tend to be more delicate in nature. The two should not be carried out simultaneously, as any secondary construction begun before site preparation has halted risks being bulldozed [62]. A collective decision to halt site preparation and begin secondary construction clearly is required here, and it is this specific unary group decision that is the focus of the experiments in this chapter.

Earlier work proposed a biologically inspired site preparation algorithm called *blind bulldozing* [64, 60]. Blind bulldozing clears a construction site out of a field of debris through the uniformly distributed plowings of a team of autonomous bulldozers. The task is complete once the site has reached a predetermined size, chosen to be sufficiently large to accommodate whatever secondary construction is intended to follow. Figure 4.3 depicts a scene from a blind bulldozing experiment along with the simple state-machine that guides the individual robots' behaviours.

The actual bulldozing is done in the wander state, in which the robots travel in straight lines. The force exerted on a robot's plow increases with the amount of debris that it has plowed up. Once this force exceeds a preset threshold, the robot enters the reorient state, in which it rotates on the spot to a new random heading, and then re-enters the wander state. Robots in the wander state also enter the reorient state whenever they encounter a teammate. The random reorientations distribute the robots' plowing uniformly about the construction site's perimeter, so the site grows evenly and adopts a circular shape as though it was being "inflated" by the robots [28]. It also makes the robots' encounters with each other well-stirred [62]. In [64] it was conjectured that the individual robots could infer the completion of the site preparation task by measuring the distance that they traveled between reorientations. If a robot travels a certain distance in a straight line before it reorients, it



(a) (b)

Figure 4.4: This figure depicts the environments of the unary decision-making experiments. On the left is a screen-shot from one of the simulated experiments. Each of the black discs is a robot. The image on the right is a photograph of a real dec-MRS making a unary decision about task-completion. These environments are static, but they are good analogs of the blind bulldozing domain towards the end of the task. In both cases, the environment is sufficiently large that the individual robots eventually will conclude that the task is complete.

knows that the diameter of the clearing is at least that great. Experiments since have demonstrated that this is a viable strategy to detect task completion [37]. However, because the robots' paths are random, each will tend to detect task completion at a different time. Furthermore, odometry errors will introduce uncertainty into each robot's individual decision-making. A group decision therefore should be employed to coordinate the transition from blind bulldozing to secondary construction.

Environment

Blind bulldozing would be employed when a construction site is covered in debris that must be cleared to prepare for secondary construction. It works by expanding an initial clearing in the debris. The rate at which the clearing grows will slow over time as the site expands and its walls become reinforced with plowed debris. Towards the very end of the task, the growth of the clearing will have sufficiently slowed that the environment can be approximated by a somewhat circular enclosed arena that is fixed in size. The experimental environments for the simulated and physical experiments are shown in Figure 4.4. The robots provide an indication of the relative scales of the environments. The robots are programmed to conclude independently that the site is large enough once they travel a sufficiently long path while in the wander state (straight-line motion). Because the motion of the individual robots is random, each will traverse such a path and conclude that the site preparation task has been completed at a different time. Without a group decision, each robot would begin secondary construction at a different time and mutual exclusivity would be violated.

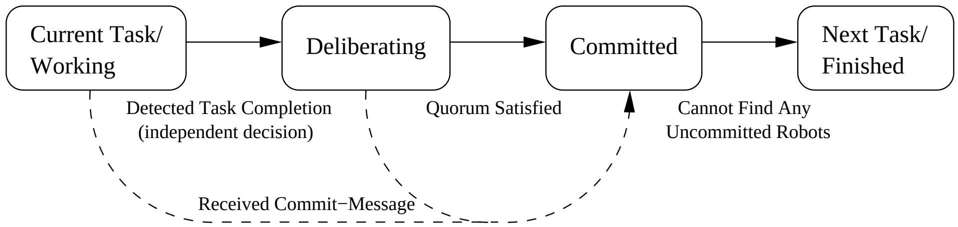


Figure 4.5: A robot's cognitive behaviour during a cooperative unary decision is divided into four states. The robots initially believe that the current group task is not yet complete, and so they work on it. When a robot decides that the current task is complete, it enters the deliberating state, in which it gathers the opinions of its teammates as they are encountered. Based on their opinions, a deliberating robot estimates the apparent consensus in favour of the current task being complete. Once it believes that this has satisfied quorum it enters the committed state, in which it instructs other robots to commit. Uncommitted robots that are told to commit do so and respond with an acknowledgment. When a committed robot no longer receives acknowledgments to its commit-messages, it concludes that all of its teammates have either committed or moved on to the next task, and it does so as well.

4.2.2 Individual Behaviours for a Unary Group Decision

In this section, the decision-making behaviour of the robots is described. This behaviour is the same in both the simulated and physical experiments. Refer to Figure 4.5 for the following discussion. Each robot begins a decision in the working state, in this case participating in the blind bulldozing task.

A robot will decide that site preparation is complete once it travels a sufficiently long path while in the wander state of the blind bulldozing behaviour, at which point it will enter the deliberating state. Robots in the deliberating state request the opinions of the teammates that they encounter regarding the state of the site preparation task (the robots continue to move about as though they were still blind bulldozing while they are making a group decision). Robots that receive these queries respond with either a "yes" or "no" vote-message, indicating that they either do or do not believe that the subtask is complete. Based on these responses, the deliberating robots estimate the apparent consensus, the proportion of their teammates that agree that site preparation is complete, and compare their estimates to the quorum threshold, Q.

Once a deliberator believes that quorum has been, met it commits. The committed robots continue to wander about, but now when they encounter a teammate, they instruct it to commit as well. If a robot has not yet committed (i.e. it is in either the working or deliberating states) and receives a commit-message, it immediately enters the committed state and responds with an acknowledgment. This acknowledgment informs the sender of the commit-message that it just encountered an uncommitted robot. Robots already committed ignore the commit-messages. Eventually, all of the robots will commit, and so the acknowledgments will cease. Committed robots conclude that every other robot has either committed or begun the next task once the time since they last received an acknowledgment reaches the preset commitment timeout, T_c. These robots leave the collective decision by

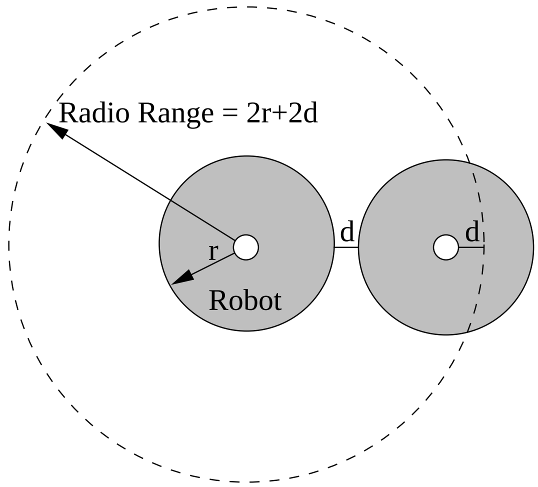


Figure 4.6: All of the communication in the simulated task-completion experiments was local and anonymous. The robots were circular, with their antennae located at their centers. Robots could detect teammates when they were a short distance (d) away, and their radio transmission ranges were set to twice their radius plus twice the teammate detection range. In practice, although it was possible for more than one robot to receive a particular teammate's transmission, 95% of the messages were one-to-one, and the remainder were mostly one-to-two.

entering the finished state. At this point, they would begin the next subtask in the mission.

4.3 Simulated Experiments

In this section, a series of simulated experiments examining the performance of unary decision-making are presented and their results examined.

4.3.1 Environment

The simulated environment was implemented using the TeamBots open-source multiple-robot simulator [85]. At the time of writing, this package was no longer supported. However, it still was available online for download and its interface strikes a good balance between simulation fidelity and ease of development. A scene from a simulated trial can be seen in Figure 4.4(a). The experimental environment was a circular arena, 12 meters in diameter.

4.3.2 Robots

The 15 individual robots in the simulated experiments were differential-drive platforms, 0.5 meters in diameter. Each robot was able to communicate with its teammates via omnidirectional local broadcast communication. The range of each robot's radio was short enough that two robots' hulls had to be within $2d = 0.1$ meters of each other in order for them to communicate. In the wander state, the robots moved in straight lines at a constant velocity, and they would enter the reorient state when an obstacle came within $d = 0.05$ meters of them on the forward-facing side. Because wheel slippage was not included in the simulations, the length of a path was calculated by multiplying the time between reorientations by the robot's velocity. Refer to Figure 4.6 for a graphical depiction of these ranges. All of the communication was anonymous; the robots were unaware of the identities of the teammates that they encountered.

When apparent consensus was introduced earlier in this thesis, analog and digital approaches for its estimation were presented. Both of these were investigated in the simulated unary decision-making experiments. The analog approach utilizes a pair of variables called the quorum and kin indices that decay exponentially with time, and the precision of analog consensus estimation is determined by the time constant τ, which sets the rate at which these indices decay. The greater τ is made, the more slowly they decay and the more precisely C_a is estimated. The digital approach computes $\tilde{C_a}$ as the proportion of the n most recent vote-messages received in response to a deliberator's queries that are affirmative. There is a fundamental trade-off between the speed and accuracy of consensus estimation [62], and so these two methods should produce similar results when calibrated equivalently.

4.3.3 Experimental Trials

Because the physical behaviours (and thus the communicative histories) of the robots in this particular domain do not depend on their decision states — they always move according to the blind bulldozing algorithm in Figure 4.3(b) — it was possible to run a series of generic simulated trials and then reparameterize them offline to generate any desired configuration of unary decision-making. These generic trials were produced as follows. Whenever a robot encountered a teammate, it would broadcast a SEND-message. This message would include its ID, as well as a rolling 8-bit integer, which would make the message uniquely identifiable. The recipient(s) of such messages would respond with a RESP-message, which included the responder's ID, another rolling 8-bit number, followed by the ID and 8-bit number that were in the original send message. RESP-messages were not responded to. For example, robot-1 might encounter a teammate and thus it would send the message "SEND.1.47". Suppose that robot-8 received this message. It would respond with the message "RESP.8.117.1.47"[2]. Robot-8's response clearly can be identified as the response to robot-1's SEND-message by the inclusion of ".1.47". The times at which messages were sent and received were logged along with the messages themselves. Additionally, robots logged the distances that they had traveled in a straight line whenever they entered the reorient state. 40 generic trials were run, each lasting 50 000 simulated time steps at a resolution of 50 milliseconds per time step.

In order to reparameterize a generic trial to generate a specific unary decision-making trial, the parameters had be specified. These included the length of the path that a working robot would have to travel in the wander state in order to enter the deliberating state, the quorum threshold, (Q), either n or τ (depending on whether digital or analog consensus estimation was to be used),

[2] This example assumes that these were the 47$^{\text{th}}$ and 117$^{\text{th}}$ messages of the initial sender and the responder, respectively.

679.65	robot-11	Path	11.3m	679.65	robot-11	Path	11.3m	
				679.65	robot-11	State	TaskComplete	
681.45	robot-04	Reoriented						
682.25	robot-11	Reoriented		682.25	robot-11	Reoriented		
683.35	robot-04	Enc_Robot						
683.35	robot-04	Path	0.5m					
683.40	robot-04	SentMsg	SEND.04.47					
683.40	robot-08	RecvMsg	SEND.04.47					
683.40	robot-08	SentMsg	RESP.08.50.04.47					
683.45	robot-04	RecvMsg	RESP.08.50.04.47					
684.65	robot-14	Path	1.8m	684.65	robot-14	Path	1.8m	
684.85	robot-02	Enc_Robot						
684.85	robot-02	Path	6.0m	684.85	robot-02	Path	6.0	
684.85	robot-13	Enc_Robot						
684.85	robot-13	Path	1.8m	684.85	robot-13	Path	1.8	
684.90	robot-02	SentMsg	SEND.02.39					
684.90	robot-13	RecvMsg	SEND.02.39	684.90	robot-13	SentMsg	QUERY	
684.90	robot-13	SentMsg	RESP.13.55.02.39					
684.90	robot-13	SentMsg	SEND.13.56					
684.95	robot-02	RecvMsg	RESP.13.55.02.39					
684.95	robot-02	RecvMsg	SEND.13.56	684.95	robot-02	RecvMsg	QUERY	
684.95	robot-02	SentMsg	RESP.02.40.13.56	684.95	robot-02	SentMsg	NO	
684.95	robot-13	RecvMsg	RESP.02.40.13.56	684.95	robot-13	RecvMsg	NO	0.4

Figure 4.7: Because the motion of the robots is independent of their decision state, a single series of 40 generic trials was run. In these, the robots sent generic messages to teammates as they were encountered to which the recipients would respond with similarly trackable messages. The lengths of the paths traveled while in the wander state also were logged. These generic logs were post-processed to generate unary decision trials with whatever parameterization was desired. This figure illustrates a portion of a generic log on the left with a post-processed version of it on the right. The first three columns of the two logs are: time of event, the robot that logged the event, and the specific event. The remaining columns are event specific data, such as the message received or transmitted, the length of a path, or the new decision state.

and the commitment timeout. With this information, and by assuming that each robot began in the working state, a simple script could be written that would read through a generic trial's log, step the individual robots through the different states of a decision and modify each line of the generic log accordingly. Before a robot had logged a sufficiently long path, all of its SEND-messages were deleted, along with any RESP-messages sent in response to them. Once a sufficiently long path has been traversed, the script would update that robot's state to deliberating, at which point all of its SEND-messages were converted to QUERY. A response to a QUERY was changed (tracked via the rolling integers and ID tags) to NO if the recipient was in the working state, YES if the recipient was in the deliberating or committed states, or eliminated altogether if the recipient had entered the finished state[3]. Via n or τ, a deliberating robot's estimate of C_a would be recomputed with the reception of each vote-message. Once this reached Q, the robot's state would be changed to committed. Committed robots' SEND-messages were changed to commit-messages, and the RESP-messages of any uncommitted robots that received these were changed to acknowledgments, in addition to updating those recipients' states to committed. Committed robots measured the amount of time since they last received an acknowledgment (or since they entered the committed state prior to the reception of any acknowledgments), and changed their state to finished once this had reached the commitment timeout, T_c. The left half of Figure 4.7 illustrates a portion of a generic log (on the left) along with a post-processed version of the same section (on the right).

[3]Finished robots did not respond to messages from teammates, since in the general case, beginning the next subtask might move them out of communication range of their teammates.

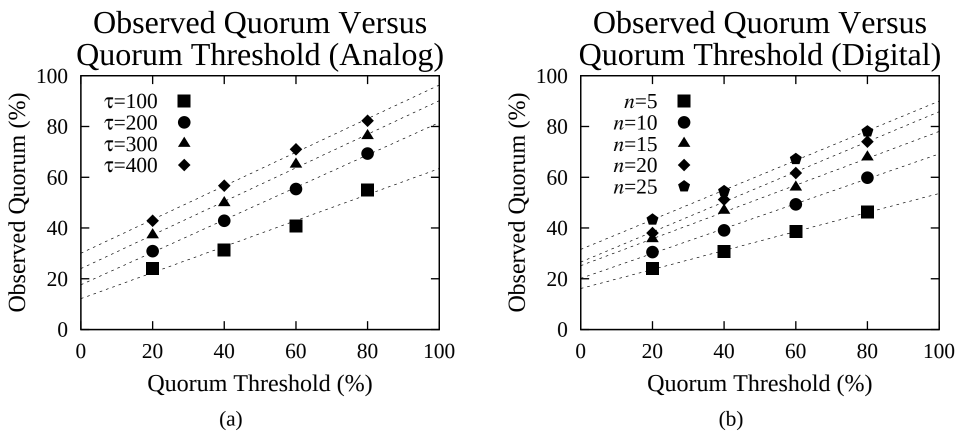


Figure 4.8: As the quorum threshold increases, the observed quorum also increases. This occurs because the quorum test delays the beginning of the commitment phase of a decision until a sufficient proportion of the robots have detected task completion for the quorum test to be likely to be positive. Increasing n or τ decreases the likelihood of false-positive quorum test, resulting in a greater observed quorum. Note that both the analog and digital approaches to consensus estimation produce similar results.

All of the reparameterized trials used a minimum path length of 11 meters. Estimates of apparent consensus were computed digitally using n equal to 5, 10, 15, 20 and 25 teammate opinions, and using the analog approach with τ equal to 100, 200, 300 and 400 seconds. The values 20%, 40%, 60% and 80% were used for the quorum threshold for each of these nine configurations.

4.3.4 Results and Discussion

The results of the unary decision-making experiments conducted in simulation are discussed here in two parts. First, quorum testing performance is examined, followed by an analysis of the commitment phase of the decision-making process.

Quorum Testing and Consensus Estimation

The individual robots cannot precisely measure the apparent consensus present amongst their teammates, and so $\tilde{C}_a$ is noisy. Therefore, the actual consensus present in a system at the time of commitment will vary somewhat. To distinguish this from some desired quorum, or the precise value of the quorum threshold, the term *observed quorum* is used here to denote it. Note that observed quorum is measured in terms of true consensus, C_t, which is equal to $\frac{N_a}{N}$, whereas $\tilde{C}_a$ is an estimate of $\frac{N_a-1}{N-1}$, so some discrepancy should be expected between the two given the relatively small population size of the simulated MRS used in these experiments.

Earlier in this thesis, when consensus estimation and quorum testing were introduced, it was suggested that Q be chosen so that it was somewhat higher than the desired quorum. For example, when $n = 15$, a value of $Q = 80\%$ was suggested to test a quorum of 50%. However, the selection of Q based on the quorum and n (or τ) is heuristic, requiring a subjective judgment on the part of the

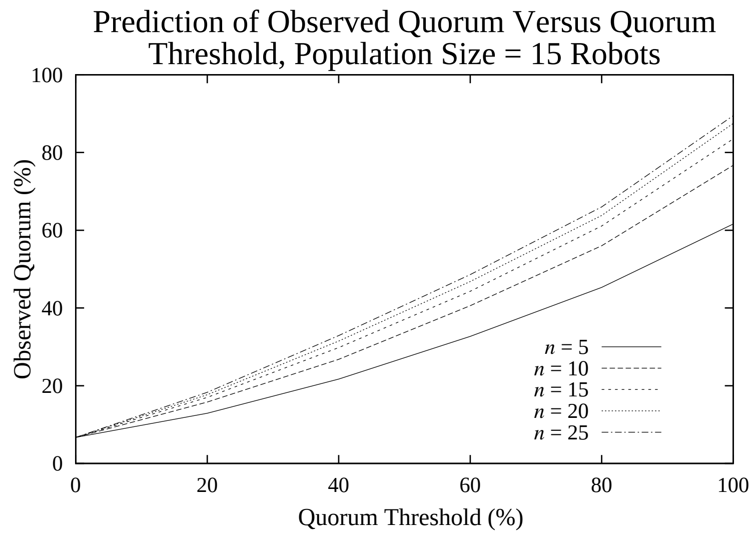
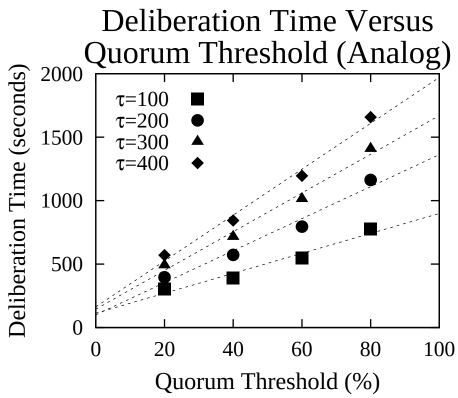


Figure 4.9: This figure presents a theoretical prediction of the relationship between the observed quorum and the quorum threshold for a multiple-robot system of the same size as the one used in the simulated experiments. The analysis used to produce this figure assumes that the rate at which vote-messages can be gathered is insignificant, but this was not the case in the experimental trials. This difference explains the discrepancy between this Figure and the real data plotted in Figure 4.8 for lower values of Q.

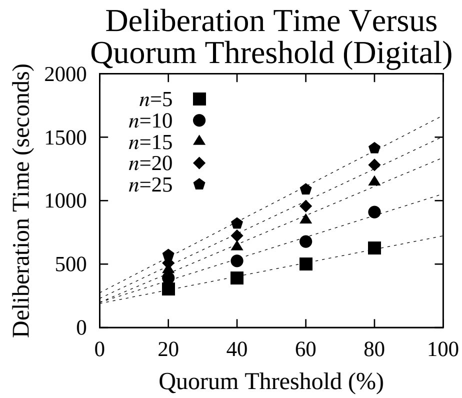
system designer. For this reason, results are presented with respect to Q, rather than some specific quorum that might have been intended by the combination of Q and n (or τ).

Figure 4.8 plots the mean observed quorum versus Q for each value of n and τ. Increasing Q increases the observed quorum, demonstrating that the quorum for a decision can be set *a priori* using both analog and digital consensus estimation. Figure 4.9 plots a theoretical prediction of the relationship between observed quorum and the quorum threshold, which appears very similar to the experimental results. There is one notable discrepancy between the theory and the reality, though. Figure 4.9 suggests that the observed quorum versus Q curves should meet where $Q = 0$ and diverge as Q increases, but the curves produced from the experimental data do not have a common y-intercept.

This difference occurs because the predicted curves do not take into account the time required by the robots to gather the vote-messages, only the increased precision in $\tilde{C}_a$ due to the greater number of vote-messages used in its calculation. As the number of vote-messages required to compute $\tilde{C}_a$ increases, so does the time required to compute it. During this period of time, consensus will tend to increase as additional robots enter the deliberating state. This accounts for the curves each having a different y-intercept. The amount by which the deliberating population size will tend to increase during the period of time necessary to receive n vote-messages depends heavily on the specific decision, the domain in which it is made, and the robots themselves, hence its not being accounted for in the theoretical prediction. As consensus increases, so does the number of robots simultaneously estimating the apparent consensus, and so the likelihood of at least one of



(a)



(b)

Figure 4.10: As the quorum threshold is increased, the likelihood of a robot prematurely committing decreases, which means that commitment will tend to be delayed until more robots have entered the deliberating state. This results in an increase in the length of the deliberation phase of a decision. As n or τ is increased, the precision in $\tilde{C}_a$ increases, and so the value of Q has a greater impact on the robots' deliberation time. The time-cost of deliberation is independent of n or τ when Q is zero (the y-intercepts of these plots) because, regardless of the precision with which C_a is estimated, $\tilde{C}_a > Q$ always will be true, and thus quorum always will be satisfied.

them concluding that quorum has been satisfied increases, too. This explains why the experimental curves bear a closer resemblance to the theoretical predictions as Q increases. If the rate at which the working robots were to enter the deliberating state was substantially reduced, so that the time required by a robot to collect teammate opinions became insignificant, a much closer resemblance between Figures 4.8 and 4.9 would be obtained. A detailed description of the derivation of Figure 4.9 is provided in Appendix A.

As the quorum threshold is increased, in turn increasing the observed quorum, the onset of the commitment phase of a decision is delayed until a greater population of robots have entered the deliberating state. This will tend to increase the reliability of a decision, since it will be predicated on a larger number of independent conclusions that the current task has been completed. Increasing n or τ increases the precision of the robots' estimates of apparent consensus, reducing the likelihood of false positive quorum tests, and therefore reducing the likelihood of premature commitment.

Because the commitment phase is delayed until the deliberating population has become sufficiently large, the system will spend a greater period of time in the deliberation phase waiting until quorum is satisfied. Figure 4.10 plots the mean deliberation time of the robots versus the quorum threshold. The deliberation time is computed as the period of time from the beginning of a trial to the beginning of the commitment phase. Increasing the quorum threshold increases the robots' deliberation time because this increases the threshold to which $\tilde{C}_a$ is compared to test quorum. Increasing the precision of $\tilde{C}_a$, accomplished by increasing n or τ, decreases the likelihood of a robot overestimating C_a, which in turn decreases the probability of prematurely committing. This is why the slopes of the regression lines in Figure 4.10 increase with n and τ. Independent of the time required to gather the additional votes necessary to compute $\tilde{C}_a$, increasing n or τ for a given quorum threshold will increase the consensus likely to be present at the time of commitment. Each of the robots will be less likely to compute a false positive quorum test, since $\tilde{C}_a$ will more closely resemble C_a.

The lines all meet at $Q = 0$ (the y-intercept), since this quorum always would be satisfied as soon as a robot entered the deliberating state, independent of the precision with which apparent consensus is estimated. A robot using $Q = 0$ will enter the committed state as soon as it detects task completion. The deliberation time corresponding to $Q = 0$ in these figures is not zero as might be expected, because the manner in which the deliberation time was defined includes the short period of time at the beginning of each trial before any of the robots detected task completion. The average time at which the first robot entered the deliberating state was 123 ± 84 seconds, and it is here that the y-intercepts occur.

Commitment Following Quorum Satisfaction

The role of the quorum test in a unary decision is to ensure that a desired minimum degree of consensus (the quorum) exists in favour of the proposed alternative to the *status quo* before it unanimously is adopted. Once one of the robots believes that apparent consensus has satisfied quorum, it enters the committed state, initiating the commitment phase of the decision. The sole purpose of this final phase is to ensure that the members of a dec-MRS making a decision unanimously adopt the proposed alternative. In general, single-hop global broadcast communication is either unavailable or impractical in a dec-MRS. Furthermore, since the individual robots move about relative to each other (stirring their system), and because they are unaware of each other's identities, explicit message-routing is impractical. The commitment phase thus is organized as a gossip-style process [10], an approach borrowed from sensor networks research [1], and well-suited to dec-MRS.

Robots enter the committed state either because they believe that quorum has been satisfied, or because they receive a commit-message from one of their committed teammates (Figure 4.5). When an uncommitted robot receives a commit-message, it responds with an acknowledgment. Robots start a timer once they enter the committed state, and this is reset whenever they receive an acknowledgment to a commit-message. If a committed robot's timer reaches a preset limit, called the *commitment timeout*, denoted by T_c, it enters the finished state and exits the decision-making process. Once all of the robots have committed, acknowledgments to commit-messages cease, since only uncommitted robots send them. Thus the robots' internal timers no longer will be reset, and they will all exit the decision as each of their timers reach T_c, unanimously adopting the proposed alternative to the *status quo*.

Because gossiping, and therefore the commitment phase, is a stochastic process, the probability of failure (e.g. some robots not receiving a commit-message before all of the committed robots exit the decision) will always be non-zero. In situations in which mutual exclusivity is not a concern, the

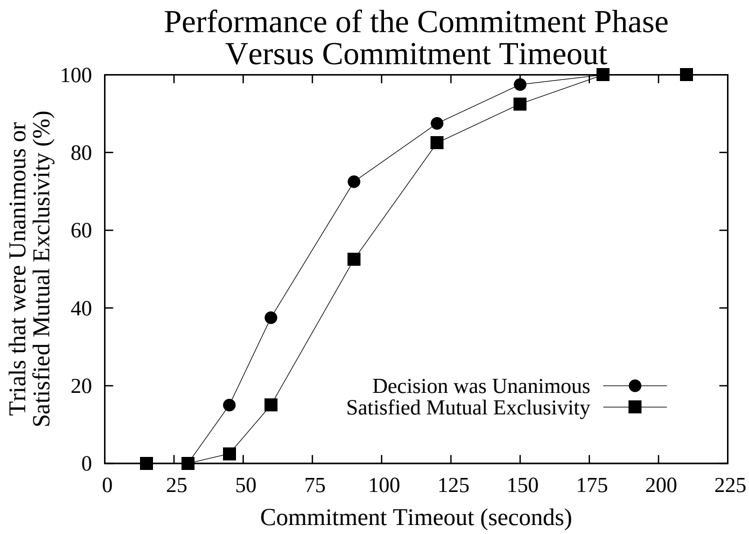


Figure 4.11: The role of the commitment phase is to induce all of the robots to accept the proposed alternative unanimously. Committed robots instruct encountered teammates to commit, and they reset a timer every time an uncommitted teammate is met. Once a committed robot's timer reaches the commitment timeout, it enters the finished state, exiting the decision. As the commitment timeout is increased, the probability of commitment reaching all of the robots increases. In order for mutual exclusivity to be respected, all of the robots must be in either the advocating or committed states before any committed robot can exit the decision.

commitment phase can be said to fail whenever one or more of the robots are left in pre-committed states after all of the committed robots have entered the finished state, as this would cleave a dec-MRS into two: one group having adopted the proposed change to the *status quo*, and the other having not. If we assume that robots halt their work on the current task as soon as they enter the deliberating state, mutual exclusivity will be satisfied only if all of the robots are in the deliberating or committed states before any robot enters the finished state. It would depend on the specific mission at hand how important mutual exclusivity was to its success.

The probability of the commitment phase achieving unanimity, or satisfying the more strict criterion of mutual exclusivity is improved by increasing the commitment timeout, T_c. Several different commitment timeouts were implemented by reparameterizing the generic simulations, and the percentages of these trials that achieved unanimity or satisfied mutual exclusivity are plotted in Figure 4.11. Both curves increase rapidly at first, but diminishing returns are encountered. 100% of the trials ended unanimously and satisfied mutual exclusivity when the commitment timeout was 180 seconds or greater. The median period of time between an individual robot's encounters with teammates in the simulations was 21.75 seconds, with first and third quartiles equal to 10.55 and 39.90 seconds, respectively. Thus a 180 second commitment timeout corresponded to approximately 4 to 17 teammate encounters per robot.

The time required to complete the commitment phase is very linear with respect to the commitment timeout, illustrated by Figure 4.12. This means that the number of messages sent by the robots

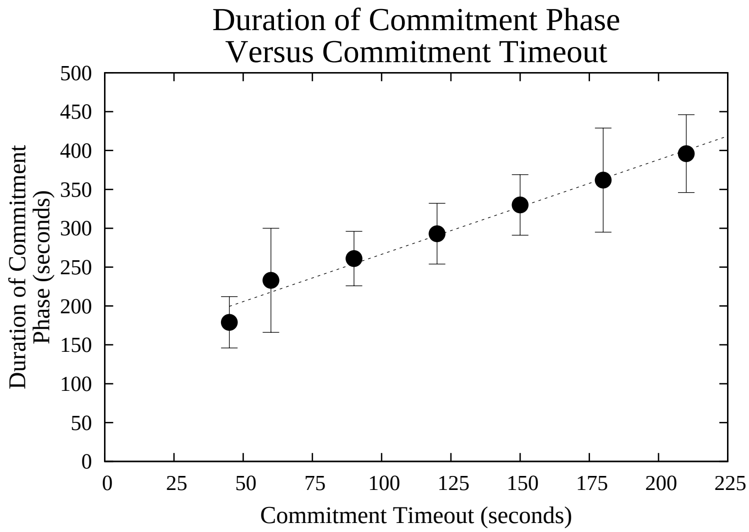


Figure 4.12: Increasing the length of the commitment timeout increases the reliability of the commitment phase of a decision, as illustrated by Figure 4.11, but it also increases the duration of the commitment phase. As shown here, this increase is linear. Because committed robots tell every teammate that they meet to commit (since they cannot discern a teammate's decision state through observation), the longer the commitment phase lasts, the more commit-messages will be sent.

during the commitment phase also increases linearly with commitment timeout, since committed robots send commit-messages to every one of their teammates that they encounter. The longer that a robot remains in the committed state, the more commit-messages it will send.

4.4 Physical Experiments

Experiments with real robots also were conducted to examine unary decision-making for task-completion. As was the case for the simulated trials, the physical experiments were carried out in a domain that mimicked that of blind bulldozing as the task reached its completion.

4.4.1 Environment

Figure 4.4(b) depicts a photograph of the experimental environment in which the physical experiments were conducted. It consisted of a hexagonal arena with sides 2.75 meters in length. Although not circular, it was sufficiently round to approximate a blind bulldozing environment near the completion of the task. Just like in the simulated experiments, the environment itself was static and no actual collective construction was carried out by the robots.

4.4.2 Robots

The robots used in the physical experiments were custom-built differential-drive platforms. Refer to Figure 4.13 for a picture of one. The robots' hulls were circular, 0.26 meters in diameter. Each of the robots utilized an omnidirectional bump sensor to detect encounters with teammates and

Figure 4.13: This figure shows one of the robots used in the physical experiments. Each robot possessed a circular bump sensor that permitted it to detect obstacles. At the rear and top of the robot is an 802.11B radio, which it used to communicate with its teammates when making a group decision.

the site's walls, and they communicated with each other using 801.11B wireless Ethernet. The physical environment was too small relative to the transmission range of 802.11B for inter-robot communication to be local, so local communication in a well-stirred MRS was simulated as follows.

When a robot in either the advocating or committed states encountered an obstacle (they were unable to discern their teammates from the walls), a teammate's IP-address was randomly selected from a list that was supplied to each robot at the beginning of the trial[4]. A query- or commit-message would then be sent to that address via TCP. The recipient of the message would then send its response to the message back to the initial sender's IP-address. In this way, the robots often would exchange messages when they were not physically near to each other, but the net behaviour of their communication was equivalent to local one-to-one communication in a well-stirred MRS. The only difference was that the communication was randomized through IP-address selection rather than random robot motion. Although the physical robots had identity information available to them about their conversation partners, it in no way contributed to their decision-making behaviour.

In order to determine the lengths of their paths, the robots measured the time between reorientations. As was the case in the simulated trials, a robot would enter the deliberating state once the time between two consecutive reorientations exceeded a preset threshold. In all of the physical trials, the robots traveled at a constant speed of 0.2 meters per second. This meant that wheel slippage would introduce noise into the determination of task completion by the individuals, a realistic addition to the trials. The robots were programmed to enter the deliberating state once they had remained in the wandering state of the blind bulldozing behaviour for more than 20 seconds, corresponding to a path approximately 4 meters in length.

[4]Each robots knew its own IP-address, and so they would never send messages to themselves.

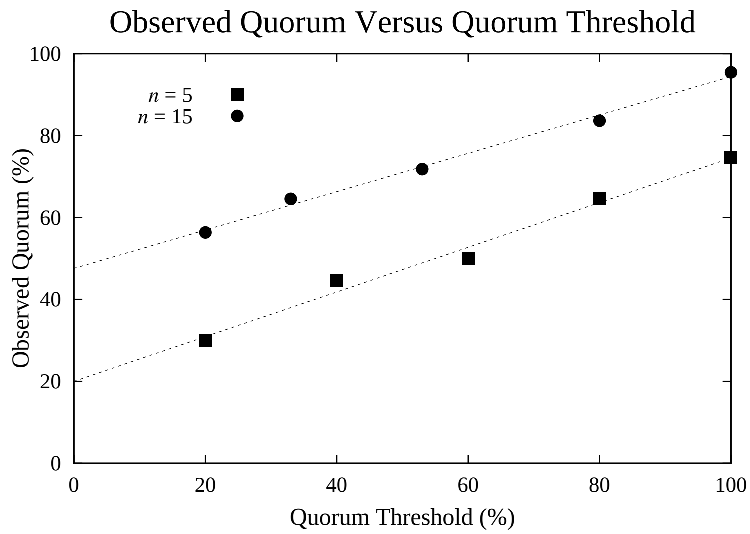


Figure 4.14: This figure plots the observed quorum versus the quorum threshold from the experiments with real robots. The data plotted here are very similar to that shown in Figure 4.8. As the quorum threshold is increased, the observed quorum increases, since the robots are less likely to overestimate C_a and prematurely commit until a sufficient proportion of their teammates also have concluded independently that the blind bulldozing task is complete.

4.4.3 Experimental Trials

Based on the analysis of Chapter 3, and the results of the simulated experiments, it is appropriate to conclude that the analog and digital approaches to consensus estimation generally are equivalent. For this reason, only digital consensus estimation was implemented in the physical experiments. The reader is encouraged to view these results as a test of anonymous consensus estimation in general: analog and digital. A series of experimental trials was run with 11 robots, varying the quorum threshold over two values of n: 5 and 15.

These trials were selected to illustrate the effects of different quorums and varying the precision of consensus estimation on the collective decision-making process in a physical implementation. The commitment timeout was held constant at 60 seconds for all of the trials.

4.4.4 Results and Discussion

The observed performance of the real robots agrees with the simulated results presented earlier and reinforces the conclusions that were drawn from them. Figure 4.14 plots the observed quorum versus the quorum threshold for the physical trials, and Figure 4.15 plots a prediction for this data given the population size of the MRS employed. The latter figure was produced in the same manner as Figure 4.9, the details of which are given in Appendix A. Once again, the observed quorum increases with the quorum threshold. In the physical experiments, the robots tended to enter the deliberating state more rapidly than they did in the simulations, and this accounts for the curves' greater deviation from the theoretical predictions when Q was low when compared with the simulated results. The same

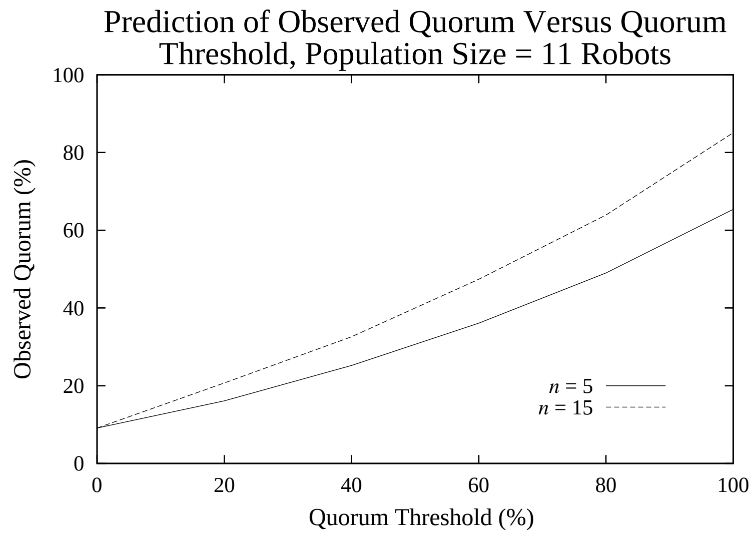


Figure 4.15: This figure plots the predicted relationship between the observed quorum and the quorum threshold for an 11-robot system, the same population size as was used in the physical experiments, for the same values of n that were employed. The actual observed quorum measured from the physical experiments is greater than the theory predicts, particularly for lower values of Q, because the theory does not take into account the time required by the robots to obtain n vote-messages. During this time, additional robots will tend to enter the deliberating state, increasing the observed quorum for a decision. If the rate at which robots were to enter the deliberating state was reduced, the data in Figure 4.14 would more closely resemble that plotted here.

explanation as was given in the discussion of the simulated results applies here, too. If the rate at which the individual robots detected task completion was decreased (i.e. increase the required time between reorientations that would cause a robot to conclude that the task was complete) relative to the rate at which they collected teammate opinions, a closer resemblance between these two figures would be observed.

Figure 4.16 plots the mean deliberation time versus the quorum threshold, Q, and the relationship shown is very similar to that predicted by the simulations (Figure 4.10). When the robots used 15 vote-messages to estimate the apparent consensus, they deliberated for longer than those that used only 5, since they were less likely to commit prematurely due to the greater precision with which they were able to estimate C_a. The regression lines meet when quorum is zero. In the physical trials, it took 31.0 ± 10.4 seconds for the first robot to conclude that the bulldozing task was complete and enter the advocating state. The y-intercepts of the regression lines in Figure 4.16 are 33.8 and 33.5 seconds, which agree with the length of this initial period in which all of the robots still were in the working state.

4.5 Summary

In this chapter, the concept of a unary decision was introduced. In a unary decision, the decision made is whether or not to adopt some new belief in place of the *status quo*. Despite its simplicity,

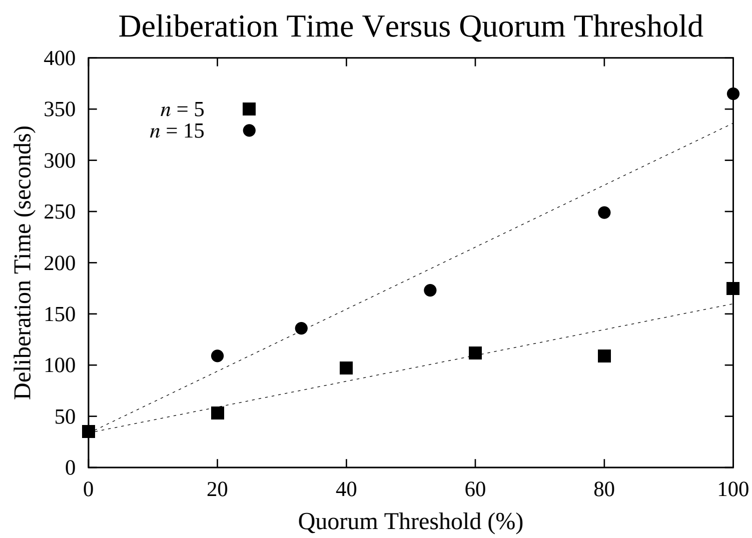


Figure 4.16: The trend of the mean observed deliberation time of the real robots very closely resembles that of the simulated trials, given in Figure 4.10. Increasing quorum increases the deliberation time, since commitment is delayed until sufficient robots are advocating in order to satisfy quorum. Given a particular quorum, increasing the accuracy of the quorum test (n) increases the deliberation time, too, because it raises the precision of consensus estimation, decreasing the chance of premature commitment. The regression lines have a common y-intercept because a quorum of zero is always satisfied, so the deliberation time in this case is independent of n.

a unary decision is very useful cognitive operation. Many problems in robotics request a robot to carry out some action when a prespecified condition is satisfied. A unary decision, then, asks whether or not that condition has been met. In other words, unary decision-making represents a kind of collective if-then operation. Another fundamental problem encountered in dec-MRS is system cohesion; ensuring that all of the robots share common beliefs. Without cohesion, the robots that compose a dec-MRS become more likely to interfere with each other, since they might each base their choice of action on very different beliefs. Unary decisions enable dec-MRS to synchronize their adoption of new beliefs, and these are adopted based on a synthesis of the robots' independent conclusions.

The experimental results presented in this chapter demonstrate that the biologically inspired decentralized decision-making framework of this thesis can be adapted to unary decision-making by dec-MRS. A particularly useful application of unary decision-making is the collective task transition problem, which also was introduced. Analog and digital implementations of anonymous consensus estimation were studied, and they were found to function equivalently. The commitment phase of the decision-making framework, which amplifies an individual robot's detection of quorum and promotes unanimity in a collective decision also was analyzed experimentally, and the results illustrate the tradeoff between the success of the commitment phase and the time required to achieve it as a function of its parameterization. It was shown that both unanimity and mutual exclusivity can be satisfied by the commitment phase.

Simple robots can indeed measure the consensus present amongst their teammates, and use this to determine whether or not a proposed alternative to the *status quo* should be adopted. Traditionally, each robot in a dec-MRS bases its decisions on its own perception of the world. By basing these decisions on the satisfaction of a quorum, the robots can leverage their redundancy rather than be victimized by it, an inclusion that is computationally within reach of even the most simple mobile robots. Even without the unifying commitment phase, quorum testing on its own is sure to have many applications in decentralized intelligent systems.

The commitment phase of a decision is triggered once one of the robots believes that quorum has been satisfied. This phase amplifies the belief that quorum is satisfied by inducing the remainder of the system's robots to commit as well, and thus a unanimous decision is made. It is through coupling of the two behaviours — consensus estimation and commitment — by the quorum test whereby intelligent cooperative decision-making emerges.

In a unary group decision, the decision makers passively measure consensus, its value set by the number of robots that independently have reached the same conclusion. In the next chapter, the decision-making framework is extended to solve the best-of-N decision-making problem by adding iterative recruitment to the robots' deliberation. This addition enables them not only to estimate consensus, but also to influence it through their active recruitment of their teammates to the various known alternatives in a decision.

Chapter 5

Collective Best-of-N Decision-Making: The Site Selection Problem

In this chapter, collective best-of-N decisions are demonstrated utilizing the proposed decision-making framework in a site selection environment. The approach used here extends the simple unary decision-making framework presented in the last chapter by introducing positive feedback into the deliberation phase via iterative recruitment. This addition enables several different alternatives to be compared by the robots so that the best one can be identified and selected. Experiments are presented that were carried out in both simulation and with real robots. Their results demonstrate that best-of-N decision-making is practical in a dec-MRS using the proposed framework of this thesis, and that accurate decisions can be made, even in the presence of noisy sensing on the parts of the individuals.

5.1 Introduction

In the previous chapter, it was shown that a dec-MRS could make an intelligent group decision based on the collected opinions of its robots through their local one-to-one interactions. The robots accomplished this feat by estimating the consensus present amongst their teammates in favour of some proposed alternative to the *status quo*, adopting it only once the consensus reached a preset quorum. These unary decisions were decisions of a yes/no nature. That is, the robots decided whether or not a single proposed alternative should replace the *status quo*, not what the alternative should be.

Unary decisions have many applications, such as collectively making a decision about the state of a group task (*e.g.* whether or not the task has been completed), but not all decisions can be captured by that model. A much more general approach to decision-making is the best-of-N framework, in which one of several proposed alternatives must be selected unanimously by a group. Rather than "whether or not", best-of-N decisions ask "which one", and it is these decisions that constitute the focus of this chapter.

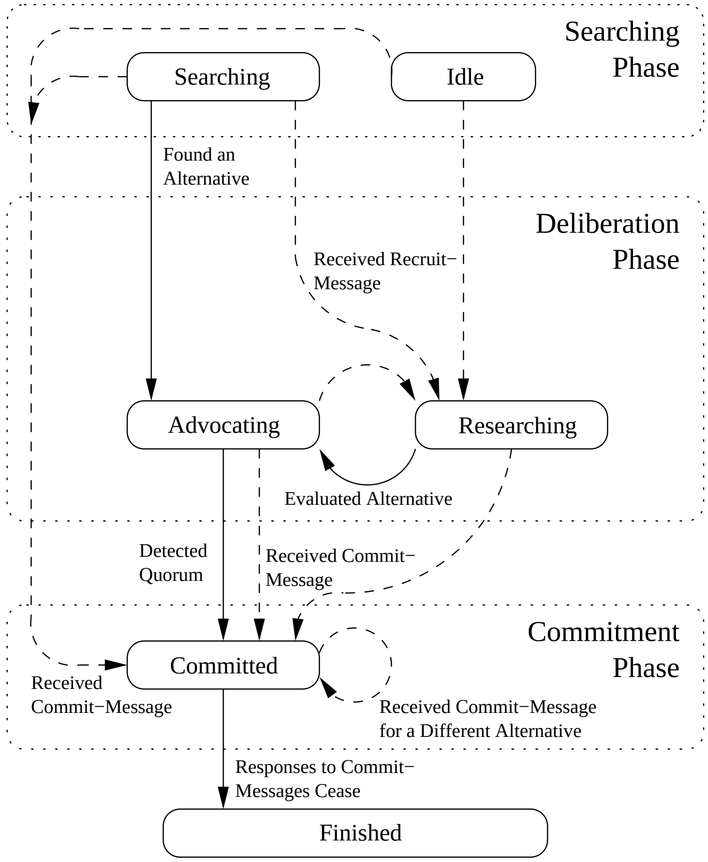


Figure 5.1: This flowchart illustrates the best-of-N decision-making framework, which is organized into three phases. In the initial searching phase, robots search for candidate solutions, called *alternatives*. Upon finding an alternative, a robot will enter the advocating state favouring it. The advocating robots iteratively recruit each other at a rate determined by their opinions of their favoured alternatives' qualities. Better alternatives induce more frequent recruitment, and so over time, the proportion of the system that favours the best alternative will tend to increase. Eventually, one of the advocating robots will conclude that the proportion of its teammates that also favour its alternative has reached the quorum, which triggers the commitment phase. In this final phase all of the robots commit to the quorum-satisfying alternative. Once no more uncommitted robots can be found they exit the process, having unanimously chosen the best of the alternatives that was found.

In some circumstances, the candidate alternatives in a best-of-N decision are known *a priori*. For example, consider a swarm of robots encountering a fork in a road along which they are traveling. The decision facing this system would be to choose the appropriate path from the two presented to it. In this best-of-N decision, the two alternatives would be known at the outset, but it would remain to identify the best one. However, there are many situations in which the available solutions to a problem will not be known *a priori*, and so the solutions must be found as part of the decision-making process. In such cases, the alternatives must be discovered, researched, and then one chosen. It is this sort of best-of-N decision-making that is addressed in this chapter.

Best-of-N decision-making requires a somewhat different approach than to unary decision-making. In a unary decision, each robot makes an individual decision about whether or not the proposed alternative should replace the *status quo*. An individual that has decided that the status quo should be replaced collects the opinions of its teammates to determine how many of them also have arrived at the same conclusion. Once a robot believes that the proportion of its teammates that agree with it has reached a preset quorum, the commitment phase is triggered and the decision is completed. In other words, unary decision-makers make independent decisions and then wait for a prespecified proportion of their teammates to do the same. The individual deliberators in a unary decision do not influence other robots' opinions; they observe them. In a best-of-N scenario, this sort of behaviour would lead to poor decisions and often stagnation. For example, if each individual decided to favour a different alternative, each of them would observe an apparent consensus of zero. A non-zero quorum would never be satisfied if this were to occur, and so the decision would stagnate. Furthermore, an obvious but poor alternative might be discovered by a disproportionately large number of robots and get selected, even though better alternatives had been found, albeit by fewer robots.

Instead, the various alternatives should be compared somehow. Unfortunately, this would be an expensive operation if each and every alternative was considered by every robot. It also would be wasteful, since much time and energy would be spent by every robot communicating/considering alternatives that ultimately would not be selected. The iterative recruitment behaviour of honeybees and *Temnothorax* ants described earlier in this thesis, however, enables a set of candidate alternatives to be compared by a decentralized system *at the system level*. Refer to Figure 5.1 for the following discussion. After the initial searching phase, several different alternatives will be known to a dec-MRS, and each robot will favour at most one of them. Some of the robots might not find an alternative, or perhaps they did not participate in the search. These robots are represented by the idle state included in the searching phase.

The deliberation phase, which consisted of a single behavioural state (deliberating) in the unary decision-making algorithm (Figure 4.5), is expanded in a best-of-N decision to contain two states: advocating and researching. It is iterative recruitment (Section 3.3) during the deliberation phase that collectively compares the known alternatives, identifying the best one. An advocating robot is said to favour an alternative, and recruits other robots to its favoured alternative as they randomly are encountered. Once recruited, a robot enters the researching state to evaluate the quality of the alternative to which it was recruited. After it has ascribed a quality to the alternative, a researching robot then enters the advocating state favouring the alternative, and will recruit additional robots to this alternative. The rate at which the advocating robots attempt to recruit their teammates is tied to their opinions of their favoured alternatives. The higher in quality an advocating robot believes its alternative to be, the more frequently it will attempt to recruit others. Therefore, the better alternatives will tend to increase in popularity while the poorer ones will be forgotten altogether.

Simultaneously, advocating robots also estimate apparent consensus and compare their estimates to the quorum threshold in order to test quorum. Unlike the deliberators in a unary decision, advocating robots in a best-of-N decision cannot assume that the other advocating robots all favour the same alternative. Each advocating robot estimates apparent consensus only for the particular alternative that it favours. If it is recruited to favour a different alternative, it will compute $\tilde{C}_a$ only for the newly favoured one. Once one of the robots believes that the apparent consensus for its favoured alternative has reached the preset quorum threshold, Q, that robot will commit to its alternative, and enter the committed phase of the decision.

Committed robots instruct every teammate that they encounter to commit as well, just like in a unary decision, eventually exiting the decision when they no longer receive acknowledgments to their commit-messages. Very rapidly, the entire dec-MRS will be induced to commit to the same alternative, thus unanimously selecting one of the alternatives found during the initial search. As a direct consequence of the positive feedback in the deliberation phase, the best alternative will be the one most likely to induce a robot to commit first, and so the commitment phase tend to will bring about the unanimous adoption of the best alternative. In this way, the best of the N alternatives found by a dec-MRS will be selected by the system as a whole.

However, because there are multiple alternatives in a best-of-N decision, and because each robot's estimate of apparent consensus will contain some error, it will be possible for more than one of the alternatives to induce commitment. A committed robot in a best-of-N decision will change the alternative to which it is committed if it receives a commit-message specifying a different alternative. When more than one alternative induces commitment, a period of attrition will follow, and the alternative with the greatest number of committed robots will tend to be selected in the end. That particular alternative is most likely to be the one that induced commitment first, which again will tend to be the best one.

5.1.1 The Site Selection Problem

Best-of-N decision-making can be applied to many different problems, but to demonstrate its performance in this work, the site selection domain is used. In this problem, introduced in [57] as "collective relocation", a dec-MRS has decided that its current home base has become inadequate (perhaps with a unary decision), and so a new one must be found. The robots do not know of any sites for a new base *a priori*, but several are located in the surrounding environment. These must be discovered by the robots during the initial search, which populates a menu of alternatives from which exactly one will be selected. Figure 5.2 illustrates a simulated site selection environment.

The deliberation phase identifies the best alternative in the decentralized menu of alternatives via its iterative recruitment. A recruited robot must travel to the site communicated by its recruiter in order to determine the site's quality for itself. Site quality could be a function of several different cues depending on the specific problem at hand. For example, in a robotic mission to the planet Mars, a dec-MRS might have to set up infrastructure for a later human mission. One task in such a mission might be to erect a solar array to generate electricity [40]. In this case, the robots would

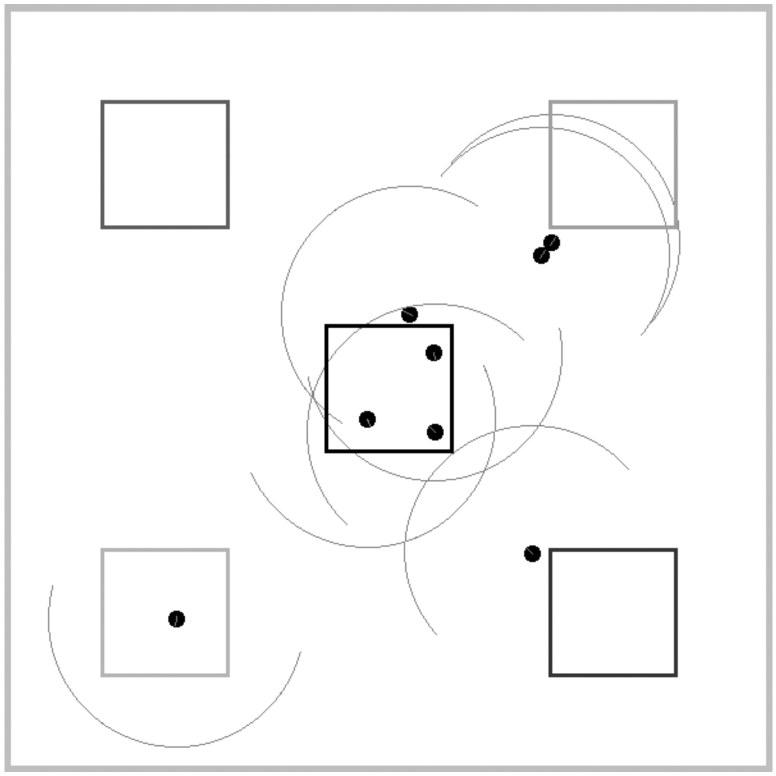


Figure 5.2: This figure presents a screenshot from a simulated best-of-N decision-making experiment in the site selection domain. The black square in the center of the environment is the robots' initial home base, and the squares in the corners are candidate sites from which the robots must select a new base. The small black circles are the robots themselves, and the arcs represent the ranges of their vision. One of the robots in this scene favours the upper right site, and is leading a teammate that it has recruited to it so that the recruit can inspect the site for itself. The rate at which the site-favouring robots recruit is based on their opinion of site quality, so the best site will tend to attract recruits more rapidly than the others, making it the most likely site to be selected by the decision's end.

need to find a site in their environment that would receive the most sunlight, while also being stable enough to support the array. In a security application, a system might have to choose locations at which to place surveillance cameras or other such equipment in order to maximize the equipment's effectiveness. Regardless of the specific scenario, a researching robot would have to evaluate the candidate site to which they were recruited and determine for itself the site's quality.

Being a domain with spatially distributed alternatives (the sites), both on-swarm (Section 3.4.2) and off-swarm (Section 3.4.1) quorum testing could be employed. When quorum is tested on-swarm, the robots estimate apparent consensus by explicitly requesting vote-messages from their their teammates as they are encountered. When quorum is tested off-swarm, robots estimate the number of robots that also are present during a visit to a favoured site. Both of these strategies were investigated in the experiments presented in this chapter

Independent of the manner in which quorum is tested, once a robot believes its favoured alternative has satisfied quorum, it will commit to the alternative, beginning the commitment phase of the collective decision. As the committed robots exit the decision by entering the finished state, they relocate to the site to which they most recently were committed, adopting it as their new base. At the end of a successful site selection decision, the robots all will have relocated to the best site that was found during the initial search.

5.2 Simulated Experiments

Analytical investigation of this decision-making behaviour is difficult, and therefore an empirical approach is employed. In this section, a series of simulated experiments are presented. These experiments investigate the effects of population size, quorum, and the mechanism of iterative recruitment on best-of-N decision-making behaviour in the site selection domain. Additionally, *off-swarm* quorum testing is demonstrated in the simulations (refer to Section 3.4.1 for a discussion of off-swarm quorum testing). This is an approach to quorum testing that closely resembles the reported behaviour of the honeybees and ants that inspired this work.

5.2.1 Environment

As described earlier, the site selection environment contains several sites, and these are the alternatives for the robots' best-of-N decisions. A screenshot from one of the simulated trials is given in Figure 5.2. Like the earlier simulations, these were implemented using the TeamBots [85] MRS simulator.

The simulated environment was a square enclosed region measuring 24 meters per side. In its center was the robots' initial home base, where recruitment and commitment actions took place. Equidistant from the base were either two or four candidate sites, depending on the trial. Each of these sites had a different quality that the robots were able to determine by visiting it. Site quality measurement was error free in the simulations, meaning that every robot that visited a given site would ascribe to it the same quality.

5.2.2 Robots

The robots in the simulated experiments were the same as those in the unary decision-making trials, except that they were able to see further to allow them to search for sites and test quorum. Each robot's vision had a range of four meters and a 180 degree field of view with its center aligned with the robot's front. The areas visible to each robot are represented by the semicircular arcs in Figure 5.2. All of the inter-robot communication was local and one-to-one.

Four of the robots began each trial in the searching state, while the rest of their teammates remained in the idle state, wandering about the initial base. The searching robots, also referred to as *scouts*, would search the environment for sites by following random walks, moving in a straight line for a random period of time followed by a reorientation to a random heading. The robots were confined by the boundaries of their environment. When a candidate site was found by a scout, it would head towards it and ascertain its quality. Next, a robot would enter the advocating state favouring that site, and spend a short period of time wandering about it in order to test quorum. Finally, the robot would return to the initial base to recruit one of its teammates. Clearly, the size of the menu of alternatives over which the robots would deliberate depended on the number of robots that participated in the initial search. Each dec-MRS deployed the same number of scouts (four) so that the observed decision-making ability of each MRS would not be affected by the ability of a given system to conduct a more or less thorough search. The smallest system contained only four robots, so a searching population of this size was used in every one of the simulated trials.

Advocating robots periodically return to their favoured sites to wander about and test quorum before heading back to the base to recruit again. Once an advocator believes that its site has satisfied quorum, it enters the committed state (still favouring its site), returns to the base and instructs the rest of its teammates to commit to its site as well. The details of quorum testing, recruitment and commitment in the simulated experiments are provided in the next three sections.

Off-Swarm Quorum Testing

Much of this thesis focuses on quorum testing that is carried out *on-swarm* (Section 3.4.2), in which the advocators actively query the robots that they encounter to ascertain what proportion of their teammates agrees with them. This is because on-swarm quorum testing is practical in a wider variety of domains. However, in the simulated best-of-N decision-making experiments, *off-swarm* quorum testing (Section 3.4.1) was investigated. When a robot tests quorum for its alternative off-swarm, it does so at some unique, well-defined location that all of the robots agree corresponds to the specific alternative favoured by that robot. This means that off-swarm quorum testing only is practical in certain spatial domains like site selection, since each alternative *is* a unique, well-defined location.

During site selection, the advocating robots spend some proportion of their time visiting the sites that they favour. These periodic visits occur immediately after an advocator has led a recruited robot to its site (discussed in the next section), or after it has been unable to find a teammate to recruit at the initial base. While at its favoured site, the robot counts the number of other robots that also are there. The robots all are identical in appearance, so to avoid counting teammates more than once in a quorum test, the advocators compute their tallies of teammates as the greatest number of robots that they observe simultaneously during a particular visit to a site. A robot that is testing quorum assumes that the other robots that are visiting the site are there because they also favour it. The size of the population visiting a site is correlated to the site's popularity within a dec-MRS, since all of the robots that favour particular a site will visit it periodically. An off-swarm quorum test concludes that quorum has been satisfied when the estimate of a site's visiting population is greater than or equal to the quorum threshold. Note that, because off-swarm quorum testing is based on an absolute count of robots rather than a proportion, the quorum in this case is not a real number $\in [0, 1]$, but instead an integer $\in [0, N-1]$[1], where N is the population size of the dec-MRS.

Iterative Recruitment

It is the process of iterative recruitment that modifies the proportion of the robots in a dec-MRS that favours the various known sites, promoting better sites over poorer ones. This in turn increases the likelihood of a robot that favours the best known site believing that quorum is satisfied, and therefore committing to it before a robot that favours a poorer site does. Therefore, the best site found will be the one most likely to be selected by the entire dec-MRS. Robots in the idle state can be recruited, as can robots in the advocating state. After an advocator has returned to the base but before it attempts to recruit one of the robots there, it will delay for a certain period of time and wander about there. It is during this delay prior to recruiting that advocating robots are themselves recruitable.

When a robot recruits a teammate, it leads its recruit to its favoured site in a follow-the-leader fashion (refer to Figure 5.2 for an example). This mimics the tandem-running of the *Temnothorax* recruitment behaviour, which allows a recruit to learn the location of the site to which it is being led [32]. In the context of a general best-of-N decision, this behaviour is analogous to an advocator *explaining* its favoured alternative to its recruit so that the recruit can evaluate the alternative for itself. Once an advocator has begun to recruit a teammate, the advocator no longer will be recruitable until it next returns to the initial base. This prevents a robot from abandoning a recruitment already in progress if another advocator attempts to recruit it. In the simulations, four versions of iterative recruitment were investigated, referred to in this thesis as *restraintive*, *discriminative*, *hybrid*, and *unbiased* recruitment. They implement the positive feedback of the iterative recruitment process in different ways, modifying the behaviours of both the advocating robots and those recruited by them.

Restraintive recruitment is a more descriptive name for the iterative recruitment strategy as it was described earlier in this thesis in Section 3.3.2. Under this model, it is the amount of time that an advocator delays prior to recruiting a teammate that is modulated to promote better sites over poorer ones. The *worse* an advocator believes its favoured site to be, the longer it will delay. As a result, recruitment towards the better sites will tend to be more frequent, and those robots advocating for poorer sites will be more likely to be recruited (to better sites) since they spend a greater proportion of their time delaying in a recruitable state at the initial base. When a robot is recruited under this model, it immediately forgets any previously favoured alternative and adopts whatever site it was led to by its recruiter as its new favoured alternative. Because the recruitment tends to shift robots that favour poorer sites to better ones, the poorer alternatives are the most likely to be forgotten by a dec-MRS[2], increasing the apparent consensus in favour of the better ones.

Contrasted with the restraintive approach is discriminative recruitment. Here, an advocator's behaviour upon returning to the initial base is independent of the quality that it ascribes to its favoured

[1] The robots do not include themselves in their counts, so the largest quorum that could be satisfied is $N-1$.

[2] The memory of a dec-MRS can be thought of as the union of its members' memories. Once the last robot favouring a particular site forgets it, the MRS as a whole forgets it.

alternative. They delay for a constant period of time before attempting to recruit a teammate, and so discriminative advocators are equally likely to recruit teammates independent of the quality of the site that they favour. They also are equally likely to be recruited, since they all spend an equal proportion of their time in a recruitable state. Preference for better alternatives is expressed instead by the recruits themselves. Rather than automatically forgetting about some previously favoured alternative, a recruited robot waits until it has evaluated the site to which it was led. If a recruit believes that the site to which it has been led to is worse than one that it previously favoured, the proposed alternative will be rejected and the recruit will continue to favour the site which it had favoured prior to the attempted recruitment. On the other hand, if the proposed alternative is at least as good as the one previously favoured by the recruit, it will be adopted and the old one forgotten. In short, a discriminative robot cannot be recruited to a site that it believes is inferior to one that it already favours. It is important to note discriminative recruitment relies on the individual robots to make direct comparisons of alternatives in order to identify the best one.

Hybrid recruitment is a combination of the restraintive and discriminative models. A hybrid advocator behaves according to the restraintive model by delaying for a period of time determined by the quality of its favoured site prior to recruiting a teammate, and recruited hybrid robots will refuse to favour a proposed alternative if it is inferior to one favoured prior to being recruited.

Finally, decisions might arise in which the candidate alternatives differ in some manner undetectable to the robots, or perhaps several equally good alternatives are discovered. In order to simulate this scenario, the unbiased model of recruitment is used. Unbiased recruits always accept whatever site they are led to as though they were following the restraintive model, and unbiased recruiters delay for a constant period of time prior to recruiting, independent of their favoured alternative's quality. Therefore, no preference is expressed for any of the sites by the individual robots. The question regarding unbiased recruitment is whether or not the decision-making process will terminate at all, or if stagnation will ensue.

Commitment

Once a robot believes that its favoured alternative has satisfied quorum, it commits to its alternative, beginning the commitment phase of the decision-making process. At the time of commitment, the robots likely will be spread across the environment. Some might be at the initial base, whereas others might be visiting other candidate sites. However, uncommitted robots periodically return to the base in order to recruit more teammates.

Committed robots therefore return to the initial base and instruct every robot that they encounter there to commit to the site favoured by the committed robots. In the simulated experiments, committed robots transmit the location of the site to teammates relative to the initial base's location. When an uncommitted robot receives a commit-message, it responds with an acknowledgment and then goes to the specified site. Every time a committed robot receives an acknowledgment it resets

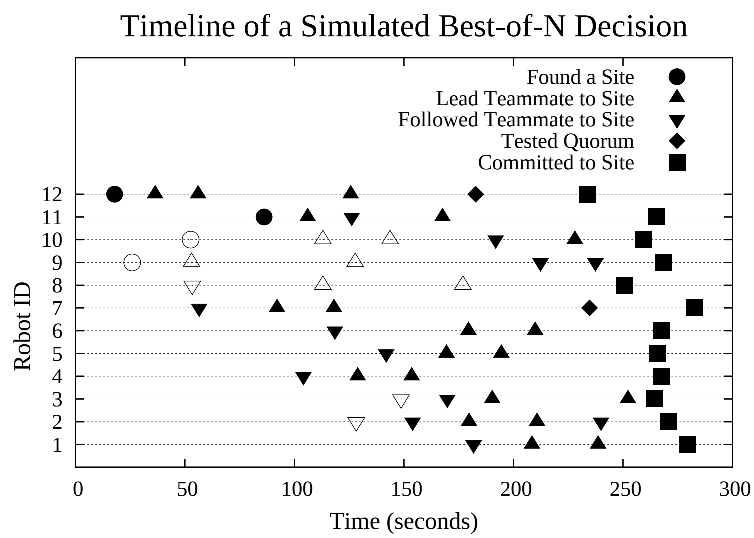


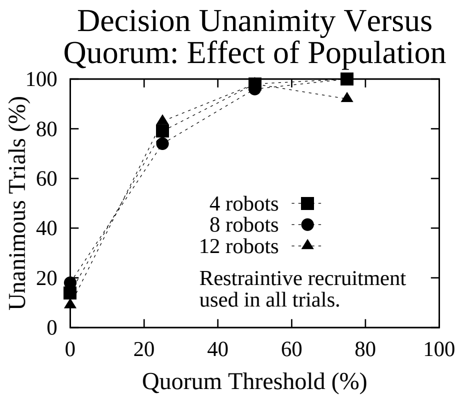
Figure 5.3: This figure presents a timeline of one of the simulated best-of-N decisions. The history of each robot is given by the sequence of symbols along the corresponding timeline. Solid and hollow symbols indicate events regarding the better and poorer sites, respectively (this particular trial compared only two sites). Once a robot found a site, the robot began to recruit teammates to it. Note that the robots that favoured the better site recruited more frequently. Over time, robots that favoured the poorer site were recruited to favour the better one, and eventually quorum was satisfied for it. After this occurred, commitment flooded throughout the dec-MRS, resulting in the unanimous adoption of the better site. This timeline presentation was inspired by a similar figure in [50].

a timer. Once a committed robot's timer reaches the preset limit (the commitment timeout, T_c), it leaves the initial base and heads to the site that triggered commitment, and exits the decision-making process.

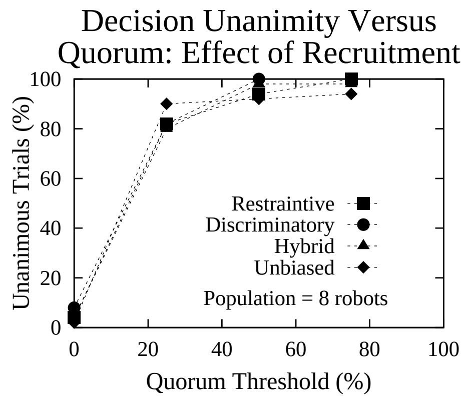
5.2.3 Experimental Trials

A series of simulated experimental trials was carried out to investigate the different recruitment strategies. The effect of quorum and dec-MRS population size also were examined by these experiments. Regardless of the population size of the dec-MRS, the number of robots that acted as scouts always was four. Robots that did not act as scouts remained in the idle state at the initial base, waiting to be recruited into the process.

Dec-MRS composed of 4, 8 and 12 robots were run in a two-site environment, and quorum was varied from zero to 75% of total system population. These trials all employed restraintive recruitment and were repeated 100 times. Restraintive, discriminative, hybrid and unbiased recruitment were implemented with an 8-robot system in a four-site environment, with each trial repeated 50 times. Quorum for this second set of experiments also was varied from zero to 75% of total system population.



(a)



(b)

Figure 5.4: It is important that a collective decision is unanimous. These graphs plot the percentages of the simulated trials that ended unanimously, regardless the particular site that was selected. In general, population size and the specific model of recruitment do not affect the ability to achieve unanimity. However, the likelihood of unanimity increases with the quorum threshold, because because a greater quorum makes commitment to multiple sites less likely.

5.2.4 Results

The general behaviour of the trials followed that outlined by Figure 5.1. Following an initial period of searching, the scouts found sites and entered the advocating state and began to recruit the idle robots and each other. As the iterative recruitment proceeded, robots favouring poorer sites were recruited to favour better ones, and the population favouring the best site found tended to increase over time. Eventually, one of the advocators would observe a quorum of robots while visiting its favoured site, which would induce it to commit. The remainder of the robots would be instructed to commit soon afterwards as the quorum-observing robot instructed them to do so. Often, once one robot had committed to its site, other advocators testing quorum at the same site would see a sudden influx of teammates that had received a commit-message. The other robots that were testing quorum at the site often would commit, because they also would observe quorum. They too would return to the initial base to tell robots there to commit. Thus a single commitment would induce other robots to observe quorum, too, leading to a chain-reaction of robots observing quorum and committing in addition to the normal commit-message flooding behaviour observed during the commitment phase. A timeline graphically depicting the history of a typical simulated best-of-N decision is given in Figure 5.3.

Decision Unanimity and Stagnation

As was the case for unary decision-making, best-of-N decisions can be characterized by their accuracy, efficiency and the time required to make them. Figure 5.4 plots the percentage of the decisions that achieved unanimity as a function of quorum. At the end of a unanimous decision, all of the robots had relocated to the same site. A decision might not achieve unanimity for several reasons.

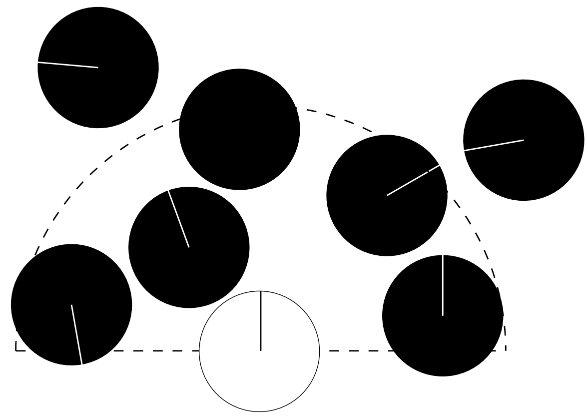


Figure 5.5: This figure illustrates how a robot's visual field of view impacts its ability to test quorum using the off-swarm method outlined in the text. While visiting its favoured site, a robot will compute the number of its teammates also there as the largest number of other robots it was able to observe simultaneously. In this example, the white robot would believe that only five other robots were present, since the other two are outside of its field of view, indicated by the dashed semi-circle. In practice, this means that larger quorums are less likely to be observed by the advocating robots, delaying the onset of commitment, or resulting in stagnation altogether.

If too low a value was used for the quorum threshold, commitment might occur before all of the scouts had found a site, and thus some of them would not return to the initial base in time to receive a commit-message before their committed teammates all relocated to the selected site. Also, when quorum was less than 50%, it was possible for multiple sites to satisfy quorum and induce commitment, which increases the likelihood of a split decision. Increasing quorum addresses both of these modes of failure.

Furthermore, there was a finite probability that none of the robots would commit before the end of a trial, even though a sufficient number of the robots favoured one of the sites. This is a problem specific to off-swarm quorum testing. For quorum to be satisfied, an advocating robot must observe a minimum number of robots while visiting its site. Even if 100% of the robots advocate for the same site, they are unlikely to present there at the same time, so an advocating robot might never have the opportunity to observe a quorum during any of its visits to its site. Even if a sufficient number of robots does assemble simultaneously at a particular site, the difficulty in accurately counting them all increases with the size of the visiting population (*i.e.* it is easy to count all of the robots at a site when there are only three, but harder to do so when there are ten, and so on), because it is harder to fit a large number of individuals into the fixed visual field of view of a robot. For example, to observe a 75% quorum, a member of the 8-robot system would have to observe six robots simultaneously, whereas a member of the 12-robot system would have to simultaneously observe nine robots to do so. This problem is illustrated by Figure 5.5. Both of these problems increase the likelihood of stagnation, and the latter worsens as the population size increases. Stagnation due to the latter phenomenon accounts for the dip in the 12-robot system's ability to make decisions when quorum was increased to 75% seen in Figure 5.4(a) [58].

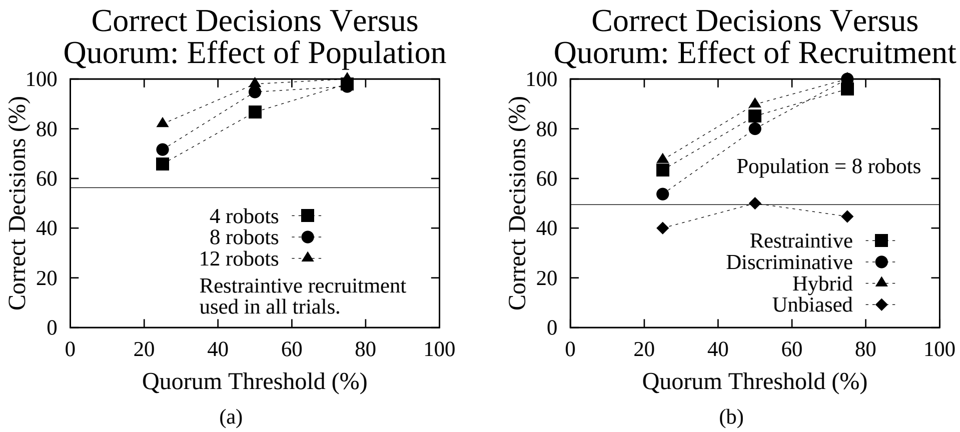


Figure 5.6: As quorum is increased, the ability of the robots to make correct decisions (in which best of the sites found by the scouts is selected at the decision's end) increases with quorum. Quorum specifies how much iterative recruitment is sufficient; once a quorum of robots is found to support a particular site, the system concludes that sufficient deliberation has transpired. Note that increasing the population size of a system also increases its ability to make correct decisions, since larger systems are less impacted by the occasional recruitment away from the best site. The model of iterative recruitment has little effect on the decision-making ability of a system, as long as it is biased in some way to so that recruitment towards the best site is the most likely.

The manner in which iterative recruitment was implemented appears to have minimal impact on the likelihood of achieving unanimity. All three of the biased models achieve similar performance as indicated by Figure 5.4(b). The slightly lower performance of unbiased recruitment also is attributable to stagnation, since it takes longer for the unbiased robots' deliberation to satisfy quorum for one of the sites.

Ability to Make Correct Decisions

Of course, it is important to make good decisions, not just unanimous ones. Figure 5.6 plots the percentage of the unanimous decisions that also were correct. Because each robot was able to determine the quality of its favoured site without any error, the correct decision was made by a system when it unanimously chose the best site that it was able to find during the searching phase. In both of these plots, the horizontal line indicates how often the robots' decisions would have been correct if they were made completely at random. Note that these baselines are not at 50% and 25% in the two- and four-site environments, since there is no guarantee that the robots will find every site in a given trial. It is iterative recruitment that promotes correct decisions in the proposed decision-making framework, and so the more iterative recruitment takes place, the more likely a correct decision becomes. This is controlled by the value of quorum. Increasing quorum increases the certainty with which a particular site could be labeled as "best" by a robot, since that conclusion would be based upon a greater consensus when higher quorums are employed. As a result increasing quorum will tend to increase the ability of a system to make a correct decision, and this precisely is what is demonstrated by Figure 5.6.

Note that the decision-making ability of a dec-MRS increases as the system's population size increases, too. In a small system, one or two robots recruited from a good site to a poorer one is much more likely to result in commitment to the poorer site than would be the case with a larger population. Consider the effect on consensus in a four-robot system when a robot favouring one site is recruited to favour another. The consensus in favour of the site that the robot originally favoured will decrease by 25 percentage points, and the consensus in favour of its newly favoured site will increase by the same amount. This is a significant change. On the other hand, consider the same scenario in a 12-robot system. The change in consensus in this case would be only 8.3 percentage points. The greater the number of robots that take part in a decision, the less of an impact each individual's actions and opinions will have.

Turning to the manner in which iterative recruitment is implemented and how this can impact the ability of a dec-MRS to make a correct decision, examine Figure 5.6(b). Not surprisingly, unbiased recruitment performs about as well as random chance. The three biased methods all see an increase in the ability of a system to make the correct decision as quorum is increased. There appears to be no advantage to direct comparisons by the individual robots, evident in the performance of discriminatory recruitment as compared to restraintive and hybrid recruitment. There is a slight interaction between the discriminative strategy and off-swarm quorum testing, since a discriminative recruit might visit a site and reject it, yet still be included in another robot's quorum test before it leaves, but it is unlikely that by removing this phenomenon that any substantial improvement in its performance would be realized.

Discriminative recruitment intuitively seems as though it should be the best approach, since it prevents recruitment from proceeding away from the best site. However, this characterization implicitly assumes that the individual robots can be relied upon to compare two alternatives accurately. Even when the robots possessed perfect sensing, which was the case in these simulations, discriminative recruitment performs no better than either the hybrid or restraintive approaches. Note that, on its own, the alternative quality dependent delay in hybrid recruitment (identical to that of restraintive recruitment) was so effective at comparing the sites that the discriminative component of the hybrid strategy largely was extraneous.

The Focus of Deliberation

When a best-of-N decision is made, two basic operations are employed: recruitment and commitment. The switch from the former to the latter occurs when one of the robots observes quorum. As in unary decisions, the role of commitment is to promote unanimity, whereas iterative recruitment in the deliberation phase promotes good decisions. At the end of a decision, all of the attention that the unselected alternatives received can be considered a waste of time and energy. A perfect decision-

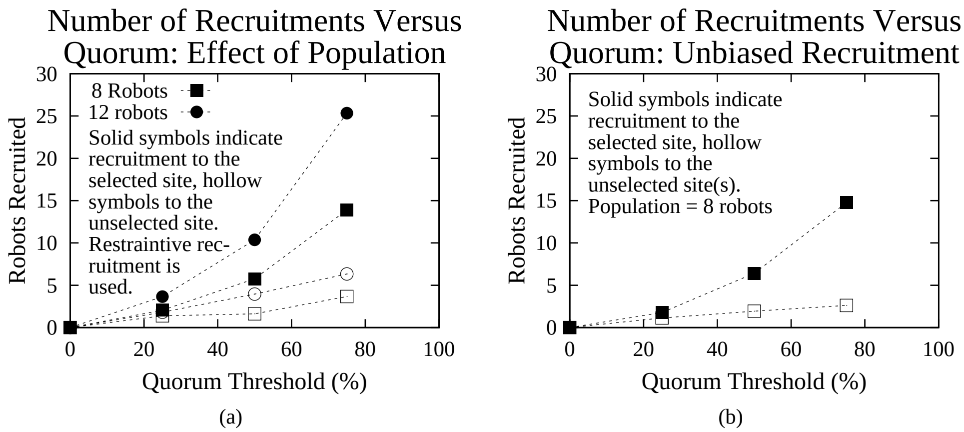


Figure 5.7: The deliberation phase of the decision-making framework compares sites by recruiting additional robots to inspect them. Ultimately, recruitment towards a site that is not selected by a system represents a waste of time and energy, and so a good decision-making algorithm should give most of its attention to the site that ultimately is selected. The plots in this figure illustrate that this is the case for the proposed decision-making framework. As quorum is increased, the selected site is seen to attract more recruitment, but recruitment to the unselected site remains minimal. Some of the system configurations are omitted from these plots to avoid clutter, but all of them follow the pattern of those shown.

making algorithm would ignore alternatives that ultimately will not be selected[3]. In reality, though, some attention must be paid to each alternative, but the less paid to unselected ones the better. The number of robots recruited to a particular site is a good measurement of the collective attention that the site attracted. Figure 5.7 plots the number of recruitments to the sites selected and unselected. The same behaviour was observed in all of the systems regardless of population. As quorum is increased, which is equivalent to demanding greater accuracy from a dec-MRS, significantly more attention is given to the site that is selected in the end, whereas the increase in recruitment towards the unselected site(s) is much less. This shows that iterative recruitment produces efficient deliberation. As the population size is increased, the total recruitment increases, but still the unselected site largely is ignored. Unbiased recruitment exhibits this behaviour, too, which might seem somewhat surprising at first. However, even though the probability of an individual robot recruiting is independent of its favoured alternative's quality, the *probability of being recruited* will be biased towards one of the sites, simply because more robots are likely to support one than the other(s) through random chance[4]. For example, if two thirds of a robot's teammates supported site-A and only one third supported site-B, the robot would be twice as likely to be recruited by a supporter of site-A. This in turn would increase the probability of other robots being recruited to site-A.

[3]This notion of a perfect decision-making algorithm can be compared to the imaginary nondeterministic function employed in complexity proofs.

[4]One might argue that an equilibrium exists when an equal number of robots favour each known site. However, this is an unstable equilibrium, and a system soon would be pushed off of it by stochastic effects.

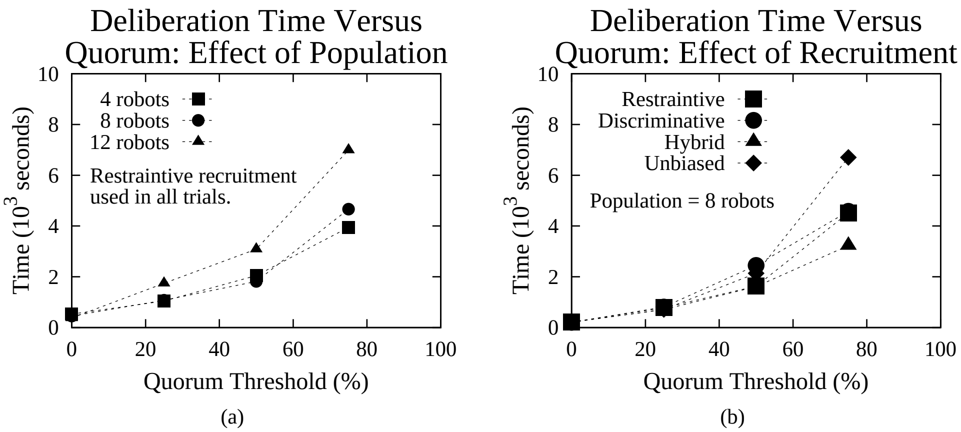


Figure 5.8: These figures plot the mean length of time that each system spent in the deliberation phase. Regardless of the number of robots that compose a dec-MRS or the kind of iterative recruitment employed, deliberation time increases with quorum. This happens because higher values of quorum required additional robots to be recruited in order to be satisfied. In each system, the number of robots that identify candidate sites is fixed, so increased deliberation is required in systems with larger population sizes.

Time Required for Deliberation

Finally, a decision-making algorithm can be judged by the time required by it to make a decision. Figure 5.8 plots the mean deliberation time of the decisions versus quorum. Deliberation time is defined here just as it was in Chapter 4: it is the time measured from the beginning of a decision-making trial until a robot commits to an alternative. This means that the time required by the searching phase is included in the deliberation time, but the commitment phase is not. Both of these phases are relatively constant in length, and so the inclusion of the former and the exclusion of the latter does not alter the trends displayed by the plots. Deliberation time increases with quorum, an observation that should be expected. As Figure 5.7 demonstrated, increasing the quorum increased the amount of recruitment in a decision. It is the additional time required by the extra recruitment that accounts for the increased deliberation time with quorum seen in Figure 5.8 (note the similarity between the shapes of Figures 5.8 and 5.7). Because there are more robots to be recruited in a larger dec-MRS, the deliberation time also should be expected to increase somewhat with population size, which it does. Once again, there is little difference between the three biased implementations of iterative recruitment.

5.2.5 Summary

The results of the simulated best-of-N decision-making experiments permit several conclusions to be drawn. First, the proposed framework of this thesis indeed enables a dec-MRS to make accurate best-of-N decisions. Increasing the quorum of a decision, which increases the duration of the deliberation phase, both increases the likelihood of a correct decision being made, and the likelihood of that decision being unanimous. Perhaps the most surprising is that no improvement in the decision-making performance of the robots was observed when direct comparisons were made of alternative quality (*i.e.* discriminative iterative recruitment), even though the individual robots in these experiments were in complete agreement about the sites' qualities. The emergent comparison of alternatives that is carried out by restraintive recruitment performed as well, if not better than the discriminative approach.

In these simulations, the individual robots tested quorum off-swarm by counting the absolute number of other robots that simultaneously visited a favoured site and compared this to the quorum. In some ways, this simplifies quorum testing, since the robots need not communicate with each other to do so. An advantage of off-swarm quorum testing, and in particular the specific manner in which it was implemented here is that premature commitment is largely eliminated. This is because a robot will not commit until it observes a quorum of robots at its favoured site *at a single instant*. If a robot can see five robots at the same time, it can be certain that there are at least five robots nearby. Therefore, the observed quorum was guaranteed to be at least as great as the quorum threshold in these simulations. This is not necessarily the case with on-swarm quorum testing, as the unary decision-making results illustrated.

However, when quorum is tested off-swarm, stagnation will become a problem as the population size of a dec-MRS or quorum increases. This was illustrated by the decrease in the decision-making ability of the 12-robot system when quorum was increased to 75% of system population[5]. Furthermore, the simplification offered by off-swarm quorum testing is of no real benefit to a robotic system. The robots already must be able to communicate with each other for the purpose of recruitment. Typically, a robot is able to communicate or it is not. If it is, then any number of different messages could be sent or received via its communication hardware with equal ease. This is very different from the communication of social insects, in which different chemicals are used to send different messages, requiring different glands to produce them and specialized receptors to detect each one. It therefore is to the insects' advantage to employ a passive quorum test since it does not require an expansion of their chemical vocabulary. The added complication of a quorum test that requires explicit communication is insignificant for a robot already able to exchange messages. Unless there is a very compelling reason not to do so, on-swarm quorum testing should be employed by a dec-MRS for collective decision-making.

[5]This result might seem somewhat odd given that the ants and bees, upon whose behaviour the proposed decision-making framework is based, employ off-swarm quorum testing, and that their populations range from a few tens to a several thousands of individuals. However, the insects base their quorum tests upon the rate at which teammates are encountered when visiting a particular site, rather than an absolute count. This population density based consensus estimation is susceptible to the same kinds of errors and thus false positive tests as analog consensus estimation, and the approach would have to be tuned to *very* specific kinds of decisions and the expected alternatives that would be available solutions. Correspondence with the primary researchers of the insects' behaviours suggests that the actual quorums thresholds employed are low, quite less than 50%, and that they vary considerably from one insect to the next [68].

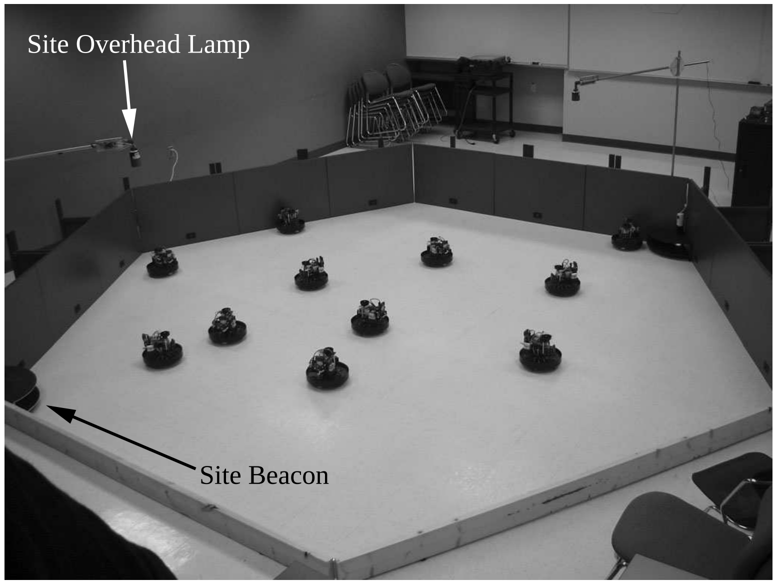


Figure 5.9: This photograph depicts the environment in which the physical site selection experiments were carried out. It was very similar to the environment of the unary decision-making experiments (a hexagonal enclosure, 2.75 meters per side), except that two candidate sites were added to it on opposite sides. These sites were the alternatives for the robots' best-of-N decision-making.

5.3 Physical Experiments

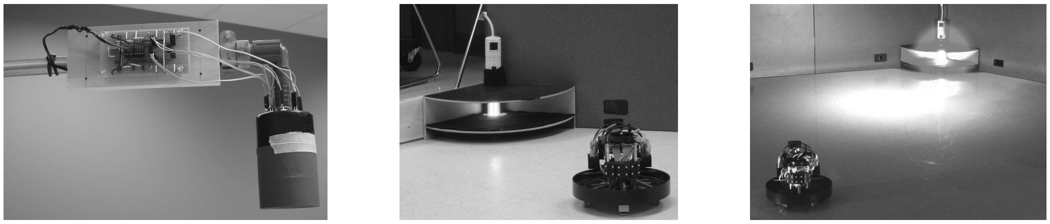
A series of experiments with real robots were carried out to further investigate best-of-N decision-making in the site selection domain. In particular, these experiments demonstrate that this approach to decision-making is practical in a real environment, and that the proposed framework is able to accommodate substantial noise in the individual robots' abilities to measure the qualities of the alternatives that they are able to find. The size of the scouting population also is varied in these experiments to examine its impact on collective best-of-N decision-making behaviour.

5.3.1 Environment

The physical best-of-N decision-making experiments were carried out in the same arena as the unary decisions. It consisted of a hexagonal enclosure measuring 2.75 meters on each side. Two sites were added to the environment to serve as alternatives over which the robots would deliberate. A photograph of the experimental environment is given in Figure 5.9. Although this photograph was taken in a well lit room, the experiments were conducted with all of the room's lights turned off except for those associated with the candidate sites, the details of which are provided next.

Sites

The goal of the robots in the site selection domain is to find and then collectively select the best available site in their environment. The sites were represented by calibrated overhead lights. As was the case in the simulated decisions, each site had associated with it a quality. In a real-world



(a) Overhead Light (b) Coloured Beacon (c) Site in the Environment

Figure 5.10: These three images show how the candidate sites were built for the decision-making experiments. At 5.10(a) is a close-up of a site's overhead light. The quality of a site is determined by its brightness. The attached circuit board controls the current to the lamp's 8-LEDs, and their brightness as a result. Because the robots were unable to localize themselves in their environment, coloured beacons were placed next to each site. One of these is shown at 5.10(b). During an experimental trial, the room was made completely dark, except for the sites' overhead lights and beacons. The photo at 5.10(c) shows what a site looked like during a trial. The illuminated spot on the ground in front of the beacon is the site itself.

application, this could be any attribute of a candidate location. For example, higher site quality might be associated with the availability of ground-water if the robots' mission was to identify potential well sites, or perhaps chemical concentration would be valued if the robots were assigned to locate a refinery leak. In the experiments, the quality of a site was represented by the brightness of a light suspended above it. A robot visiting a brightly lit site would tend to perceive it as higher quality than it would a dimly lit one. Although this is intended to serve as an abstract site quality here, this behaviour might be useful if the robots' mission was to identify the best place at which to deploy a solar array [40].

Refer to Figure 5.10 for the following description of the sites' construction. Each site's overhead light was provided by an array of eight white light emitting diodes (LEDs) connected to an adjustable current source (Figure 5.10(a)). The light emitted by an LED is very linear with the current that passes through it, permitting the brightness (and thus the perceived quality) of a site to be set precisely. Each LED array was housed in an ABS plastic pipe cap with a piece of white tissue paper fastened over its open end to serve as a diffuser. A cowl constructed from card was slid over the pipe cap, enabling the size of the spot projected on the ground to be controlled, which in turn determined the physical size of a site (*i.e.* the diameter of the illuminated spot on the ground). Two sites were placed on opposite sides of the experimental environment. The currents supplied to the two were 17.5 milliamperes and 4.0 milliamperes. These are referred to as the *better site* and the *poorer site*, respectively.

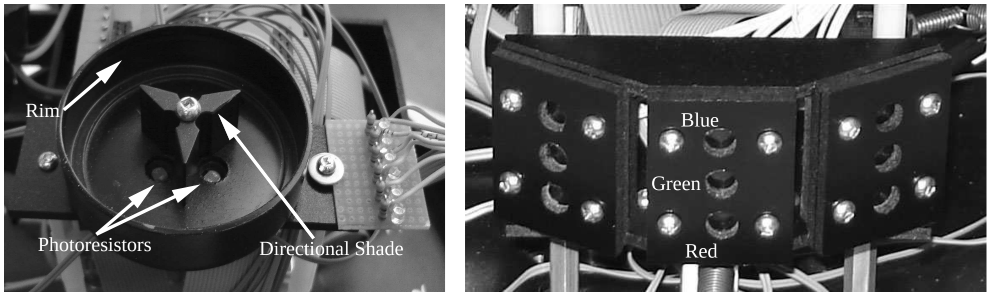
Due to the non-linear response of the cadmium-sulfide photoresistors that were used in the robots' overhead light sensors, the two sites were much closer in quality from the robots' points of view than these two currents suggest. The sites' currents were adjusted along with their shades so that the two sites were equally easy for the robots to find, and so that there was some overlap in their qualities as perceived by the robots. Figure 5.12 presents the quality of the two sites according to the robots. Each robot recognized that the better site indeed was brighter than the poorer one, but they were perceived to be sufficiently similar in quality that an individual robot's direct comparison of two was unlikely to be precise. The distributions presented in Figure 5.12 are based on the robots' assessments of site quality during the experimental trials presented later in this Chapter.

The robots possessed no means of localizing themselves within the environment (odometry would have been too inaccurate, whereas a more rigourous approach would have exceeded the individual robots' computational capabilities), so the follow-the-leader behaviour used in the simulations was impractical for the physical experiments. Instead, coloured beacons were placed on the ground adjacent to each site. These beacons made the sites uniquely identifiable and provided a common ontology for the robots to refer to them (*i.e.* a robot could recruit a teammate to its favoured site by communicating the colour of the site's beacon as part of its recruit-message). Each beacon was illuminated by a fluorescent work light covered with a coloured theatre gel, enclosed in a wedge-shaped enclosure lined with black felt (Figure 5.10(b)). The enclosures minimized the light from the beacons that reflected off of the ceiling, reducing the beacons' interference with the robots' searching behaviour and measurements of site quality. One of the beacons was blue and the other was red, and these were placed on the floor of the arena adjacent to better and poorer sites, respectively. When it is more convenient to do so, the sites are referred to by the colour of their beacon: the *blue site* and the *red site*. A completed site with a robot next to it in the darkened environment is shown in Figure 5.10(c)

5.3.2 Robots

The same robots were used for the physical best-of-N decision-making experiments as were used for the unary decisions. In addition to the 360 degree bumper sensor that was used to detect obstacles and teammates, the robots also possessed other sensors. On the top of each robot were three photoresistors pointed straight up, arranged in an equilateral triangle with one of its vertices pointing towards the robot's rear (refer to Figure 5.11(a)). A triangular piece of plastic[6] was placed at the centroid of the photoresistor triangle, rotated 180 degrees relative to the resistors. The plastic triangle selectively shaded the photoresistors like a sundial, so that when an overhead light was placed near a robot, the relative azimuth to the light could be determined. The magnitude of each photoresistor's response represented the length of a vector pointing from the center of the assembly towards the photoresistor itself. By summing the three vectors, a gradient was computed. The angle of the gradient provided the heading towards the overhead light, whereas its length allowed a robot to decide whether or not it was sufficiently close to the brightest point under the overhead light to make a reliable measurement of its brightness (the length of the gradient would be zero when directly under an overhead light). A short circular rim surrounded the entire assembly to prevent lights from being seen until they were at least 30 degrees above the sensor's horizon, reducing the interference

[6] A small piece of a three-sided engineering scale painted black.



(a) Overhead Light Sensor

(b) Beacon Sensors

Figure 5.11: In order to find, measure and identify sites, the robots were outfitted with upward-pointing site sensors and forward-pointing beacon sensors. The sensory elements in all of these were cadmium-sulfide photoresistors. 5.11(a) shows the overhead site sensor. Three photoresistors (the one in the rear cannot be seen in this image) were arranged in a plane with a triangular shade separating them. Their relative responses to an overhead light allowed a robot to compute direction to the point on the ground directly under a site's overhead light, where a measurement of its quality should be made. At 5.11(a) can be seen a robot's beacon sensors. Here, a column of three photoresistors, each covered by a different coloured gel (red, green and blue) allowed the robot to determine which coloured beacon it was facing. Each robot had three of these to increase the beacon sensor's field of view.

from the coloured beacons. This sensor, referred to simply as the *overhead light sensor*, allowed the robots to find the candidate sites in the environment and center themselves under the sites' lights. This centering operation allowed them to measure a site's quality at its brightest point.

Once a robot had found a site, its next task would be to determine the colour of its beacon. Beacon sensors also were assembled from photoresistors[7] (refer to Figure 5.11(b)). An array of three photoresistors was arranged into a column, with each one covered by a different coloured piece of theatre lighting gel (red, green or blue). Each element was connected to an 8-bit analog to digital converter, so the array was able to see a single pseudo-RGB[8] pixel. Each robot was equipped with three of these arrays; one pointing forward and one each pointing 30 degrees to either side. Because of the wide tolerances typical of photoresistors, all of the robots' beacon sensors were calibrated to the actual site beacons. This ensured that, when any two robots observed the same beacon, they would agree about its identity. The robots were able to detect a beacon once they were within approximately three meters of it.

Communication

Because the robots used in the physical best-of-N trials were the same as those used to demonstrate unary decision-making, communication amongst them was carried out using 802.11B wireless Ethernet. Again, the range of these radios was global in the experimental environment, so local peer-to-

[7]Photoresistors are known to be non-linear, have wide tolerances and slow responses. However, in the experimental environment, they were sufficiently fast for the purposes of these experiments and are incredibly cost effective, an important consideration when eleven robots must be equipped with them.

[8]The term "pseudo-RGB" is used because the passbands of the red, green and blue theatre gels used, combined with the responses of the photoresistors did not match that specified for true RGB.

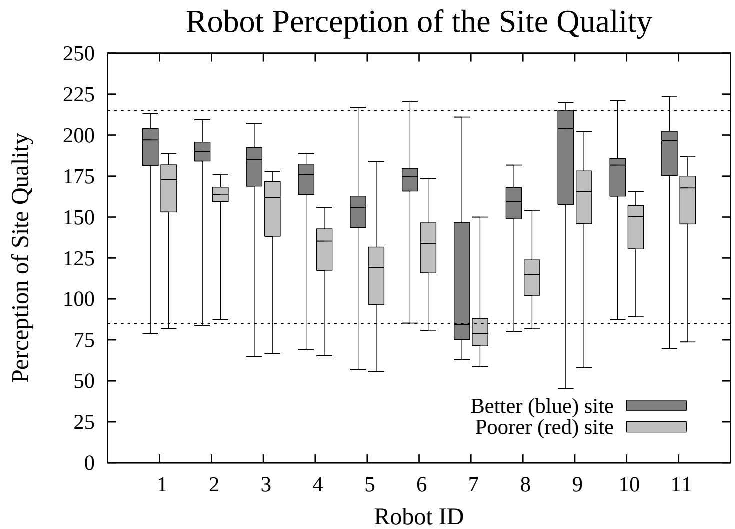


Figure 5.12: Unlike the simulated site selection experiments, the real robots' perception of site quality was noisy. This figure plots each of the eleven robots' opinions of site quality. The median, minimum, first and third quartiles, and maximum readings of each site's quality are plotted. All of the robots agreed that the blue site was better than the red site, although most had noisy enough perception of site quality that a single robot's opinion would be unreliable. The horizontal dotted lines indicate the perceived site qualities above or below which the robots' inter-recruitment delays would saturate (see Figure 5.13).

peer communication had to be simulated. This was accomplished in the same way as it was for the physical unary decision-making experiments. When a robot encountered an obstacle (being unable to tell the difference between a teammate and the wall) on its front half and communication was appropriate, it would randomly select a teammate's IP-address from a list supplied to it at run-time and send a peer-to-peer message to that teammate via TCP. The recipient would respond (if appropriate) with a peer-to-peer message back to the sender's IP-address. The resulting random peer-to-peer communication [59, 61] mimicked that of a well-stirred dec-MRS. IP-addresses were only used to address messages, and were not used for any other computation (*i.e.* received opinions were not associated with the identities of their senders).

5.3.3 Robot Behaviours

As illustrated by Figure 5.1, the best-of-N decision-making framework employs several distinct robot behaviours. In this section, the individual robot behaviours are described in detail, presented in the order that they appear in a best-of-N decision.

Searching/Idle

Robots began a decision in either the searching or idle states. The searching robots, also referred to as *scouts*, search the arena for candidate sites. The search behaviour is identical to the wander/reorient behaviour of the robots in the task completion experiments. The robots travel in straight

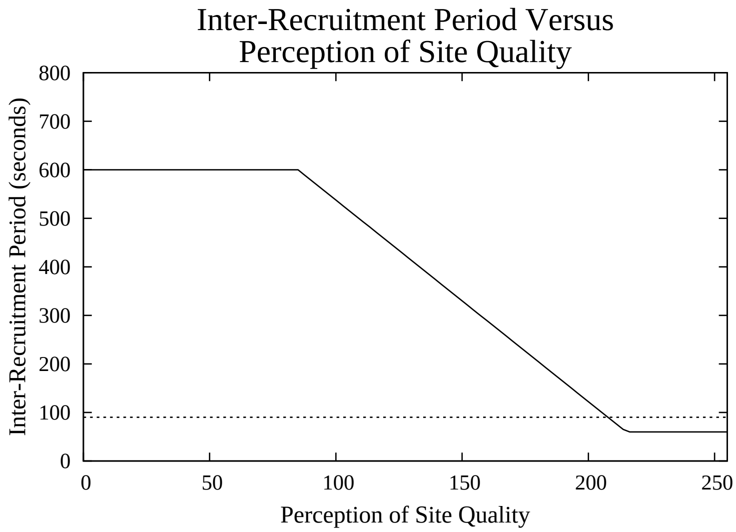


Figure 5.13: Restraintive recruitment was used by the robots in the physical experiments. In this approach, the advocate robots delay for a certain period of time between attempting to recruit teammates to favour their site; the better a robot believes its site to be, the less time it will delay, and thus the more frequently it will recruit. The solid line in this figure shows the relationship between a robot's perception of site quality and the amount of time that it delays between attempting to recruit. Additionally, to reveal any biases in the experimental environment itself, trials were run with unbiased recruitment, given by the dotted line. If one of the sites was easier to find, then this would be revealed by these unbiased trials.

lines, and reorient to random headings when they encounter an obstacle (a teammate or the environment's walls). As they wander, the scouts continually measure the intensity of the light above them. When this exceeds a preset threshold, a searching robot knows that it has found the edge of a site. It then moves towards the center of the site (its brightest point) by following the overhead light gradient. Once the robot has moved to within a preset distance of the center of the site (determined by the length of the gradient computed from the overhead light sensor intensities, not physical distance), the site's quality is recorded as the average of the three intensities reported by the overhead light sensors' elements. Finally, the robot rotates on the spot to identify the site's beacon. A site's quality and beacon having been found, a scout becomes an advocating robot, entering the deliberation phase of the decision favouring the site that it found. As was explained earlier, in the general case of a best-of-N decision, some scouts might not find an alternative, or perhaps some robots would not be equipped to participate as scouts. These robots are represented by the idle state. Idle robots sit motionless at the side of the arena until they are recruited into a decision.

Iterative Recruitment and Researching

Once a robot has entered the advocating state, the deliberation phase of a best-of-N decision begins. The advocating robots iteratively recruit their teammates to their favoured sites, modifying the apparent consensus in favour of those sites that have been identified. In the simulated decisions, four variations of iterative recruitment were examined. Restraintive recruitment performed very well in those experiments, and since it was the only biased model of recruitment that did not necessitate direct comparisons of site quality by the individual robots), it was the only iterative recruitment strategy implemented in the physical trials.

Robots that have found a site and entered the advocating state recruit their teammates to it at a rate that is based on the site's quality (*i.e.* the brightness of the overhead light). The better a robot believes its site to be, the more frequently it will recruit teammates to it. The rate of recruitment by a robot is controlled by the period of time between a robot's attempts to recruit teammates, called the inter-recruitment period, T_r. The relationship between the perceived quality of a site and the inter-recruitment period in the physical experiments is given by Figure 5.13. T_r decreases linearly from a maximum value as perceived site quality increases. Note that there are both maximum and minimum perceived site qualities (215 and 85), and that the inter-recruitment period is constant for site qualities above or below these values (60 and 600 seconds, respectively). Limiting the minimum value of T_r prevented recruitment from proceeding too rapidly. Recall that robots concluded that they had found the center of a site once the magnitude of the overhead light gradient was sufficiently close to zero. However, a robot might also observe a short gradient if it got lost and wandered into complete darkness. The areas of the environment away from the sites registered as approximately 50 on the robots' overhead sensors, so 85 was selected as a practical minimum site quality.

A recruit-message specifies only the colour of the beacon next to the site favoured by its sender. A recruitable robot (*e.g.* a robot in the searching, idle, or advocating states) that receives a recruit-message immediately enters the researching state in which it searches for a site near the specified beacon. Note that, because advocators do not need to lead their recruits to their favoured sites, recruiting is an atomic act for them. Therefore, an advocating robot cannot be interrupted mid-recruitment, and thus they are always recruitable. The search behaviour of a researching robot consists of a random walk similar to that described for the scouts, except that a researching robot will periodically rotate on the spot, scanning for the target beacon. If the target beacon is observed, the robot will reorient to face it and then resume straight-line motion. This scanning/reorienting behaviour biases a researching robot's search to the area nearby the target beacon, increasing its likelihood of finding the associated site. If a researching robot happens to find a different site, it will move away from it before resuming its search for the particular site specified by its recruiter. Although recruiting is an atomic action, researching a site once recruited is not. Therefore, researching robots could not be recruited. Instead, they would respond to recruit-messages as though they were queries for their opinions. Once a researching robot found the site that it was looking for, it would measure its quality and adopt it as its favoured site and forget about any other site that it might previously have favoured. The robot, now in the advocating state favouring the new site, would iteratively recruit robots that it encountered every T_r seconds based on its own opinion of its favoured site's quality. If a researching robot was unable to find the specified site in a predetermined period of time (set to 180 seconds in all of the physical trials), it would give up and revert to whatever state it had been in prior to being recruited. This time limit rarely was reached during the experimental trials. If an advocating robot received a recruit-message instructing it to research a site that it already favoured, the recruit-message would be responded to as though it was a query for its opinion and the robot would not enter the researching state.

Quorum Testing

As the advocating robots iteratively recruit each other, they also estimate the popularity of the sites that they favour and compare their estimates to the quorum threshold. The apparent consensus in favour of a particular site was estimated using the digital approach described in Section 3.4.2, by computing the proportion of the n most recently received teammate opinions that were agreeing. Teammate opinions were gathered via query-messages. A query-message in a best-of-N decision is equivalent the question "Do you favour alternative X?", where X is the particular alternative favoured by the querying robot. The response to this query would be "yes" if the queried robot favoured alternative X, or "no" if it did not[9]. Each advocating robot only tests quorum for the particular alternative that it favours, which means that the number of known candidate alternatives in no way affects the complexity of quorum testing. It also decreases the likelihood of false positive quorum tests for unpopular sites, since fewer robots test quorum for them.

When a robot was recruited successfully, it would forget about any previously favoured alternative. An essential part of this process was forgetting any previously received vote-messages, since these would be relevant only to the previously favoured alternative. However, a robot would begin to favour a new alternative and forget previous teammate opinions only after it had found the site to which it was recruited and entered the advocating state favouring it. Although robots in the researching state would not query the teammates that they encountered for their opinions, they would respond to query-messages, and did so as if they had not been recruited. For example, suppose that a robot favoured the red site, and then received a recruit-message for the blue site. The robot would enter the researching state and begin to search for the blue site, but until it found the blue site and entered the advocating state favouring it, the robot would continue to respond to queries as though it still favoured the red site. This choice of behaviour for researching robots is somewhat arbitrary. Alternatively, researching robots could have been programmed to ignore query-messages altogether. Given that researching robots tended to find their target sites relatively quickly, the overall impact of this decision on the collective decisions likely was insignificant. Robots responded to recruit-messages as though they were queries, too. This behaviour ensured that robots that recruited more frequently would not be deprived of their teammates' opinions for their quorum tests. Once an advocating robot believed that the apparent consensus for its favoured site had reached the quorum threshold, it would enter the committed state.

[9] Robots also would answer "no" if they favoured no alternative at all.

Commitment

This is the final behavioural state through which the robots would pass before exiting a decision, and it essentially is the same in a best-of-N decision as commitment in a unary decision. When an advocator commits because it computed a positive quorum test, it continues to favour the alternative that it did as an advocator. Instead of querying encountered teammates for their opinions, a committed robot instructs them to commit, and it includes in its commit-messages the identity of its favoured site as the particular site to which the recipient of the message should commit. Unless the recipient has entered the finished state and exited the decision, it always obeys, committing to whatever site was specified. These robots in turn tell the robots that they encounter to commit, too. If the recipient of a commit-message was not already committed to the specified site, it would respond with an acknowledgment, otherwise no response would be given. In this way, if more than one site satisfied quorum and triggered commitment, a period of attrition would follow after which one of the commitment-inducing sites would remain and thus be selected unanimously. The site with the most committed robots at the beginning of this attrition would be the most likely to be chosen in the end since it would have had a head start building up a corps of robots committed to it.

Once a committed robot has been in the committed state for longer than the commitment timeout without receiving any acknowledgments to its commit-messages, it will conclude that all of its teammates have committed to the same site as it and will enter the finished state, exiting the collective decision. If the commitment timeout is made sufficiently long, all of the robots will be very likely to exit the decision at approximately the same time, each believing that every other robot had exited the decision similarly committed. This behaviour was illustrated by the results of the simulated unary decision-making experiments.

5.3.4 Experimental Trials

A series of experimental trials was conducted with an 11-robot system to examine decision-making performance in the site selection domain. Apparent consensus was estimated using $n = 15$ samples, and the quorum threshold (Q) was set to 33%, 53%, and 80%. These configurations were repeated with four and eleven (all) of the robots participating in the initial search for candidate sites. When only four robots scouted, the remaining seven would begin in the idle state. The 11-scout, $Q = 80\%$ configuration was repeated with the inter-recruitment period set to a constant 90 seconds (indicated by the dashed line in Figure 5.13). This configuration implemented unbiased decision making, the performance of which served as a control trial that revealed any bias in the experimental environment. An 11-scout, $n = 5$, $Q = 20\%$ system also was implemented. With so many robots testing such a low quorum based upon such inaccurate estimates of apparent consensus, the purpose of these trials was to induce commitment to both sites in order to observe the resulting attrition during the commitment phase. The robots were lined up along the two walls of the arena that were the furthest from the two sites at the beginning of each trial. The commitment timeout was 60 seconds in every

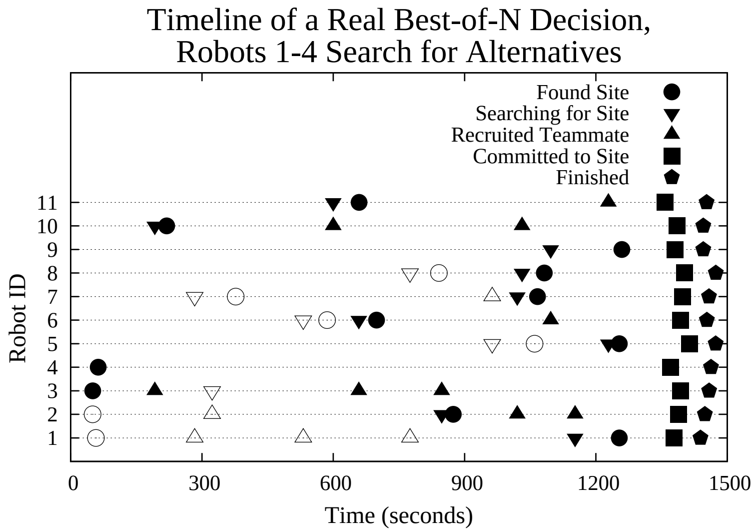


Figure 5.14: This timeline depicts a best-of-N decision using eleven real robots, the first four of which acted as scouts. Solid and hollow symbols refer to the better and poorer sites, respectively. Two of the scouts find the better site and two find the poorer site. Even though robot-4 makes a poor evaluation of the better site's quality, this error eventually is overcome by other robots recruited to the better site, and the final decision is unanimous in favour of it. This illustrates the self-correcting nature of the proposed decision-making framework. When a robot's timeline indicates that it recruits a teammate, the recruited teammate can be identified by an upside down triangle of the same colour in its timeline at the same time. For example, robot-3 recruits robot-10 to the better site at 200 seconds, and robot-10 finds the site soon after.

trial, and every one of the nine experimental configurations was repeated between 20 and 26 times.

5.3.5 Results

The results of the physical experiments are presented in this section in the same order as they were for the simulated trials. Figures 5.14 and 5.15 present timelines of two trials, illustrating typical decision-making behaviour. In both of these, the quorum threshold was 80% and apparent consensus was estimated using $n = 15$ samples. In Figure 5.14, only four of the robots scouted for sites, whereas all eleven did so in the trial shown in Figure 5.15.

Once a robot found a site, it began to recruit its teammates with a frequency determined by its own opinion of its favoured site's quality. As soon as one of the robots believed that quorum had been met, it committed to its favoured site, and its teammates rapidly followed suit as commit-messages flooded the system. After all of the robots had committed to the same site, the acknowledgments to commit-messages ceased, and approximately 60 seconds later (the length of the commitment timeout), the robots began to enter the finished state having unanimously selected one of the sites. When all of the robots scouted, the early stages of the deliberation phase appear somewhat chaotic, but over time, the consensus in favour of the better site increases, and ultimately it is chosen. Note also that more robots find the poorer site in Figure 5.15, yet the better site still is selected by the system as a whole by the trial's end. The individual robot and system-level behaviours illustrated by

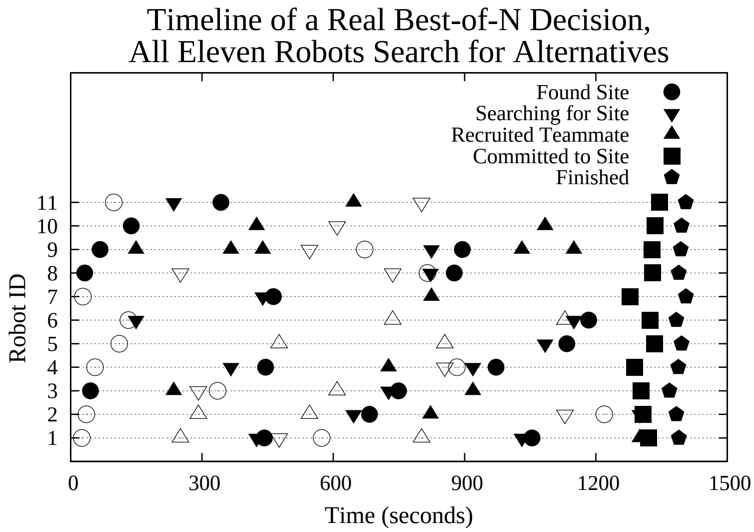


Figure 5.15: This timeline illustrates a best-of-N decision in which all eleven of the robots acted as scouts. Overall, recruitment is more frequent by the better-site favouring robots, so the proportion of the robots that favour that site tends to increase. Eventually, robot-7 determines that quorum for the better site has been satisfied (note that it initially favours the poorer site, and that it's opinion of the better site actually is quite low) and it commits, inducing the rest of the dec-MRS to follow suit. Once all of the robots have committed to the same site, responses to commit-messages cease, and the robots all exit the decision unanimously favouring of the better site.

these timelines is typical of that observed in the physical trials.

Ability to Make Correct Decisions

Every one of the physical decision-making experiments ended unanimously, including those that experienced attrition during the commitment phase. However, not every experiment ended with a correct decision. Whether or not the robots selected the better (blue) site by the end of a trial is not a good indicator of decision-correctness, since it was possible in a given trial that none of the scouts would find that site. Furthermore, because the robots' perception of site quality was noisy, their opinions of the two sites should be taken into account when labeling a particular decision as correct or not. For example, if most of the robots that found the blue site undervalued it and *vice versa*, then from the robots' points of view, it could be argued that choosing the red site would be the correct decision to make. However, the later in deliberation phase that a robot measured a site's quality, the less of an impact that robot's opinion would have on the decision. Therefore, the correctness of a decision is not black and white. Instead, the following heuristic was used to make this classification: *in a given trial, a decision was correct if and only if it unanimously selected the alternative that received the single highest evaluation by any robot.* Note, however, that using the average evaluation of a site, or the best evaluation by a scout (since scouts make their evaluations of site quality early in a trial) instead of the aforementioned maximal rule does not significantly alter the trends observed, so the following discussion of decision-making ability does not hinge upon this

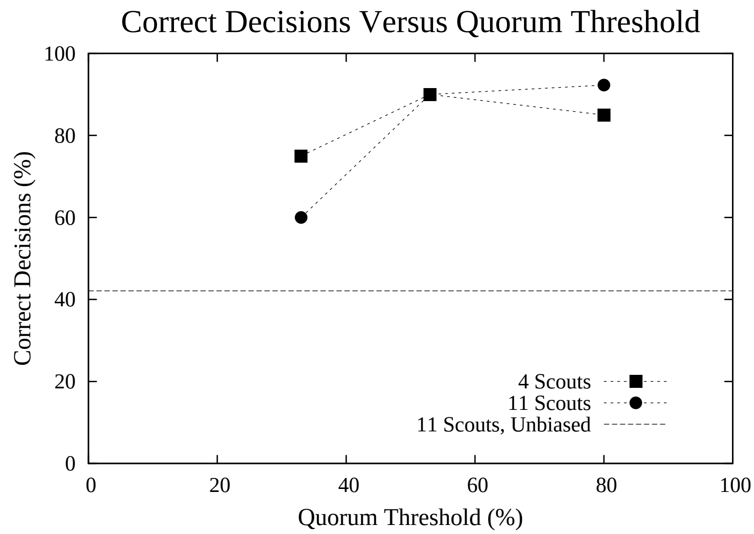


Figure 5.16: This figure plots the proportion of the best-of-N decisions that chose the best site found. In general, the ability of a dec-MRS to make correct decisions increases with quorum. When quorum is low, the results of the initial search for sites can adversely affect performance when too many scouts are involved. Raising quorum overcomes this problem, but reveals another mode of failure. Too few scouts allows stochastic effects to influence the outcome of a decision, reducing performance. The horizontal dotted line indicates how accurately an unbiased dec-MRS was able to make the decisions. Even when quorum was as low as 33%, biased iterative recruitment made good decisions much more likely than random chance.

definition.

Figure 5.16 plots the percentage of the trials in which a correct decision was made. Both scouting populations made the correct decision at least 60% of the time, and this increases to at least 80% when the quorum threshold is increased to 50% or more. The size of the scouting population deployed by a system appears to affect its decision-making ability, particularly when quorum is less than 50%. When quorum is greater than 50%, the system with more scouts performs better, although the difference between the performances of the two scouting populations is somewhat less here than it was when $Q = 33\%$.

When all of the robots participated in the initial search for sites, there was a chance that quorum would be satisfied without any recruitment having to take place. For instance, if all eleven of the scouts found a site, a quorum threshold of 33% would have been satisfied for one of the sites since at least six of the robots would have favoured the same site, leading to an apparent consensus of 50% ($\frac{6-1}{11-1} \times 100\%$). Because the scouts were equally likely to find either site, the particular site that would have satisfied quorum in this case would be determined by random chance. In practice, however, the scouts did not all find sites at the same time, so some of them would have had an opportunity to recruit before one of them believed that quorum had been satisfied. Therefore, iterative recruitment still influenced the outcome of the decisions, although its ability to do so was much reduced. Eight of the twenty eleven-scout, $Q = 33\%$ trials made incorrect decisions, and in three of these no recruitment occurred at all. The remaining five incorrect decisions involved only one, two or three recruitments. Thus iterative recruitment, the process that compares the known alternatives, was largely absent in this configuration. When the scouting population was reduced to four, however, recruitment *always* was necessary to satisfy quorum. Even if all four scouts happened to find the same site (6.25% probability), the apparent consensus would only have been 30% ($\frac{4-1}{11-1} \times 100\%$). Because iterative recruitment had a greater opportunity to promote the better site in the four-scout configuration when the quorum threshold was 33%, they performed better than their eleven-scout counterpart [63].

The situation is reversed when quorum is increased, as the eleven-scout system performed as well or better than the four-scout system when the quorum threshold $\geq 50\%$. A greater quorum was much less likely to be satisfied by the outcome of the searching phase alone, and so iterative recruitment always played a role regardless of the size of the scouting population. When the four-scout trials in which incorrect decisions were made are examined in greater detail, a second mode of failure becomes evident. Recall that a dec-MRS knows only what its individual members know. If all of them were to forget about a particular site, that site would be forgotten by the dec-MRS as a whole and thus would not be selected. The following example illustrates how deploying a small number of scouts could negatively impact a dec-MRS' decision-making ability. Consider a system in which only four robots are scouts. The probability of three of them finding one site and only one finding the other is 50%. In half of these cases, three robots would find the red site and one would find the blue site, and it is likely that the quality assigned to the blue site by the latter robot would be greater than any of the qualities assigned to the red site. If the lone blue-site-favouring robot was recruited successfully to favour the red site, the blue site would be forgotten by the dec-MRS altogether, resulting in an incorrect decision according to the aforementioned definition. This is because the site that was assigned the highest quality over the course of the decision (the blue one) was forgotten when its lone advocator was recruited to the other site. This mode of failure accounts for 60% of all of the four-scout trials in which an incorrect decision was made. Increasing the scouting population decreases the likelihood of a site being found by just one or two robots, making this kind of failure much less likely.

Furthermore, because the early stages of the four-scout system's deliberation phases were based on fewer evaluations of site quality (compare Figures 5.14 and 5.15), the noise in the individual robots' overhead site sensors would have resulted in noisier decision-making behaviour than that of the eleven-scout system. The slight dip observed in the four-scout system's ability to make correct decisions likely can be attributed to this noise. With sufficient trials upon which to base the plotted means, it is likely that the four-scout system would display a leveling off in its ability to make correct decisions as quorum is increased from 50% to 80%, if not a slight increase.

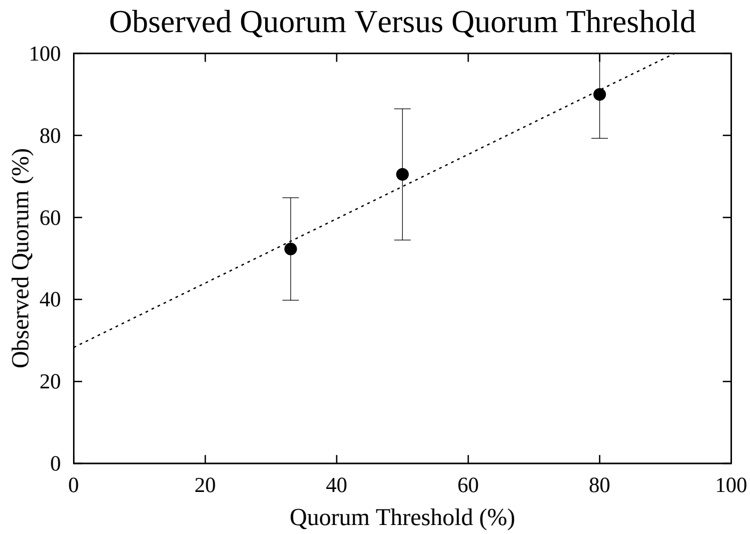


Figure 5.17: This figure illustrates the relationship between the observed quorum and the quorum threshold (Q) from the best-of-N decision-making trials. Notice that the observed quorum here is greater than it was in the unary decisions (Figure 4.14). As the individual robots gather each other's vote-messages, the also iteratively recruit each other, changing the apparent consensus for each alternative. Because a robot's estimate of apparent consensus is based on the average value of C_a during the period over which the n most recent vote-messages were received, and because the apparent consensus for the best alternative (the one most likely to induce commitment first) tends to increase over time, $\tilde{C}_a$ will tend to underestimate C_a, and thus the observed quorum will tend to be greater than expected.

Observed Quorum in a Best-of-N Decision

An alternative's popularity reaching quorum signifies to a robot that favours it that the deliberation phase's iterative recruitment has identified its alternative as the one that should be selected by the dec-MRS as a whole. As was illustrated in Chapter 3, if iterative recruitment is allowed to continue for a sufficiently long period of time, all but one of the known alternatives will be forgotten. In practice, however, a system might not be able to afford that much time for deliberation. Therefore, a quorum threshold less than 100% is used to terminate the collective comparison of alternatives before complete consensus is achieved, when one of them has been identified as *good enough*. In Chapter 4, the notion of the observed quorum was introduced. The observed quorum is the consensus (true consensus, not apparent consensus) in favour of the alternative that first induces commitment at the time that the first commitment occurs. Figure 5.18 plots the observed quorum for the four-scout best-of-N decision-making trials versus the quorum threshold[10].

As was the case in the unary decision-making results, the observed quorum increases with the quorum threshold, demonstrating that a desired consensus necessary for commitment can be set *a priori*. However, if this figure is compared to Figure 4.8 or Figure 4.14, one will notice that the observed quorum is significantly greater for the best-of-N decisions than it was for the unary

[10] The observed quorum of the eleven-scout trials is much noisier, particularly when the quorum threshold is low, so it was omitted from this plot.

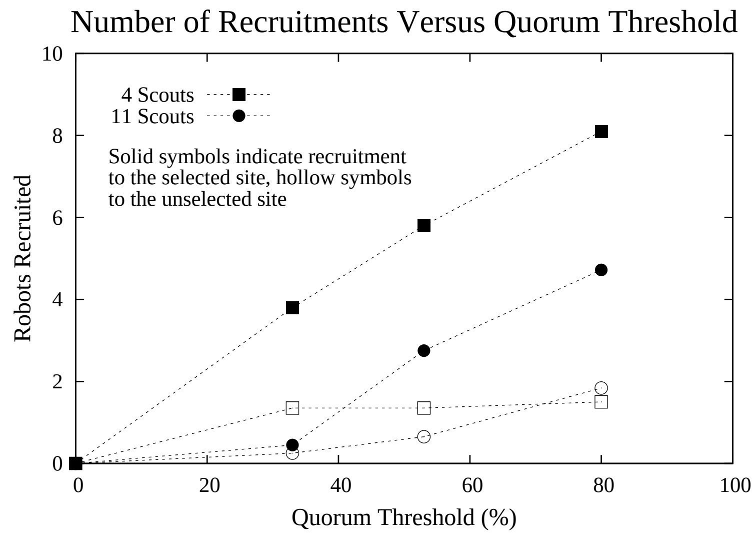


Figure 5.18: A good decision-making algorithm will minimize the amount of time and energy spent considering alternatives that are unlikely to be selected in the end, because this represents a waste of time and energy. This figure illustrates that, regardless of the size of the searching population, considerably more attention (in the form of the number of robots recruited) is paid to the site ultimately selected by a decision than the unselected site, and that this increases with quorum.

decisions. It also is greater than the theoretical predictions of Appendix A suggest it should be, too. To understand why this is so, one must look at the two decision-styles' deliberation phases, as this is the principle manner in which they differ. When a unary decision-maker enters the deliberating state, it begins to collect its teammates' opinions to compute $\tilde{C_a}$, its estimate of apparent consensus. As more of its teammates enter the deliberating state, the apparent consensus will increase, and so will $\tilde{C_a}$. Because the robots' estimates of apparent consensus contain some error, it is likely that a robot will believe that the apparent consensus has reached the quorum threshold before this actually occurs. On the other hand, in a best-of-N decision, while the individual robots estimate the apparent consensus, they also are recruiting each other, actively modifying it. Robots that favour inferior alternatives are likely to be recruited and thus have to start collecting teammate opinions all over again for some new alternative, so it will be less likely for a robot that favours a poor alternative to commit. Furthermore, while a robot that favours a good alternative is gathering vote-messages, the popularity of its alternative will tend to increase, increasing the value of observed quorum should it or an agreeing advocator commit to that site. Therefore, the observed quorum for a given value of the quorum threshold will tend to be greater than expected, but the degree to which it will exceed the prediction likely depends on the domain and the particular decision being made.

The Focus of Deliberation

The focus of the deliberation phase tended to be the site that ultimately was selected by the robots. Figure 5.18 plots the mean number of robots successfully recruited by advocating robots to the

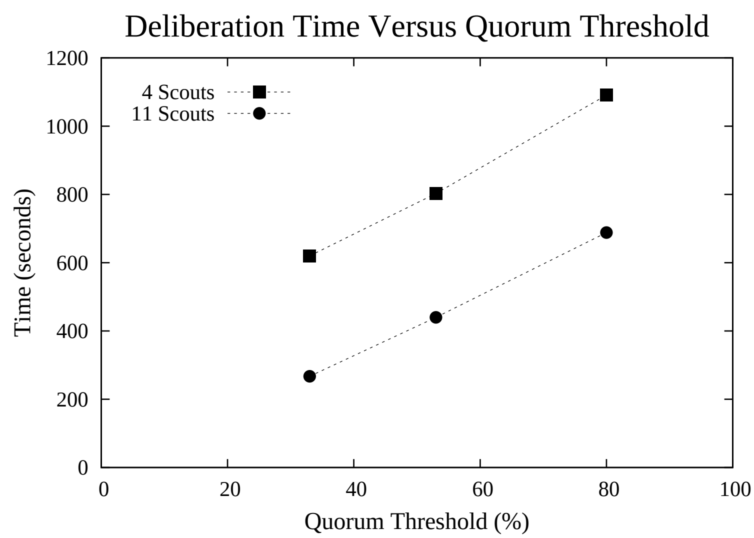


Figure 5.19: Increasing quorum demands greater accuracy from a decision, so the robots spend more time in the deliberation phase of the framework, where the best-of-N is determined. The more robots participate in the initial search for sites, the greater the apparent consensus in favour of the known sites will tend to be at the beginning of deliberation, so less recruitment (and therefore a shorter deliberation phase) would be required to satisfy a given quorum.

site that ultimately was selected, and to the site that was not selected. Regardless of the size of the scouting population, significant recruitment occurred in favour of the selected site, while the unselected site largely was ignored by the advocators. This shows that the proposed decision-making algorithm is able to discard and thus not waste time or energy considering alternatives that ultimately will not be chosen by the robots. Fewer robots needed to be recruited when eleven of the robots acted as scouts, since those that were able to find a site on their own (half of which tended to find the site that was selected in the end) did not need to be recruited into the deliberation phase. Of course, as the data plotted in Figure 5.16 demonstrates, the ability to make correct decisions can suffer because of this unbiased head start.

Time Required for Deliberation

The amount of time required to complete a decision also is affected by the quorum threshold and the size of the scouting population. Figure 5.19 plots the mean deliberation time of the robots versus the quorum threshold. As the simulations predicted earlier in this chapter, increasing the quorum threshold increases the amount of time required by the robots to compare the known alternatives. Decreasing improvements in decision-making accuracy (Figure 5.16) require an increasing amount of time to realize, meaning that diminishing returns are encountered. This result illustrates why a 100% quorum might not be desirable in practice. Increasing the size of the scouting population decreases the deliberation time by a constant amount, independent of quorum. Because fewer recruitments are required by the eleven-scout system in order to satisfy quorum (Figure 5.18), its

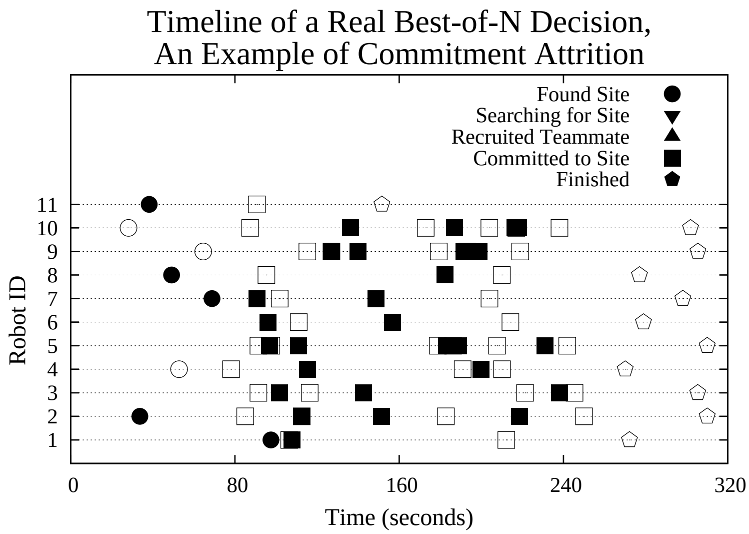


Figure 5.20: This final timeline demonstrates what happens if two sites satisfy quorum. The robots committed to each instruct every robot that they meet to commit to their favoured site, but they switch sites when they receive a commit-message referring to the other one. Normally, the site that induced commitment first would be selected in the end. In this trial, the two were committed to so rapidly that neither gains the upper hand. Robot-11 prematurely decides that unanimity has been achieved and exits the decision early. It is only due to luck that the rest of its teammates end up committed to the same site as it. Attrition in the commitment phase is best avoided altogether by making quorum higher and ensuring that it is measured with sufficient samples to make the measurements accurate.

deliberation phase will tend to be shorter. Note the similarity between the shape of the two curves in this Figure 5.19 and those plotting the number of robots recruited to the site ultimately selected in Figure 5.18. It is the time required by the additional recruitment to the selected site that largely determines the duration of the deliberation phase, since these by far are the most numerous.

Attrition During the Commitment Phase

All of the experimental trials ended unanimously. Furthermore, even when quorum was as low as 33% only one of the two sites ever induced commitment as long as apparent consensus was measured with sufficient precision. It is possible, however unlikely, for two different alternatives to induce commitment in a decision. This might occur when two very similar alternatives are found and quorum is less than 50%, or when the apparent consensus is estimated inaccurately. When this occurs, it falls to the commitment phase to ensure that unanimity is achieved. Because a committed robot will not exit the committed state until it believes that all of its teammates have committed to the same alternative as it, a period of attrition will ensue, during which the robots committed to the different alternatives send each other commit-messages. Recall that when a robot receives a commit-message referring to an alternative to which it is not already committed, it commits to that alternative and responds with an acknowledgment. The perceived quality of a site plays no role in the commitment phase, since committed robots send every robot that they encounter a commit-message. In a war of attrition, it is the side with the greatest number of combatants that will emerge victorious, and this also is the case when attrition occurs in the commitment phase. The alternative that induces commitment first will have a head start, and during this time its committed corps will swell in number. Therefore, the first site to induce commitment will tend to be the one that is selected at the end of a decision. Of the 20 $n = 5, Q = 20\%$ trials, commitment attrition was observed in just eight of them. In every case, the site that induced commitment first was the one that was selected by the robots in the end.

A timeline of the trial that experienced the longest period of attrition is depicted in Figure 5.20. The squares appear on each robot's timeline whenever it committed to an alternative, and most of the robots can be seen to switch back and forth several times between the two sites. None of the robots are supposed to enter the finished state until all of their teammates have committed to the same site as them, but this particular timeline reveals a danger in commitment attrition. Robots will conclude that all of their teammates have committed to the same site when they have been committed to the same alternative for longer than T_c seconds without receiving any acknowledgments to their commit-messages. However, it is possible for a particular committed robot to encounter only those teammates that are similarly committed and thus decide that it can exit the decision safely. Robot-11 can be seen to do this so at approximately 150 seconds into the trial, at which point it enters the finished state and exits the decision. It is only through luck that the site ultimately selected by the other robots later on is the same as the one chosen by robot-11. This untidy behaviour occurs because both sites induced commitment at almost the same time, giving neither one the opportunity to gain the upper hand before the other induced commitment. In the other trials that experienced attrition, the second site to induce commitment did so somewhat after the first, so it quickly was forgotten. Increasing the length of the commitment timeout would make this sort of premature exit from a best-of-N decision less likely, but commitment attrition should be avoided as a general rule whenever possible by choosing n and Q such that commitment to multiple alternatives will be unlikely. A reliable quorum test for a quorum $\geq 50\%$, such as $n = 25, Q = 80\%$ would satisfy this criterion well for dec-MRS with populations up to at least 100 robots.

5.4 Summary

This chapter extended the unary decision-making strategy presented in the Chapter 4 to accommodate best-of-N decisions. Unlike a unary decision, in which a single proposed alternative is accepted or rejected, a best-of-N decision identifies and selects the single best alternative from a decentralized list. This list is created by the robots' initial search. The candidate alternatives are compared via iterative recruitment, and this process is terminated once one of the alternatives is believed to have obtained a quorum of support from the robots. The belief of a single robot that quorum has been satisfied by the particular alternative that it favours is amplified by the final phase of a decision, commitment, which induces the entire dec-MRS to adopt that alternative as the result of the collective decision.

Experiments using both simulated and physical robots to make best-of-N decisions were carried out in a site selection domain. These demonstrated that the emergent, three-phase decision-making framework can make good best-of-N decisions, even when the individual robots' sensing is noisy. In the simulated experiments, off-swarm quorum testing was investigated. This operation more closely resembles the manner in which honeybees and *Temnothorax* ants estimate consensus and test quorum, which is based on the sizes of the populations visiting the favoured sites. Only certain domains, such as site selection can take advantage of off-swarm quorum testing. The only real benefit that it offers is that it eliminates the need for inter-robot communication to test quorum. However, given that iterative recruitment still necessitates inter-robot communication, the off-swarm strategy offers no real savings to a dec-MRS in practice. Furthermore, it is limited to small quorums, the robots must know the absolute population size of their system, and it increases the likelihood of decision stagnation. For all of these reasons, on-swarm consensus estimation and quorum testing is much better suited to collective decision-making by a dec-MRS.

Somewhat surprisingly, best-of-N decisions exhibited a greater observed quorum than similarly calibrated unary decisions. That is, for a given quorum test (*i.e.* n and Q), the consensus observed at the time of commitment tended to be greater than expected in the best-of-N operation. This observation makes sense, however, when the impact of iterative recruitment and its interaction with consensus estimation is considered more closely. Iterative recruitment, which compares the candidate alternatives during the deliberation phase of a decision increases the apparent consensus in favour of the best known alternative by inducing robots that favour inferior alternatives to abandon them for the best alternative. Therefore, robots that favour poorer alternatives are less likely to commit to them, both because the apparent consensus in favour of them will tend to be both low and decreasing, and because robots will tend to be recruited to favour better alternatives before they have favoured an inferior alternative long enough to make an error. At the same time, the robots' estimates of consensus lag reality. When a robot commits to the best alternative (the most likely outcome), the actual apparent consensus will tend to be somewhat greater than the robot's estimate, since it likely will have increased over the period of time that the robot collected the necessary vote-messages to make its estimate.

Both the simulated and real robots improved their decision-making accuracy when either the quorum threshold or the number of teammate opinions used to compute $\tilde{C}_a$ was increased. Increasing either of these parameters increased the amount of iterative recruitment that would take place in a decision, increasing the reliability of the process that identifies the best alternative. Of course, accuracy and speed are competing interests, as was illustrated in Chapter 3. Therefore, it was no surprise that the increased decision-making accuracy achieved by increasing quorum came at the cost of additional deliberating time.

The variations on iterative recruitment investigated by the simulated best-of-N decisions demonstrated that direct comparisons of candidate alternatives by the individual robots are unnecessary, as the emergent positive feedback of restraintive recruitment, in which no direct comparisons ever are made, performed just as well as (if not better than) discriminative recruitment. Direct comparisons can be dangerous in a system composed of many simple individuals, since it introduces the potential for individual stubbornness, which could lead to stagnation during collective deliberation. Furthermore, real robots will have noisy sensors (illustrated by Figure 5.12), so the individual robots should not be assumed able to directly compare two candidate alternatives. There is one exception to this statement, however. Although it was not the case in the experiments described in this chapter, scouts could directly compare candidate alternatives *before* they are recruited for the first time. This is because it is the duty of the scouts to populate the menu of candidate alternatives. In this way, a scout could increase the likelihood of informing its teammates of a good alternative when it enters the deliberating phase. The key point to make here is that the scout would be comparing two alternatives that *it* had found.

The size of the scouting population was observed to have a significant impact on the accuracy and precision of collective decision-making. If too many robots scout for a decision that has only a few alternatives, then the outcome of their searching behaviour alone might satisfy quorum without any collective comparison every taking place. However, if too few scouts are deployed, then each alternative found will be held in the collective memory of a dec-MRS only tenuously. In general, a greater population of robots will be able to perform a better comparison of a given set of alternatives, since the susceptibility of such a population to stochastic errors will be less than that of a smaller dec-MRS. However, the quorum for a decision should be made sufficiently large that collective deliberation will be given a reasonable opportunity to identify the best alternative. For domains in which multiple alternatives are likely to be found, a reliable test for a quorum of 50% would suffice, but when only two alternatives are expected (*e.g.* deciding between two paths at a fork in a road), a greater quorum is advised. Iterative recruitment eventually will bring about the unanimous support of one of the candidate alternatives, so fear of stagnation should not prevent the use of a large quorum when domain-specific knowledge is lacking.

Chapter 6

Discussion and Conclusions

In this, the final chapter of this thesis, the contributions of this work are summarized, and directions for future work are provided.

6.1 Contributions

This work has focused on the ability of dec-MRS to make cohesive, system-level decisions through a sequence of emergent behaviours. It is the use of emergent behaviours to build up the three phases of the decision-making framework that is particularly important, since it means that the entire process depends only on the local interactions of the individual robots. This is the mode of operation of a decentralized system like a dec-MRS, so the approach to decision-making presented in this work is *ideally* suited to such systems.

Much of the existing research with dec-MRS also takes advantage of emergence and peer-to-peer interactions, but it treats the robots as a kind of end effector, things that are to be commanded, rather than an intelligent collective that senses its environment and computes its own cohesive response. That is, most work has viewed the micro-macro link [76] as a one-way street, in which local behaviours combine to produce a globally observed emergent behaviour with no direct link back from the macro to the micro. In some cases, individual robots have been programmed to respond indirectly to their macroscopic behaviour by tuning their individual behaviours to the environment that they collectively modify (stigmergy), but this is a brittle arrangement [6].

Perceptual cues were introduced in [48], and the idea was that individual robots, instead of physically responding to certain stimuli, might instead react cognitively. For example, a robot might respond to the Sun rising by switching to its daytime behaviours, switching back to its nighttime mode of operation after the Sun had set. In contrast, the quorum test employed by the deliberating robots to determine when sufficient deliberation had taken place is a *social cue*. These robots did not predicate their choice of behaviour on some varying aspect of their environment, but instead directly upon the global state of their team. Macro was sensed by the micro, and the latter's response brought about a phase change in the former. This completes the feedback loop (see Figure 6.1), enabling a

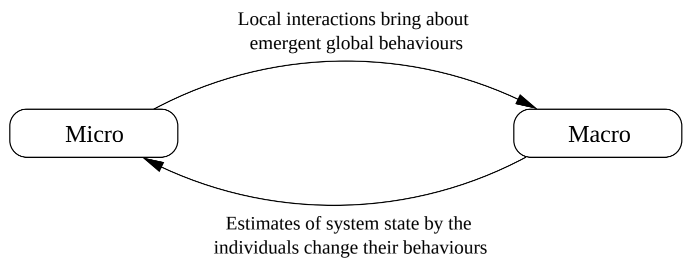


Figure 6.1: In many dec-MRS, the micro-macro link is unidirectional. The individual robots interact with each other and the environment, and a global (macroscopic) behaviour emerges. The system-level decision-making framework of this thesis uses consensus estimation and quorum testing to complete the loop, enabling the robots to predicate their behaviours directly upon their collective state.

dec-MRS to respond directly to itself, and thus predicate its collective behaviour on its state.

Completing this loop is significant. It means that dec-MRS can be viewed as intelligent entities in and of themselves, rather than loose collections of coordinated individuals. In this way, a dec-MRS mission could be designed as though it was any other robotic mission, and the dec-MRS could be treated as though it was any other robot. The illusion of centralized control provided by closing the micro-macro loop is a useful layer of abstraction for designers. When it makes sense to do so, a cohesive dec-MRS could be treated as an intelligent and autonomous *individual*, while still retaining all of the robustness and scalability of a decentrally structured entity.

6.2 Issues for Future Study

Although this work has demonstrated that a dec-MRS can indeed make intelligent, system-level decisions, there still are issues that remain to be investigated further.

6.2.1 Asynchronously Initiated Decisions

Two kinds of collective decisions have been discussed in this work. These are the unary decision and the best-of-N decision. Robots became involved in a unary decision after they independently had decided that some alternative to the *status quo* should be adopted by their entire system. In a best-of-N decision, the robots collectively compared a set of proposed alternatives and adopted the best one, but it implicitly was assumed by the best-of-N decision framework that the *status quo* was inadequate, and that it therefore must be replaced by the best alternative that could be found. This belief was simultaneously adopted by all of the robots at the behest of an external central controller: the operator that began each decision-making trial.

A more general approach to decision-making would be a combination of these two approaches. As individuals notice the need for a collective decision, they should seek out alternatives. Other robots that do not believe that a decision is required, however, will not participate in the search. If these robots are recruited by an advocator, they might still conclude that the *status quo* is sufficiently satisfactory that the alternative proposed by the advocator should be rejected, and thus they should re-enter the idle state after researching it. Therefore, an advocator might attempt to recruit many of its teammates, but if it is the only robot that believes that a decision is necessary, all of them would reject it and no progress would be made by its actions. This ability of idle robots to reject recruitment would prevent a dec-MRS from being subjected to a decision every time that one of its robots believed that a decision was needed.

If an advocator detects such stagnation (*e.g.* the apparent consensus for its favoured alternative was not increasing), it must conclude that its belief that a collective decision was required was incorrect, and therefore it too should abandon its decision and revert to the idle state. However, if a robot believes that a collective decision is required (*i.e.* it is in either the searching or advocating states), it should not reject recruitment and instead should be an active participant in the iterative recruitment process. Commitment need not be modified, and is the same here as before.

With this slightly modified framework, a dec-MRS could initiate either unary or best-of-N decisions asynchronously, as circumstances demanded, as recognized by the individual robots. Note that here, for quorum to be satisfied and thus commitment to occur, a sufficient proportion of the robots must agree not only that one of the alternatives is sufficiently good, but also that the *status quo* is sufficiently bad that a collective decision is required to replace it. If the individual robots of a dec-MRS were able to recognize the need for different types of decisions, then this general framework would enable a dec-MRS to autonomously interact with its environment as though it were a superorganism, providing the illusion of central control. This synthesis has yet to be implemented, and it is left here as a starting point for future study.

6.2.2 Recovery from Incomplete Commitment

A system that is based around stochastic processes will never enjoy 100% reliability. For example, some of the individual best-of-N decisions described in Chapter 5 were incorrect. Even a deterministic algorithm, however, once deployed in the real world would have some probability of error, since the deterministic responses of such a system would still be in response to the stochastic nature of the real world. One potential problem has not been addressed explicitly by the decision-making framework proposed by this work, and that is what should be done in the event of an incomplete commitment.

Just like any other gossip-style algorithm, commitment might fail to induce every robot to commit to the same alternative. The probability of this occurring can be decreased by increasing the length of the commitment timeout, but the probability of failure can never be reduced to zero. How should a robot respond to an incomplete commitment? There is an implicit question of economy here, since it must also be asked at what point should the loss of a few robots be a concern? When this is a minor inconvenience, it might be best to leave recovery to the lost robots themselves. For example, if a robot in a large a convoy traveling through a dangerous area gets lost, it probably is not worth risking the entire convoy to recover it. In other circumstances, the individual robots might be much more valuable, and thus the system would have a vested interest in retaining them at all costs. Given that robustness to individual failure is one of the most common arguments in favour of dec-MRS, it is more likely that the former scenario will be encountered than the latter. Nonetheless, autonomous recovery from this mode of failure would be a worthwhile problem to investigate.

6.3 Final Thoughts

A decentralized multiple-robot system *can* make system-level decisions through local peer-to-peer interactions. Furthermore, the individual participants in these decisions need not be complex, yet they collectively are able to make good decisions, better than they would be able to individually. In fact, a close examination of the decision-making framework proposed by this work reveals that the simplest case is not one robot, but two, as social interaction with at least one other robot is required in order for an individual to conclude that quorum has been satisfied. Implicit in this strategy is the assumption by each robot that it is not alone, that it is a member of some greater community that extends beyond the boundaries of each individual. In many ways, the key to making decisions in a decentralized system appears to be the practice of restraint: not acting on one's own conclusions, but based on those of one's teammates, thereby freeing the collective from the tyranny of over-eager, error-prone individuals. It is difficult not to be reminded of our own brains, composed of a myriad of individual nerve cells connected together, participating in their own peer-to-peer interactions to make their own system-level decisions. We compute *with* our brains, not *in* them. In the same way, this thesis has demonstrated that we can compute *with* a dec-MRS, not just within it.

Bibliography

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. *Computer Networks*, 38:393–422, 2002.

[2] Walter S. Avis, Patrick D. Drysdale, Robert J. Gregg, Victoria E. Neufeldt, and Matthew H. Scargill. *Gage Canadian Dictionary*. Gage Learning Corporation, 1983.

[3] Tucker Balch and Lynne E. Parker. Guest editorial. *Autonomous Robotics*, 8(3):207–208, 2000.

[4] R. Beckers, J. L. Deneubourg, and S. Goss. Trail laying behaviour during food recruitment in the ant *Lasius niger* (l.). *Insectes Sociaux*, 39:59–72, 1992.

[5] R. Beckers, S. Goss, J. L. Deneubourg, and J. M. Pasteels. Colony size, communication and ant foraging strategy. *Psyche*, 96:239–256, 1989.

[6] R. Beckers, O. E. Holland, and J.L. Deneubourg. From local actions to global tasks: Stigmergy and collective robotics. In R. Brooks and P. Maes, editors, *Artificial Life IV*, pages 181–189. MIT Press, 1994.

[7] Daniel S. Bernstein, Shlomo Zilberstein, and Neil Immerman. The complexity of decentralized control of Markov decision processes. In *UAI '00: Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, pages 32–37, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[8] E. Bonabeau, M Dorigo, and G. Theraulaz. Inspiration for optimization from social insect behaviour. *Nature*, 406:39–42, July 6 2000.

[9] Andrew F. G. Bourke and Nigel R. Franks. *Social Evolution in Ants*. Monographs in Behavior and Ecology. Princeton University Press, Princeton, New Jersey, 1995.

[10] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory, Special Issue of IEEE Transactions on Information Theory and ACM Transactions on Networking*, 52(6):2508–2530, 2006.

[11] S. Camazine, K. P. Visscher, J. Finely, and R. S. Vetter. House-hunting by honey bee swarms: Collective decisions and individual behaviors. *Insectes Sociaux*, 46:348–360, 1999.

[12] Scott Camazine, Jean-Louis Deneubourg, Nigel R. Franks, James Snyed, Guy Theraulaz, and Eric Bonabeau. *Self-Organization in Biological Systems*. Princeton Studies in Complexity. Princeton University Press, 1st edition, 2001.

[13] Jason D. Campbell, Padmanabhan Pillai, and Seth Copen Goldstein. The robot is the tether: Active, adaptive power routing for modular robots with unary inter-robot connectors. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, pages 4108–15, Edmonton, Alberta Canada, August 2005.

[14] Y. Uny Cao, Alex S. Fukunaga, and Andrew B. Kahng. Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, 4:1–23, 1997.

[15] V. Cerny. A thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985.

[16] Iadine Chades, Bruno Scherrer, and François Charpillet. A heuristic approach for solving decentralized-POMDP: assessment on the pursuit problem. In *SAC '02: Proceedings of the 2002 ACM Symposium on Applied Computing*, pages 57–62, New York, NY, USA, 2002. ACM Press.

[17] Larissa Conradt and Timothy J. Roper. Consensus decision making in animals. *TRENDS in Ecology and Evolution*, 20(8):449–456, 2005.

[18] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. In *PODC '87: Proceedings of the sixth annual ACM Symposium on Principles of distributed computing*, pages 1–12, 1987.

[19] J. L. Deneubourg, S. Goss, N. R. Franks, A. Sendova-Franks, C. Detrain, and L. Chrétien. The dynamics of collective sorting: Robot-like ants and ant-like robots. In *Proceedings of First International Conference on Simulation of Adaptive Behavior*, pages 356–365, 1990.

[20] Marco Dorigo and Alberto Colorni. The ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Cybernetics and Man – Part B*, 26(1):1–13, 1996.

[21] Marco Dorigo, Gianni Di Caro, and Luca M. Gambardella. Ant algorithms for discrete optimization. *Artificial Life*, 5(2):137–172, 1999.

[22] A. Dornhaus and N. R. Franks. Colony size affects collective decision-making in the ant *Temnothorax albipennis*. *Insectes Sociaux*, 53:420–427, 2006.

[23] A. Dornhaus, N. R. Franks, R. M. Hawkins, and H. N. S. Shere. Ants move to improve: Colonies of *Leptothorax albipennis* emigrate whenever the find a superior nest site. *Animal Behaviour*, 67:959–963, 2004.

[24] Gregory Dudek, Michael Jenkin, and Evangelos Milios. *Robot Teams: From Diversity to Polymorphism*, chapter 1: A Taxonomy of Multiple Robot Systems, pages 3–22. A K Peters, Ltd., 2002.

[25] Alessandro Farinelli, Luca Iocchi, and Daniele Nardi. Multirobot systems: A classification focused on coordination. *IEEE Transactions on Systems, Man and Cybernetics–Part B: Cybernetics*, 34(5):2015–2028, 2004.

[26] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42:142–166, 2003.

[27] N. R. Franks, A. Wilby, B. W. Silverman, and C. Tofts. Self-organizing nest construction in ants: Sophisticated building by blind bulldozing. *Animal Behaviour*, 44:357–375, 1992.

[28] Nigel R. Franks and Jean-Louis Deneubourg. Self-organizing nest construction in ants: Individual worker behaviour and the nest's dynamics. *Animal Behaviour*, 54:779–796, 1997.

[29] Nigel R. Franks, Anna Dornhaus, Jon P. Fitzsimmons, and Martin Stevens. Speed versus accuracy in collective decision making. *Proceedings of the Royal Society of London B*, 270:2457–2463, 2003.

[30] Nigel R. Franks, Eamonn B. Mallon, Helen E. Bray, Mathew J. Hamilton, and Thomas C. Mischler. Strategies for choosing between alternatives with different attributes: Exemplified by house-hunting ants. *Animal Behaviour*, 65:215–223, 2003.

[31] Nigel R. Franks, Stephen C. Pratt, B. Mallon, Eamonn, Nickolas F. Britton, and David J. T. Sumpter. Information flow, opinion polling and collective intelligence in house-hunting social insects. *Philosophical Transactions of the Royal Society of London Series B*, 357:1567–1583, 2002.

[32] Nigel R. Franks and Tom Richardson. Teaching in tandem-running ants. *Nature*, 439:153, 2006.

[33] Brian P. Gerkey and Maja Matarić. Sold!: Auction methods for multirobot coordination. *IEEE Transactions on Robotics and Automation*, 18(5):758–768, 2002.

[34] Brian P. Gerkey and Maja Matarić. Multi-robot task allocation: Analyzing the complexity and optimality of key architectures. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2003)*, pages 3862–3868, 2003.

[35] Deborah M. Gordon. *Ants at Work, How an Insect Society is Organized*. The Free Press, 1999.

[36] Piyush Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, 2000.

[37] Owen Holland. Personal correspondence, March 2006.

[38] Owen Holland and Chris Melhuish. Stigmergy, self-organisation, and sorting in collective robotics. *Journal of Adaptive Behaviour*, 5(2):173–202, 1999.

[39] Owen E. Holland. Grey Walter: The pioneer of real artificial life. In Christopher Langton, editor, *Proceedings of the 5th International Workshop on Artificial Life*, pages 34–44. MIT Press, 1997.

[40] Terry Huntsberger, Guillermo Rodriguez, and Paul S. Schenker. Robotics challenges for robotic and human mars exploration. In *Proceedings of ROBOTICS2000*, pages 340–346, 2000.

[41] Henry M. Robert III, William J. Evans, Daniel H. Honemann, and Thomas J. Balch. *Robert's Rules of Order Newly Revised*. Da Capo Press, 10^{th} edition, 2000.

[42] Auke Jan Ijspeert, Alcherio Martinoli, and Aude Billard. Collaboration through the exploitation of local interactions in autonomous collective robotics: The stick pulling experiment. *Autonomous Robots*, 11(2):149–171, 2001.

[43] David Kempe, Jon Kleinberg, and Alan Demers. Spatial gossip and resource location protocols. In *Proceedings of the 3 Annual ACM Symposium on Theory of Computing*, pages 163–172. ACM, 2001.

[44] David Kempe and Jon M. Kleinberg. Protocols and impossibility results for gossip-based communication mechanisms. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 471–480. IEEE Computer Society, 2002.

[45] Jelle R. Kok and Nikos Vlassis. Distributed decision making of robotic agents. In *Proceedings of the 8th Annual Conference of the Advanced School for Computing and Imaging*, pages 318–325, 2003.

[46] Kurt Konolige, Deiter Fox, Charlie Oritz, Andrew Agno, Michael Eriksen, Benson Limketkai, Jonathan Ko, Benoit Morisset, Dirk Schulz, Benjamin Stewart, and Regis Vincent. Centibots: Very large scale distributed robotic teams. In M. H. Ang and O. Khatib, editors, *Experimental Robotics IX: The 9th International Symposium on Experimental Robotics*, pages 131–140, 2006.

[47] C. Ronald Kube and Eric Bonabeau. Cooperative transport by ants and robots. *Robotics and Autonomous Systems*, 30(1/2):85–101, 2000.

[48] C. Ronald Kube and Hong Zhang. The use of perceptual cues in multi-robot box pushing. In *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, pages 2085–2090, 1996.

[49] Michael Lederman Littman. *Algorithms for Sequential Decision Making*. PhD thesis, Department of Computer Science, Brown University, 1996.

[50] E. B. Mallon, S. C. Pratt, and N. R. Franks. Individual and collective decision-making during nest site selection by the ant *Leptothorax albipennis*. *Behavioral Ecology and Sociobiology*, 50:352–359, 2001.

[51] James McLurkin, Jennifer Smith, James Frankel, David Sotkowitz, David Blau, and Brian Schmidt. Speaking swarmish: Human-robot interface design for large swarms of autonomous mobile robots. In *Proceedings of the 9th International Symposium on Distributed Autonomous Robotic Systems*, 2006.

[52] Chris Melhuish, Owen Holland, and Steve Hoddell. Collective sorting and segregation in robots with minimal sensing. In *Proceedings of the Fifth International Conference on the Simulation of Adaptive Behavior*, 1998.

[53] Chris Melhuish, Owen Holland, and Steve Hoddell. Convoying: using chorusing to form travelling groups of minimal agents. *Robotics and Autonomous Systems*, 28:207–216, 1999.

[54] Melissa B. Miller and Bonnie L. Bassler. Quorum sensing in bacteria. *Annual Review of Microbiology*, 55:165–99, 2001.

[55] George E. Monahan. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.

[56] Martin Mundhenk, Judy Goldsmith, Christopher Lusena, and Eric Allender. Complexity of finite-horizon Markov decision process problems. *J. ACM*, 47(4):681–720, 2000.

[57] Chris A. C. Parker and Hong Zhang. Biologically inspired decision making for collective robotic systems. In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, pages 375–380, September 2004.

[58] Chris A. C. Parker and Hong Zhang. Collective decision making: A biologically inspired approach to making up all of your minds. In *Proceedings of the 2004 IEEE International Conference on Robotics and Biomimetics (ROBIO 2004)*, pages 250–255, August 2004.

[59] Chris A. C. Parker and Hong Zhang. An analysis of random peer-to-peer communication for system-level coordination in decentralized multiple-robot system. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, pages 398–403, 2006.

[60] Chris A. C. Parker and Hong Zhang. Collective robotic site preparation. *Adaptive Behavior*, 14(1):5–19, 2006.

[61] Chris A. C. Parker and Hong Zhang. A practical implementation of random peer-to-peer communication for a multiple-robot system. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, pages 3730–3735, 2007.

[62] Chris A. C. Parker and Hong Zhang. Consensus-based task sequencing in decentralized multiple-robot systems using local communication. In *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, pages 1421–1426, July 2008.

[63] Chris A. C. Parker and Hong Zhang. Cooperative decision-making in decentralized multiple-robot systems: the best-of-N problem. *IEEE/ASME Transactions on Mechatronics*, 14(2):240–251, 2009.

[64] Chris A. C. Parker, Hong Zhang, and C. Ronald Kube. Blind bulldozing: Multiple robot nest construction. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, pages 2010–2015, October 2003.

[65] Christopher A. C. Parker. Distributed robotic construction without communication. Master's thesis, Department of Computing Science, University of Alberta, 2002.

[66] Lynne E. Parker, Yi Guo, and David Jung. Cooperative robot teams applied to the site preparation task. In *Proceedings of the Tenth International Conference on Advanced Robotics*, volume 5, pages 71–77, 2001.

[67] Maxim Peysakhov, Christopher Dugan, Pragnesh Jay Jodi, and William Regli. Quorum sensing on mobile ad-hoc networks. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*, pages 1104–1106, 2006.

[68] Stephen Pratt. Personal correspondence, 2007.

[69] Stephen C. Pratt. Quorum sensing by encounter rates in the ant *Temnothorax albipennis*. *Behavioral Ecology*, 16(2):488–496, 2005.

[70] Stephen C. Pratt, Eamonn B. Mallon, David J. T. Sumpter, and Nigel R. Franks. Quorum sensing, recruitment and collective decision-making during colony emigration by the ant *Leptothorax albipennis*. *Behavioral Ecology and Sociobiology*, 52:117–127, 2002.

[71] Stephen C. Pratt and David J. T. Sumpter. A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences*, 103(43):15906–15010, 2005.

[72] Stephen C. Pratt, David J. T. Sumpter, Eamonn B. Mallon, and Nigel R. Franks. An agent-based model of collective nest choice by the ant *Temnothorax albipennis*. *Animal Behaviour*, 70:1023–1036, 2005.

[73] D. V. Pynadath and M. Tambe. The communicative multiagent team decision problem: Analyzing teamwork theories and models. *Journal of Artificial Intelligence Research*, pages 389–423, 2002.

[74] D. V. Pynadath and M. Tambe. Multiagent teamwork: Analyzing the optimality and complexity of key theories. In *Proceedings of International Conference on Autonomous Agents and Multi-Agent Systems*, pages 873–880, 2002.

[75] Robocup robotic soccer. http://www.robocup.org, April 2005.

[76] R. Keith Sawyer. Artificial societies: Multiagent systems and the micro-macro link in sociological theory. *Artificial Societies*, 31(3):325–363, 2003.

[77] Thomas D. Seeley. Consensus building during nest-site selection in honey bee swarms: The expiration of dissent. *Behavioral Ecology and Sociobiology*, 53:417–424, 2003.

[78] Thomas D. Seeley and Susannah C. Buhrman. Group decision making in swarms of honey bees. *Behavioral Ecology and Sociobiology*, 45(1):19–31, 1999.

[79] Thomas D. Seeley and Susannah C. Buhrman. Nest-site selection in honey bees: How well do swarms implement the "best-of-n" decision rule? *Behavioral Ecology and Sociobiology*, 49:416–427, 2001.

[80] Thomas D. Seeley and P. Kirk Visscher. Choosing a home: How the scouts in a honey bee swarm perceive the completion of their group decision making. *Behavioral Ecology and Sociobiology*, 54:511–520, 2003.

[81] Thomas D. Seeley and P. Kirk Visscher. Group decision making in nest-site selection by honey bees. *Apidologie*, 35:101–116, 2004.

[82] Thomas D. Seeley and P. Kirk Visscher. Quorum sensing during nest-site selection by honeybee swarms. *Behavioral Ecology and Sociobiology*, 56:594–601, 2004.

[83] William Shakespeare. Romeo and Juliet, 2:2, 1597.

[84] D. A. Shell and M. J. Matarić. On foraging strategies for large-scale multi-robot systems. In *Proceedings of IEEE/RSJ Internationsl Conference on Intelligent Robots and Systems (IROS 2006)*, pages 2717–2723, 2006.

[85] Teambots home page. http://www.teambots.org, April 2005.

[86] Guy Theraulaz and Eric Bonabeau. Modelling the collective building of complex architectures in social insects with lattice swarms. *Theoretical Biology*, 177:381–400, 1995.

[87] P. Kirk Visscher and Scott Camazine. Collective decisions and cognition in bees. *Nature*, 397:400, February 1999.

[88] N. Vlassis, R. Elhorst, and J. R. Kok. Anytime algorithms for multiagent decision making using coordination graphs. In *Proceedings of International Conference on Systems, Man and Cybernetics*, 2004.

[89] W. Grey Walter. An imitation of life. *Scientific American*, 182(5):42–45, 1950.

[90] Christopher M. Waters and Bonnie L. Bassler. Quorum sensing: Cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology*, 21:319–346, 2005.

[91] Jens Wawerla, Gaurav S. Sukhatme, and Maja J. Matarić. Collective construction with multiple robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002)*, pages 2696–2701, 2002.

[92] Jan Wessnitzer and Chris Melhuish. Building adaptive structure formations with decentralized control and coordination. In *Proceedings of the 7th International Conference on Simulation of Adaptive Behavior*, pages 371–372, 2002.

[93] Jan Wessnitzer and Chris Melhuish. Collective decision-making and behaviour transitions in distributed ad hoc wireless networks of mobile robots: Target hunting. In W. Banzhaf *et al.*, editor, *European Conference on Artificial Life (ECAL) 2003*, pages 893–902. Springer-Verlag, 2003.

Appendix A

The Relationship between the Observed Quorum and the Quorum Threshold

When a robot measures the apparent consensus amongst its teammates for the alternative that it favours, it cannot determine it precisely. Instead, an estimate is computed, denoted $\tilde{C_a}$. It is this estimate that it compares to Q, the quorum threshold. Given the unavoidable noise in $\tilde{C_a}$, robots are likely to make errors when testing quorum. This appendix describes the approach used in this thesis to predict the relationship between the quorum threshold and the observed quorum, the latter being the actual consensus present in a MRS when the very first robot believes that $C_a \geq Q$. Only

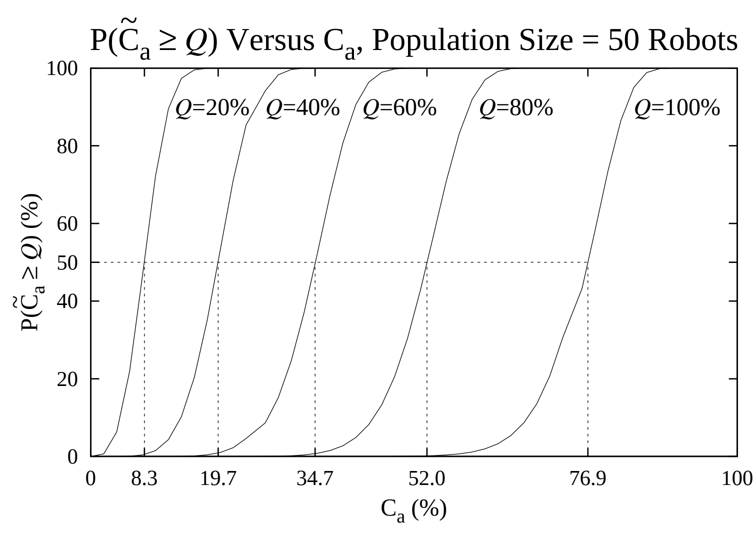


Figure A.1: The curves in this figure plot the probability of at least one of a system's N robots believing that quorum has been satisfied. Each curve corresponds to a different value of the quorum threshold, Q. The likely value of the observed quorum for a given value of Q is the apparent consensus (horizontal axis) that corresponds to a 50% likelihood of believing that quorum is satisfied. These values are indicated by the intersections of the dashed lines in this figure.

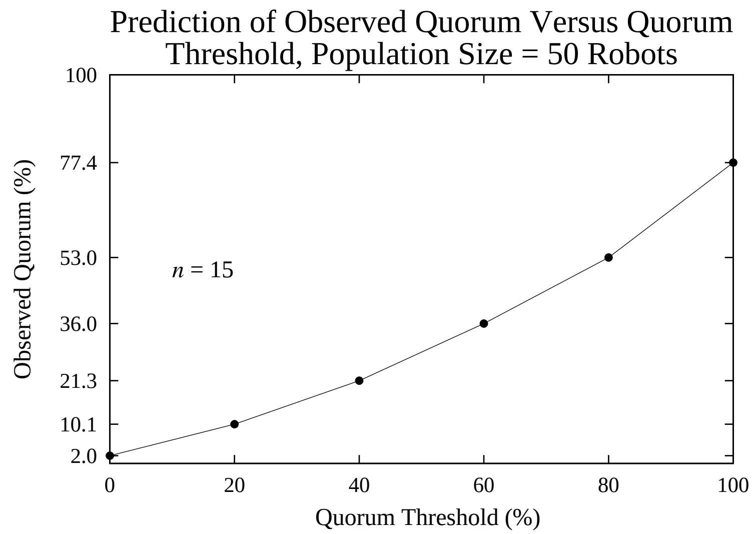


Figure A.2: The values of observed quorum read off of Figure A.1 are in the units of apparent consensus. These are converted to true consensus with Equation A.2, which are then plotted against the value of Q corresponding to the curve in Figure A.1 from which they were read. Because a robot will always believe that quorum is satisfied when $Q = 0$, the observed quorum for this particular quorum threshold will be $\frac{1}{N}$.

the digital approach to consensus estimation is used here, but given the results of Chapter 3, the predictions derived here should be considered equally applicable to a similarly calibrated analog quorum test.

Given the population size of a dec-MRS and the number of teammate opinions used by each robot to compute $\tilde{C}_a$, the probability of at least one robot computing $\tilde{C}_a \geq Q$ is given by Equation A.1[1]. Plotting Equation A.1 with $n = 5$ for several values of Q yields Figure A.1. Note that the curves in Figure A.1 take into account the number of robots that compose a dec-MRS, as does the entire derivation in this Appendix.

$$P(\tilde{C}_a \geq Q) = 1 - \left[\sum_{i=\lceil nQ \rceil}^{n} \binom{n}{i} (C_a)^i (1 - C_a)^{n-i} \right]^{(\lfloor C_a(N-1) \rfloor + 1)} \tag{A.1}$$

Each curve in Figure A.1 plots the probability of a robot $\tilde{C}_a \geq Q$ given the value of Q as C_a is varied from zero to 100%. By assuming that the most likely observed quorum for a given value of Q occurs when $P(\tilde{C}_a \geq Q) = 50\%$, the observed quorum for such a configuration can be read off of the horizontal axis of Figure A.1 at the point where the corresponding curve reaches a height of 50%. For example, the leftmost curve corresponds to $Q = 20\%$. On that curve, when $P(\tilde{C}_a \geq Q) = 50\%$, $C_a = 8.3\%$.

The values read off of the horizontal axis of Figure A.1, however, are measured in the units of apparent consensus. This is because the individual robots within a dec-MRS do not know the

[1]This is the same equation as Equation 3.24, reproduced here for convenience.

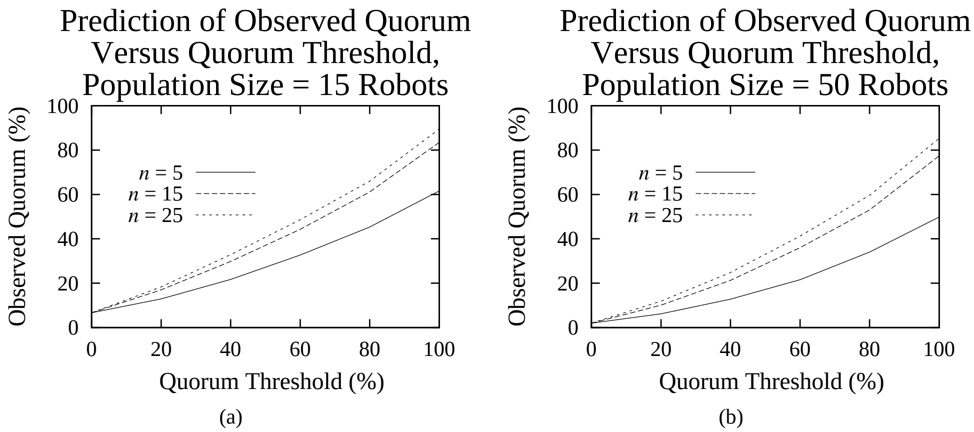


Figure A.3: These graphs plot the predicted relationship between observed quorum and the quorum threshold for two different population sizes for different values of n. As n is increased, the probability of a robot overestimating C_a decreases, which results in an increase in the observed quorum. Increasing the number of robots in a system decreases the observed quorum, both because apparent consensus is less of an overestimate of true consensus as N increases, and also because the probability of at least one of the robots making an error in its estimate of apparent consensus will tend to increase.

population size of their system, and so they cannot measure true consensus. However, when the observed quorum is reported, it is measured in the units of true consensus, since this is how it would appear to an agent observing it from outside of the system, such as a designer. Apparent consensus can be converted to true consensus by Equation A.2.

$$C_t = \frac{(N-1)C_a + 1}{N} \tag{A.2}$$

Returning to the leftmost curve of Figure A.1, for which $Q = 20\%$, the observed quorum equals $\frac{(50-1)\times 8.3\%+1}{50} = 10.1\%$. Therefore, when the robots of a 50-robot dec-MRS each compute $\tilde{C}_a$ using the 15 most recently received teammate opinions, and test quorum by comparing $\tilde{C}_a$ to $Q = 20\%$, an observed quorum of 10.1% should be expected. By repeating this process for different values of Q, the relationship between the observed quorum and the quorum threshold is built up, point by point. Figure A.2 portrays the relationship computed from the curves of Figure A.1. The intercept of this curve with the vertical axis is more easily found. Because $\tilde{C}_a \geq 0$, a robot will always believe that quorum is satisfied when $Q = 0$, independent of n. Therefore, the observed quorum will be $\frac{1}{N}$, since the first robot to test quorum is guaranteed to conclude that it has been met when $Q = 0$.

Increasing n increases the precision with which each robot estimates apparent consensus. Robots that use a larger value of n will be less likely to believe that quorum is satisfied for a given $C_a < Q$. Of course, because the intercept with the vertical axis is independent of n, each curve corresponding to a different value of n will share a common intercept, and they will diverge as Q increases. This is precisely the relationship that is shown in the graphs of Figure A.3.

The population size of a system also impacts the relationship between observed quorum and the quorum threshold, as is evident when Figures A.3(a) and A.3(b) are compared. Figure A.3(a) corresponds to a 15-robot system, whereas Figure A.3(b) corresponds to a 50-robot system. Increasing N lowers the curves for two different reasons. First, as N increases, apparent consensus is less of an overestimate of true consensus. This effect is most noticeable for low quorum thresholds. In particular, the intercept of Figure A.3(a) with the vertical axis is $\frac{1}{15}$, whereas it is $\frac{1}{50}$ in Figure A.3(b). Second, as N increases, more robots will simultaneously compute an estimate of C_a and compare their estimates to Q, and so the likelihood of at least one of them overestimating C_a also will increase. This also lowers the expected value of observed quorum for a given quorum threshold.

The prediction of the relationship between the observed quorum and the quorum threshold derived in this appendix ignores one fact, and that is that it takes time for the robots to collect the n teammate opinions that they use to compute $\tilde{C}_a$. During the period of time in which the robots gather vote-messages from each other, C_a will tend to evolve as the robots independently change their opinions based on their observations of the world, or as they recruit each other in a best-of-N decision. If, as was the case in the experimental domain of Chapter 4, C_a tends to increase with time, then during the time required to collect the extra teammate opinions demanded by a greater value of n, C_a will increase, and the observed quorum will increase along with it. The more rapidly C_a tends to change relative to the time required to collect teammate opinions, the more the experimentally derived relationship between the observed quorum and the quorum threshold will tend to deviate from the predictions presented here.