A multi-robot system for adaptive exploration of a fast changing environment: probabilistic modeling and experimental study

Aude Billard[2,1], Auke Jan Ijspeert[2,1], Alcherio Martinoli[1]

[1] LAMI, Swiss Institute of Technology (EPFL)
IN-F Ecublens, 1015 Lausanne, Switzerland
Tel: +41-21-693-39-07, Fax: +41-21-693-52-63
[2] - Instituto Dalle Molle di Studi Sull'Intelligenza Artificiale,
Corso Elvezia 36, CH-9600 Lugano
Aude.Billard@epfl.ch,
http://diwww.epfl.ch/w3lami/team/billard/

Abstract. Is it more efficient to use one or several robots? Will the performance of a group of robots working in a collaborative task be enhanced if the robots can communicate with one another? What learning abilities should the robot(s) be provided with for adapting to a continuously changing environment? We address these three issues in a specific task, namely learning the topography of an environment whose features change frequently. We propose a theoretical framework based on probabilistic modeling to describe the system's dynamics. The adaptive multi-robot system and its dynamic environment are modeled though a set of probabilistic equations. The model gives an explicit description of the influence of the variables of the system, namely the number of worker robots, the frequency of environmental changes and the environment's configuration, on the data collecting performance of the group. It is then used to determine boundaries for these system's variables within which the learning is successful. Further, we implement the multi-robot system in experiments with a group of Khepera robots and in simulation using Webots, a 3-D simulator of the Khepera robots. The robots are controlled by a distributed architecture. Each robot's controller is based on an associative memory type of learning algorithm. Results show that the algorithm allows a group of robots to keep an up-to-date account of the environmental state when this changes regularly. Finally, the predictions of the probabilistic model are compared to the results of simulations and physical experiments. It is found that the model shows both a good qualitative and quantitative correspondence to these results. This suggests that a probabilistic model can be a good first approximation of a multi-robot system when this system's behavior is stochastic in nature.

Keywords: Learning in a dynamic environment - multi-robot system - probabilistic modeling - local and global communication

1 Introduction

Numerous works on autonomous robot systems investigate the questions of 1) whether it is more efficient to distribute the area of expertise needed for performing a complicated task between several robots rather than designing a unique expert robot [6, 10, 19]; 2) whether the use of explicit communication could improve the performance of a group of robots in a collaborative task ([1], [2], [9], [20], [25]); 3) what learning abilities should the robot(s) be provided with for adapting to a continuously changing environment [8, 17, 21, 24].

We address these three issues in a specific task, namely learning the topography of an environment whose features, the locations of objects, change frequently. A group of *worker robots* search constantly the environment. The robots are provided with an associative memory which allows them to store the locations of the objects as they detect them. They transmit to each other the coordinates of the objects' locations by locally broadcasting the location when finding an object or when meeting another robot. The information gathered by each robot is also transmitted to a static *database robot* which each robot visits regularly. The database robot keeps an up-to-date account of the global state of the environment of which each robot has only a partial knowledge. We study the system's performance in a dynamic environment, in which the locations of the objects change with a constant frequency.

We propose a theoretical framework based on probabilistic modeling to describe the system's dynamics. The adaptive multi-robot system and its dynamic environment are modeled though a set of probabilistic equations. The model gives an explicit description of the influence of the variables of the system, namely the number of worker robots, the frequency of environmental changes and the environment's configuration, on the data collecting performance of the group. It is then used to determine boundaries for these system's variables within which the learning is successful.

Further, we implement the multi-robot system in experiments with a group of Khepera robots and in simulation using Webots, a 3-D simulator of the Khepera robots [16]. The robots are controlled by a distributed architecture. Each robot's controller is based on an associative memory type of learning algorithm. The algorithm is simple and, as such, makes no contribution to connectionist architecture. It is however well indicated for the task as results show that the algorithm allows a group of robots to keep an up-to-date account of the environmental state when this changes regularly and very frequently (we study periods of changes of a few seconds).

We investigate different communication types (local versus global, point-to-point and broadcast protocols) in order to show the generality of the learning system. We use simple communication protocols (see section 2.3) and make no attempt to define the optimal communication strategy for the given situation. For other works in that direction, the reader might refer, for instance, to Yoshida et al. [25] who determine theoretically and in simulations the optimum communication range for a multi-robot system. See [1, 9, 22] for robotics studies of the importance of communication in robot's control.

The work presented in this paper brings three new contributions to research in collective robotics. First, it develops an abstract representation of a multi-robot system by modeling it as a set of probabilistic equations. In this respect, our work was inspired by studies of multi-agent systems which developed probabilistic models to represent biological systems (ants' society [5]) and robotic systems ([12, 25]). To our knowledge such a framework has seldom been applied to collective robotics problems (to the exception of [12, 25]) and never to a mapping task, nor has it been compared to physical experiments (exception being our previous work [12]). Finally, the present work brings a new contribution compared to [12], as here we develop a parameter free probabilistic model; that is the system's dynamics is completely and explicitly described by the model. In [12], the system's dynamics was implicitly defined by a set of probabilities determining the system's state transitions. The state of the system along time was then evaluated by running simulations.

The second contribution of this work lies in its giving different levels of modeling of the problem and its comparing the predictions of each model. We propose first an abstract/theoretical representation of the problem through the probabilistic model and then compare this to a practical approach of the problem through a sensor-based simulation and physical experiments. While there exist several theoretical modelings of multi-agent systems ([7]) and numerous multi-robot experiments, it is seldom the case that theoretists come to implement their model or that roboticists present a theoretical prediction of their experimental results. This is not due to these authors' neglect, but to the important constraints in cost and time required to develop the hardware of physical experiments. In our experiments, the physical set-up of the experiments is highly simplified as we use "tiny" robots (requiring not much room) and two virtual sensors (see section 2)[1]. The theoretical model we propose is a first attempt to model a simple collective task.

The third contribution of our work is to investigate mapping of a *dynamic* environment. There have been several studies of multi-robot systems engaged in the mapping of a environment (e.g. [4, 23]) in which, however, the considered environment was static. Note though that those studies often investigated environments whose complexity was much higher than ours. We use round and square arenas with no obstacles apart from the robots. Our study aims only at giving an example of how a probabilistic model can be used to model multi-robot systems whose dynamics is stochastic in nature.

The rest of this paper is divided as follows. Section 2 describes the experimental set-up and the robots' controllers used in the experiments. Section 3 gives the equations of the probabilistic model for the two sets of experiments (local and global communication). Section 4 presents the results of the first set of experiments (simulations using local communication range) and compares those to the predictions of the probabilistic model. Section 5 reports on the second set of experiments (simulations and physical experiments with global communica-

[1] The physical experiments of this paper were made possible thanks to the expertise of Alcherio Martinoli.

tion). Section 6 concludes the paper with a short summary and discussion of the results.

2 The experimental set-up

2.1 Set-up for experiments with local communication

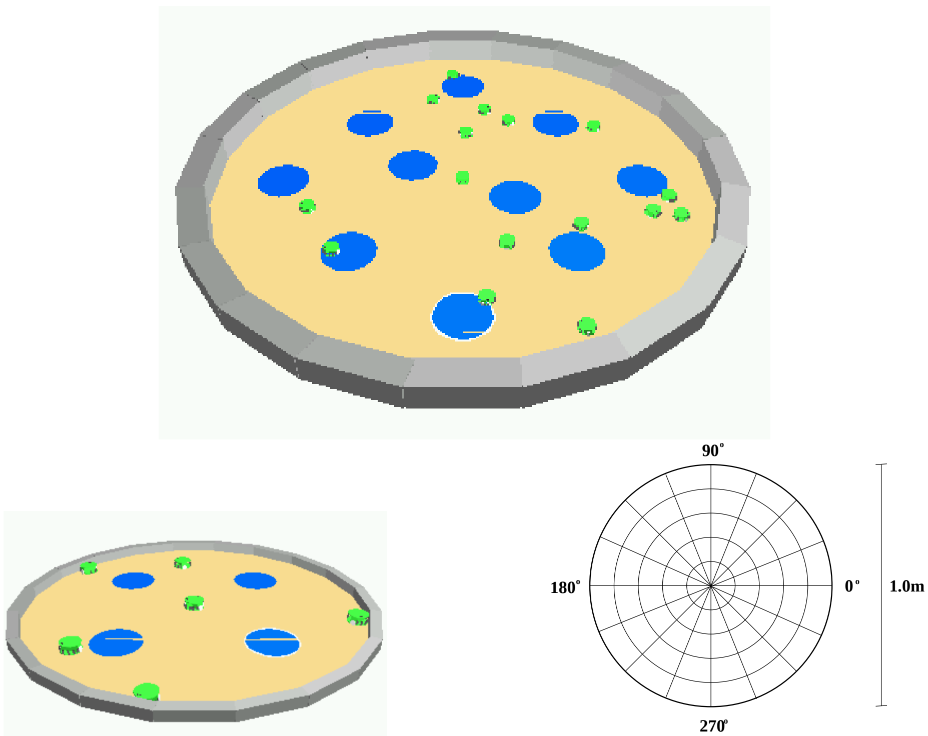

Fig. 1. Set-up of simulations of section 4: Arenas of 2 meters (top) and 1 meter (bottom left) of diameter with respectively 5 and 15 *worker* robots. The database robot stands in the center of each arena. There are 10 and 4 objects in the big (top) and small (bottom left) arenas respectively represented as patches of 0.1m and 0.07m diameter lying on the floor. **Bottom right** Division of the small arena into 80 zones.

The experiments with local communication (section 4) are carried out in Webots [16], a 3-D simulator of the Khepera [18] robots. In the first set of experiments (section 4) simulation studies are realized in two circular arenas of 1 meter and 2 meters diameter respectively as shown in figure 1. The Khepera robot is round with a diameter of 5.5cm. Thus, we study the exploration strategy in environments which are 1600 and 400 times the robot's size.

The simulator gives a relatively faithful representation of the Khepera robots, by incorporating imprecise movements of the robots' wheels (slipping) and introducing noise in the robots' sensors measurements as measured on the real Khepera robots. Each robot is provided with 9 infra-red (IR) sensors (8 are used to detect other robots and the arena walls, the 9th IR is activated only by the walls and allows to distinguish between robots and walls), a detector of ground color (used to distinguish between zones with/without objects), a radio transceiver (418 MHz + baudrate 9600), a compass with 5^0 degrees precision and one odometry counter on each wheel.[2] Compass and odometry sensors are used by the robots to determine their location relative to the center of the arena. The robots reset their position to the correct one each time they meet the database robot or hit a wall in the arena. The odometry errors are therefore contained within a range of up to 10 percent error.[3] The objects' locations, given as an angle and a distance relative to the center, are determined following a scaling of the arena into $5 \cdot 16 = 80$ (small arena) and $10 * 16 = 160$ (big arena) zones, see schema of figure 1. Thus, the objects' locations are known within a precision of 22.5 degrees (for the angle) and 10 cm (for the distance). Note the new object locations after each update are chosen randomly in the simulations. The database robot is a static Khepera robot placed in the middle of the arena (see figure 1). The robots' controller is described below in section2.3.

2.2 Set-up for experiments with global communication

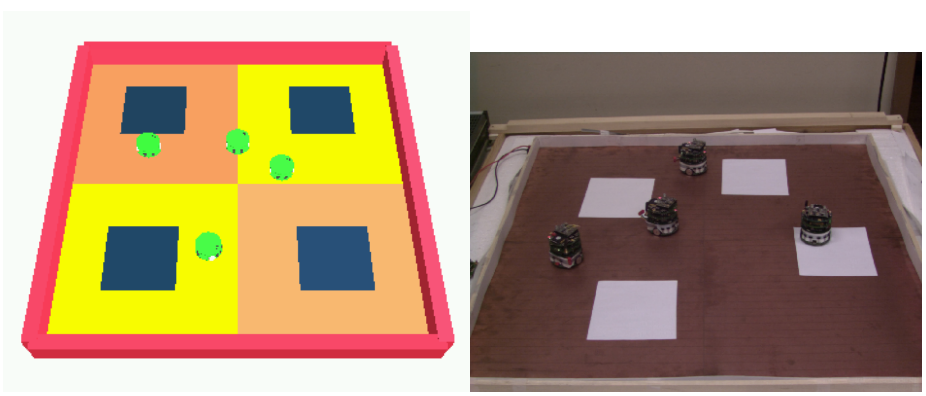

Fig. 2. **Set-up of section 5:** Square arena of 78 by 78 cm with 4 robots (Webots simulation (left) and physical set-up (right)). There are 4 objects represented as square patches.

[2] All sensors used in the simulations exist in miniature and could be used for the real Khepera robots.

[3] For a wide discussion on odometry error for a similar system, see [14, 15]

The experiments with global communication (section 5) are carried out in an square arena of 78 by 78 cm. A photo of the physical set-up and a picture of its implementation in Webots is given in figure 2. The experiments are realized with 1 to 4 Khepera robots. A technical description of the robots can be found in [18]. In the experiments, the robots are provided with eight infra-red sensors (six in the front and two in the back with which they perform obstacle avoidance, see section 2.3). They communicate via radio using a Motorola radio transceiver (418 MHz, 4800 Baud rate). In these experiments, the database robot is a radio transceiver placed on the bench and connected to a external Workstation. The database robot (Khepera or Sun station) runs a C program with the same learning algorithm as the one used by the worker robots (it will be described in section 2.3). The learning performance of the group is evaluated in sections 4 and 5 in reference to the data recorded by the database robot.

The robots also use a KPS (Khepera Positioning System [11]) for determining their position with 5mm precision and orientation (virtual compass with a precision of 5 to 10 degrees).[4] The robots are continuously powered by the arena's floor, which is covered with an electrified copper bands (the robot runs on 5V using about 400 mA) [13]. The electrical contacts with the ground are not perfect (contact may break when the robot makes a sharp turn or strong acceleration). These occasional electrical shorting are compensated by the robot's on-board battery. This set-up provides the robot with an autonomy open-end, tested up of 3 days [11]. Although the robots are provided with ground color detectors and can thus distinguish between patches on the floor (the objects in the simulation) and the rest of the arena, the objects in the physical experiments are defined virtually (that is, the locations of the patches are predefined in each robot's controller). Because the object locations had to be changed every 50 seconds (see section 5), it was not possible to change the patches manually and to keep the required frequency without disturbing too much the robots (the experimenter's hand is an obstacle for the robot). It was also not possible to develop new hardware for that pupose (set of flashing lights). In these experiments, the broadcast of the robots is recorded by a radio receiver attached to a Sun station.

2.3 The robots' controllers

In both experiments, all worker robots have the same controller which consist of five modules (see figure 3): 1) an *obstacle avoidance module* which consists of a one-layer real value feed forward neural network with eight input units (one for each infra-red sensor measurement) and two output units for the two motors (speed proportional control of the motors); 2) a *memory-based exploration*

[4] Note that in the simulations, the robots compute their position by odometry. Odometry was also possible on the real Khepera robots, as the robots are also provided with wheel counters. However, because the robots could not distinguish between a wall and another robot, we could not use the reset strategy of the simulation. As the experiments lasted for about 45 minutes, the odometry errors would have become enormous and the virtual compass was not sufficiently precise to compensate for this.

module which determines the robot's direction of travel when crossing the border between two zones of the arena (following the division represented in figure 1); each robot keeps track of the number of times it has crossed each zone; when it estimates that it has reached the border between two zones, the robot turns towards the zone it has less visited so far; the angle of turn is chosen randomly between 0 and PI 3) a *communication module* for the broadcast of the object locations; in the experiments on local communication (section 4), a robot can communicate in two occasions: when it discovers an object, it broadcasts locally (within a limited range) the location of the object; when it meets another robot, if transmits (using point-to-point protocol, i.e. with acknowledgement from the receiver robot) the location of one object chosen randomly over all locations it knows; in the experiments on global communication (section 5), the robots communicate only when they discover an object (broadcast with infinite range). 4) an *odometry module* which calculates the robot's position relative to the database (center of the arena) given the measurements of the wheels' counters and of the compass; 5) a *learning module* which consists of a bidirectional associative memory; the robots keeps track of the objects' locations by associating the two outputs of the odometry module which are the angle θ and the distance ρ, the polar coordinates of the robot relative to the center of the arena. Each connection of the module between an angle and a distance measurement is bidirectional and is associated with two parameters, a weight $w_{ij} = w_{ji}$ and a time parameter $\tau_{ij} = \tau_{ji}$. The associative module takes two binary (1/0) vectors as inputs; the vectors encode the robot's measures of angle and distance (following the arena's scaling, see figure 1); there is one active bit per vector at any point of time (e.g. *dist-vect*= $[00100] = 30[cm]$ and *dist-vect*= $[00010] = 40[cm]$). The weights w and time parameters τ are two matrices of 10 by 16 units (for simulations in the big arena) and of 5 by 16 units (for simulations in the small arena). The experiments starts with all weights w and time parameter set to zero. The learning algorithm is a system of three rules:

1. **Learning by seeing:**
   If the robot detects an object, then
   $$w_{\theta,\rho} = 99 \text{ and } \tau_{\theta,\rho} = t$$
   where t is the time measured by the clock of the robot[5].
2. **Forgetting:**
   If the robot crosses a location given by θ, ρ such that $w_{\theta,\rho} > 0$ but does not detect a object, then
   $$w_{\theta,\rho} = 1 \text{ and } \tau_{\theta,\rho} = t$$
3. **Learning by hearing:**
   If the robot hears the location of an object as told by another robot, then:
   $$\text{If } \tau'_{\theta',\rho'} > \tau_{\theta',\rho'} \text{ then } w_{\theta',,\rho'}, = \frac{1}{2} \cdot (w_{\theta',\rho'} + w'_{\theta',\rho'}) \text{ and } \tau_{\theta',\rho'} = \tau'_{\theta',\rho'}$$

[5] The clock is incremented at each processing cycle and is set at zero when the experiment starts.

$\theta^{'}, \rho^{'}, w^{'}_{\theta^{'},\rho^{'}}, \tau^{'}_{\theta^{'},\rho^{'}}$ are the distance, angle, weight and time parameter transmitted by the emitter robot.

The learning for one robot is evaluated by counting the number of correctly memorized object locations. Learning is successful when this number is equal to the number of different locations. An object located at the coordinates $\{\theta, \rho\}$ is considered as correctly memorized when the weight $w_{\theta,\rho}$ is greater than a threshold H. H is calculated at each time step as a function of the current value of all the weights w: $H = \frac{\frac{M(w)-m(w)}{2}+M_A(w)}{2}$, where $M(w) = \max_{w>0}(w)$ and $m(w) = \min_{w>0}(w)$ are the maximum and minimum values of weights for all $w > 0$. $M_A(w) = \mathrm{mean}_{w>0}(w)$ is the arithmetic mean calculated over all $w > 0$. H estimates the threshold between the important weights (close to 99) which correspond to strong correlations and the small weights (close to 1) which are noisy or discarded correlations. The calculation is based on the hypothesis that the distribution of the correct and incorrect memorizations are uniform, that is equiprobable on all pairs. This is to some extend correct as there is no a-priori bias on the number of times the robots can observe each object. In other words, we assume that the number of weights above and under the threshold will be equal on average. Thus, the threshold should be the median value of weights. However, as the process of forgetting and learning moves the weights continuously from one side to the other of the threshold, it is important also to consider the mean value of the weights which indicates the level of transitions. The average of mean and median values is a plausible and simple (but not optimal) approximation of the actual value of the threshold.

Note that there is no total forgetting of an object's location, that is, the weight associated to the corresponding pair of coordinates never returns to zero. The minimal value for a weight which has been updated once is 1. This value can never be greater than the threshold; therefore, a location associated with a weight equal to 1 is always discarded, i.e.considered as no longer valid. Keeping track of all objects' locations which have once been discovered results in the robots constantly checking the validity of the location (following rule 2). Forgetting and relearning of locations is thus made faster as compared to a strategy which would only increase the weights of newly discovered locations and never decrease those of no longer valid ones.

Note that the calculation the division of the weights by 2 (in rule 2) is to some extend arbitrary and one might query whether it would not be more efficient to store directly the newly told location. It is however not certain that this would be the case. There is uncertainty as to whether what the robot is being told is still true. What the robot has seen is "more true" than what it is told. As the later is a combination of what the communicative robot has seen and has been told about. The imprecision in the objects' detection adds up through the telling process. The division by two is a simple (but not optimal) way of representing this imprecision.

More studies should be done to determine the optimal values for the algorithms parameters. However, the fact that the probabilistic model, which assumes a perfect learning algorithm, match so well the experimental results (as it will be demonstrated in sections 4 and 5) suggests that using an optimal learning algoritm (whose convergence might be faster) would not necessarily bring a significative improvement of the learning of the group. For a discussion of the similitude/difference this algorithm bears to other connectionist models, the reader can refer to the description of DRAMA[3] (a connectionist architecture for on-line learning of spatio-temporal regularities and of time series in autonomous robots) of which the present algorithm is a simplification.

The comparison of the time parameters in rule 3 prevents a robot from replacing information it has acquired earlier with that transmitted by another robots when the transmitted information is older than its own. The database robot's controller consists of the same learning module as that used in the worker robots. In the experiments on local communication (section4), when worker and database robots meet, the worker robot transmits to the database robot its two matrices of weights and of time parameters (all w and τ, while in a two worker robots communication, only one pair (w, τ), corresponding to one location, is transmitted). Following rule 3, the database robot updates its knowledge by calculating the mean value between its current set of weights (collected from another robot) and those newly transmitted, iff the new information is more recent than its current one. The database robot transmits then back to the worker robot the mean matrix of weights and time parameters. After a meeting with the database, a worker robot has therefore the same global knowledge of the environment as the database. This speeds up the forgetting process as the robot can then verify more locations (all the locations which have been recorded by the group) than only those it had stored itself. This exchange of matrices weights does not occur in the experiments on global communication (section 5). In that case, the database robot stores the locations of the newly discover objects as soon as this one is broadcasted by the robot which has discovered it, as the robots' broadcast is audible by all (unless interference). All worker robots can pick up this signal. Thus, worker robots and databse robot have almost the same knowledge at all time (small differences exist due to the interferences) and there is no need for the database robot to send any information in return to the one it picks up.

3 The probabilistic model

In this section, we determine sets of probabilistic equations to model the learning dynamics of the experiments using local communication (section 4) and global communication (section 5). The model is based on the assumption that the information gathering process (learning of the locations) is essentially a stochastic process based on simple geometrical considerations. It assumes that the exact trajectories of the robots, the details of the learning and communication events can be ignored and that the result of the learning can be represented as a set of probabilities of occurrence.

3.1 Local communication

The aim is to define an equation which will allow us to determine T the minimal time for the database robot to learn the locations of N_s objects, given that there are N_r worker robots, that the arena has size A and that an object covers a surface S_s.

We define the building blocks or fundamental probabilities of the model by considering the geometrical configurations of the system. We define the probability of meeting the database robot (P_{db}) as the ratio of the surface of detection of the database robot by another robot S_d over the arena's surface A, i.e. $P_{db} = S_d/A$. Similarly, the probabilities of meeting another robot ($P_r = S_r/A$), of passing across a source ($P_s = S_s/A$) or of being in the range of communication of another worker robot ($P_c = S_c/A$) are the ratios of the surfaces of each of these objects over the arena's surface ($A = \pi \cdot (r^2)$, $r = 0.5[m]$ or $r = 1[m]$ for small/big arenas), $S_d = \pi \cdot 0.1^2[m^2]$ (small arena),, $S_d = \pi \cdot 0.15^2[m^2]$ (big arena) $S_r = \pi \cdot 0.1^2[m^2]$, $S_s = 0.0038[m^2]$ (small arena), $S_s = 0.0078[m^2]$ (big arena), $S_c = \pi \cdot 0.3^2[m^2]$.

Let $P_{\text{success}}(N_r, N_s, T)$ be the probability that the event "the database robot has recorded N_s locations" has occurred after a time T. This event is true if each of the N_s locations have been seen by at least one robot and been transmitted to the database robot at least once in a time T, i.e.:

$$P_{success}(N_r, N_s, T) = 1 - (1 - P_{L-success}(N_s, T))^{N_r}$$

$P_{\text{L-success}}(N_s, T)$ is the probability that a first event "all N_s locations have been transmitted" has occurred within a period $T - t_1$, given that a second event "all N_s locations have been learned" has happened in a period t_1. This conditional probability can be expressed as follows:

$$P_{\text{L-success}}(N_s, T) = P(\textit{see DB in } T - t_1 \textit{ (and) learn object in } t_1)$$

The two events are independent, thus the probability of their co-occurrence for a given pair $\{t_1, T - t_1\}$ is the product of each event's probability. The total probability is the sum over all possible pairs $\{t_1, T - t_1\}$ (time is discretised) of this product:

$$P_{\text{L-success}}(N_s, T) = \frac{\sum_{t_1=1}^{T} P_{\text{learn-object}}(N_s, t_1) \cdot P_{\text{see-database}}(T - t_1)}{\sum_{t_1=1}^{T} P_{\text{learn-object}}(N_s, t_1)} \quad (1)$$

The probability of meeting the database robot, $P_{\text{see-database}}(T - t_1)$ in equation 1, is the probability of crossing the surface S_d within a period $T - t_1$:

$$P_{\text{see-database}}(T - t_1) = 1 - (1 - P_{db})^{T - t_1}$$

$P_{\text{learn-object}}(N_s, t_1)$ is the probability that the event "a robot has learned N_s object locations in a time t_1" is true. A robot learns about an object's locations if the robot either sees the object P_{S} or hears its location from another robot P_{h}.

$$P_{\text{learn-object}}(N_s, t_1) = (1 - (P_{\text{not-learn-object}})^{t_1})^{N_s}$$

$$P_{\text{not-learn-object}} = (1 - P_{\text{S}}) \cdot (1 - P_{\text{h}})^{N_r - 1}$$

The probability of hearing an object's location from another robot's broadcast is the probability that the three following events are true: 1) the listener robot is within an area S_c around the emitting robot (S_c is the surface within which the communication is audible) and 2) the emitting robot broadcasts the particular location, 3) no other robot out of the $N_s - 1$ (excluding the emitting robot, including the listener robot) is simultaneously emitting in that same area (P_{Interf} is the probability of this event).

$$P_h = P_{\text{Hear}} \cdot P_{\text{Interf}}$$

$$P_{\text{Interf}} = (1 - P_{\text{Hear}})^{N_s \cdot N_r - 1}$$

$$P_{\text{Hear}} = (\frac{Sc}{A}) \cdot (1 - P_{\text{Not-emit}})$$

Event 2 is true if the emitter robot either sees the object (it then broadcasts the location) or if the emitter robot meets another robot which transmits it that particular location. In the later case, the emitting robot chooses 1 location among the $2 * N_s$ it knows, which included the N_s correct and no longer valid locations. Event 2 can occur only if the robot has seen that object within a time $t_o < t_1$ before meeting another robot. It follows:

$$P_{\text{Not-emit}} = (1 - P_s)^{t_1} \cdot (1 - \frac{1}{2 \cdot N_s} \cdot \frac{\sum_{t_o=1}^{t_1} (1 - (1 - P_s)^{t_o}) \cdot (1 - (1 - P_r)^{t_1 - t_o})}{\sum_{t_o=1}^{t_1} 1 - (1 - P_s)^{(}t_1 - t_o)})$$

In the above equations, the unity of surface is the meter and the unity of time corresponds to the time needed to cover the surface S_s (which is the minimal surface considered in the equations). In order to convert the value of time in seconds, T has to be multiplied by $\frac{S_s}{V_r * D_R}$, where $V_r = 0.16[m/s]$ is the maximal speed of the robots and $D_r = 0.055[m]$ is the diameter of the robot.

3.2 Global communication

Experiments of section 5 are modeled by determining the probability that the event "N_s locations are correctly transmitted in a time T". The difference with the previous model is that, in this experiment, a robot broadcasts an object location only when it sees the object and not when it meets another robot. The above event is true when the broadcast of each of the N_s locations (made by one of the robot which was visiting the surface S_s) has been correctly picked up by the database robot at least once (i.e. no interference occurred). Referring to the reasoning regarding the previous model, the probability of this event is:

$$P_{success}(N_r, N_s, T) = (1 - (1 - (P_s \cdot (1 - P_s)^{N_s \cdot (n-1)}))^{n \cdot T})_s^N \qquad (2)$$

$P_s = S_s/A$. An object surface is $S_s = 0.35 * 0.35/4[m^2]$ and the arena surface is now $A = 0.78 \cdot 0.78 = 0.61[m^2]$.

4 Experiments with local communication

This section reports on simulation experiments carried out in round arenas as shown in figure 1. Communication is local, that is the broadcast of the robot can be picked up only within an area $S_c = 2 \cdot \pi \cdot 0.3^2[m^2]$. First, a set of simulations studies is carried out in a static environment in order to determine the minimal time delay $T$for learning all the four and ten locations of each arena. The results of these experiments are compared to the prediction of the probabilistic model of section 3 in order to evaluate the model's validity. Further, we evaluate the performance of the multi-robot system at adapting to a dynamic environment by carrying simulations while varying the period of change of the object locations.

4.1 Probabilistic model versus simulations

Webots simulations were carried out in a static environment (i.e. the locations of the objects did not change). The number of worker robots was varied from 1 to 10 and from 1 to 15 in the small and big arenas respectively. For each set-up (i.e. for a given arena and a given number of robots) 10 different runs were carried out with a different random seed. A run simulated 1000 seconds. We measured the mean time delay after which the database robot knew all 4 (small arena) and 10 (big arena) object locations. In figure 4, we compare the prediction of the probabilistic model and the results of the simulations. As one would expect, the more robots the faster the learning. However, the relation between these two variables is not linear and the increase of time efficiency saturates for important numbers of robots. Thus, if one would consider implementing the system in a real robotic set-up based on these results, one would determine the optimal number of robots by comparing the gain in time efficiency to the cost increase when augmenting the number of robots.

For both small and big arenas, the results of the probabilistic model and of the simulations are qualitatively and quantitatively similar. This means that, although the probabilistic model is a crude representation of the system, it approximates well the correlations between the main system's variables. Two aspects of the simulations are, however, not represented by the model: 1) the probabilistic model assumes a uniform coverage of the space, where all points of the space are visited with the same frequency; this does not take into account the boundary effects due to the walls which make the center of the arena (i.e. the database robot's location) an area more often visited than the exterior of the arena (this effect is augmented by the exploration strategy which gives preference to moving towards the database robot); in order to represent this effect in the model, we increased S_{db} compared to its real geometrical value so that we obtained the same probability of meeting the database robot ($P_{db} = S_{db}/A$) as that measured in the simulations. 2) The probabilistic model assumes that learning of an object's location is perfect, i.e. when a robot sees the object, it learns its location; this does not take in account the imprecise determination of the location due to odometry error[6].

The probabilistic model is a good first approximation of the studied system. It allows one to determine the optimal efficiency of the system in the ideal case; for instance, it can be used to estimate the minimal number of robots, as well as the minimal battery life time they should be provided with for the robots to collect a given number of informations from an environment of a given size. However, more realistic simulations, such as those realized in Webots, need to be carried out, in order to evaluate the importance of the above mentioned aspects which are not represented by the model.

4.2 Learning in a dynamic environment

Simulations were carried out in a dynamic environment, in which the objects changed locations periodically. For each arena, 15 different periods P of environmental change were tested, 8,16,24,...,120 seconds in the small arena and 28,56,...,420 seconds in the large arena[7]. In each case, 3 runs were carried out for a duration corresponding to $5000 \cdot P$. We ran simulations with groups of 1, 3 and 5 robots in the small arena and 5, 10 and 15 robots in the big arena. Figures 5 top left and 5 top right show the mean number of correctly and incorrectly learned locations over the whole run for the small and big arenas respectively. The results for each three configuration of robots are superimposed. For P less than 40sec. (small) and 140sec. (big), the database knows on average about 50% of the correct locations, while still taking for correct almost 50% of the locations which are no longer valid. For those periods, the environment changes faster than the minimal time delay T required for the robots to learn all the locations. The minimal T was measured in the simulations of section 4.1 (see figure 4) as a minimum of 40 and 120 seconds for small and big arenas respectively (the measures were consistent with the probabilistic predictions). For P greater than 40 seconds (small arena) and 140 seconds (big arena), the proportion of correctly learned locations increases steadily while the proportion of incorrectly learned locations decrease by the same proportion. There is almost no difference between the three different robot configurations in each case. This is due to the fact that the minimal time delay for finding all objects is almost the same for

[6] The odometry errors are in fact negligible given the resetting strategy, see section 2.1.

[7] The minimal steps of 8 and 28 seconds were chosen because it was estimated to be the minimal time to finding an object; it was calculated as the mean distance between the objects divided by the distance traveled by a robot in one time step.

these three robot configurations (see figure 4)). Thus, it appears that there is almost no benefit in use 15 rather than 5 robots in the big arena and 5 robots rather than 1 in the small arena. However, figure 4 shows that it is more efficient to using 5 robots rather than 1 in the big arena.

Figures 5 bottom shows the progression of the learning of the database robot along a run (results of the simulation in small (left) and big (right) arenas with 5 and 10 robots respectively for $P = 40 and 140$). The curve varies from zero (no locations known yet) to the maximum (4 and 10 correctly known locations in small and big arenas). In the simulations, the objects are not displaced simultaneously. The period at which each object is displaced is constant (it is P) but the phase at which the first displacement occurs is different for each object. This explains the fact that the learning curve does not always decrease until zero (the new locations being discovered and transmitted before all locations have been changed). These results demonstrate that a multiple robots system based on an associative memory learning algorithm, as described in section 2.3, is successful at learning the topography of an environment and updating its knowledge when the environment constantly changes.

5 Experiments with global communication

In this section, experiments are carried out with robots communicating globally in a square arena (see figure 2) with four objects. Similarly to experiments of section 4, we first carried out a set of simulations studies and physical experiments in a static environment in order to determine the minimal time delay for learning the four locations. The results of these experiments were then compared to the prediction of the probabilistic model of section 3 in order to evaluate the model's validity. Then, we carried out sets of physical experiments in a dynamic environment when varying the period of environmental change around the minimal one determined by the study in static environment. These experiments were meant to validate the learning algorithm in a dynamic *physical* environment when using global communication. Experiments of section 4.2 validates the algorithm in local communication.

5.1 Learning in a static environment

50 runs were carried out in simulation and then in a physical set-up using groups of 1,2,3 and 4 robots (10 runs for each robot configuration, with different random seeds in the simulations and different starting positions for the robots in the physical experiments). Figure 6 shows superimposed the mean time delay for learning all the four locations.

We observe a good agreement between the three plots, with the different implementations (physical set-up, Webots simulations and probabilistic model) showing the same behavior qualitatively. As expected, the mean value of minimal time T decreases with an increase of the number of robots. The standard deviations also decrease with an increase of the robots' number. This means that the influence of the randomness of the robots' walk on the learning success decreases when there are more robots, as it is compensated by the learning redundancy due to having more robots.

Note that in order to get a good qualitative fit of the probabilistic model, the probability of seeing a source $P_s = S_s/A$ had to be increased by 20% compared to its original value (see section 3)[8]. This value corresponds to that measured in the simulations. As mentioned in section 4.1, the boundary effects (bumping of the robots on the wall) are not represented in the probabilistic model. These effects are taken in account by reducing the dimension of the visited surface of the arena A, increasing consequently the value of P_s. Thus, as in section 4.1, it appears that the probabilistic model is a good first approximation of the system's dynamic. However, simulations or physical experiments need to be carried out in order to evaluate correctly the primitive probabilities of the system (an estimation based on geometrical considerations being not sufficient, especially in small arenas in which boundary effects have more importance).

The good correspondence between the results of the Webots simulations and those of the physical experiments further confirm the validity of that simulation (the later had already been demonstrated in a multiple robots clustering task [12]). It is however important to stress the fact that the physical experiments still used a number of artifacts, such as a virtual definition of the object locations and the use of an external positioning system instead of odometry calculation (this was done in the simulations). Further experiments using robots relying only on physical sensors should be carried out in order to validate the learning system for real world applications.

5.2 Learning in a dynamic environment

12 (4 times 3) runs were carried out in the physical set-up of Khepera robot, using groups of 1,2,3 and 4 robots respectively with 3 different periods of environmental changes (50, 100 and 200 seconds). A run lasted for 10 environmental changes. Figure 7 shows the learning performance (mean number of correctly and incorrectly learned locations over the whole run) of the 4 groups of robots for the 3 update periods. The plotted data are the locations recorded by the database robot (here the stand-alone radio station connected to a Sun station) following the broadcast of the worker robots (see explanations of section 2.2). The learning performances of the four different groups of robots are qualitatively and quantitatively similar; the correctedness of the learning improves when the period of environmental changes increases (leaving more time for the robots to discover all sources). The more robots, the better the learning on average, i.e. the percentage of learning success (figure 7 top) and of learning failure (figure 7 top) are bigger and smaller respectively the more robots we use. However, the

[8] If the probability P_s is not increased, the probabilistic model gives an estimation of T 20% bigger than the measured one.

gain in using more robots is not important (the standard deviations of the four curves superimpose[9]), as this was shown previously in figure 6.

Figure 8 shows the number of correctly learned locations (as recorded by the Sun station) along a run for the four different robot configurations. In all graphs, the period of change was 100 seconds, which is bigger or equal to the average minimal time required for learning all locations (as shown in figure 6). We observe that the fluctuations of the learning decreases as the number of robots increases. This correlates to the observation made in section 5.1 that the standard deviation of T (shown in figure 6) decreases with an increase of the number of robots. This means that the redundancy in the learning due to using more robots reduces the variability of the results due to the randomness of each robot's trajectory (imprecise and highly variable infra-red sensor readings result in random movements of the robot, see section 2.3).

6 Conclusion

This paper presented a multi-robot system capable of learning the topography of an environment whose features changed regularly. A learning algorithm was proposed, composed of learning and forgetting processes. It was implemented in simulation with 1 to 15 robots and in a real set-up of 1 to 4 Khepera robots. Results showed that the multi-robot system was able of keeping an up-to-date account of the environmental state when this changes regularly. The change period was of the order of a few seconds to a few hundreds of seconds. The system's dynamics was modeled through probabilistic equations. The equations give an explicit description of the correlations between the variables of the system, namely the number of robots, the frequency of environmental change and the environments' configuration. It was used to determine the minimal time delay for correct learning in each of the different configurations of robots and environments (round and square arenas of different sizes with different number of objects) which were tested in simulation and with physical experiments. The prediction of the probabilistic model were shown to agree qualitatively and quantitatively to the results of these experiments. Thus, the probabilistic model is a good first approximation of the system. However, simulations or physical tests remain necessary for determining the exact values of some of the basic probabilities in the model.

7 Acknowledgement

Many thanks to Luca Gambardella and anonymous reviewers for useful comments which helped improving the writing of this paper. Lots of thanks to the LAMI, EPFL which provided the facilities and technical support for these experiments. This research was supported by a grant of the Swiss National Research Foundation, for the project "A methodology for collective robotics design".

[9] We did not plot the error bar of the graph in figure 7 for reason of clarity of the figure, as these bar superimpose.

References

1. T. Balch & R. C. Arkin, (1994), 'Communication in reactive multiagent robotic systems', *Autonomous robots Journal*, 1, pp. 27-52.
2. Billard A. & Dautenhahn K., (1999), 'Experiments in social robotics: grounding and use of communication in autonomous agents', accepted for publication by *Adaptive Behaviour Journal* (August 1998).
3. Billard, A. and Hayes, G., 1999, "DRAMA, a connectionist architecture for control and learning in autonomous robots", *Adaptive Behaviour Journal*, vol. 7:1.
4. W. W. Cohen, (1996), 'Adaptive mapping and navigation by teams of simple robots', *Robotics and autonomous systems*, 18, pp. 411-434.
5. E. Bonabeau, G. Theraulaz & J-L Deneubourg, (1998), 'Fixed response thresholds and the regulation of division of labor in insect societies', *Bulletin of Mathematical Biology*, **60**, pp. 753-807, Article No. bu980041.
6. G. Dudeck, M. Jenkin, E. Milios and D. Wilkes, (1996), 'A taxonomy for multi-agent robotics', *Autonomous Robots*, 3, pp.375-397.
7. J. M. Epstein and R. Axtell, "Growing Artificial Societies: Social Science from the Bottom Up", The MIT Press, Cambridge, MA, 1996.
8. Floreano, D., (1998), 'Evolutionary Robotics in Artificial Life and Behavior Engineering', in T. Gomi (Ed.), Evolutionary Robotics, Ontario (Canada): AAI Books, 1998.
9. M. J Mataric, (1998), 'Using Communication to Reduce Locality in Distributed Multi-Agent Learning', *Journal of Experimental and Theoretical Artificial Intelligence*, special issue on Learning in DAI Systems, Gerhard Weiss, ed., **10(3)**, Jul-Sep, pp. 357-369.
10. M. J Mataric, (1995), 'Issues and Approaches in the Design of Collective Autonomous Agents', *Robotics and Autonomous Systems*, **16(2-4)**, Dec. 1995, pp. 321-331.
11. A. Martinoli, "Swarm Intelligence in Autonomous Collective Robotics: From Tools to the Analysis and Synthesis of Distributed Control Strategies", PhD Thesis, Computer Science Department, EPFL, Lausanne, oct. 1999. To appear.
12. A. Martinoli, A. Ijspeert, F. Mondada, (1999), 'Understanding collective aggregation mechanisms: from probabilistic modelling to experiments with real robots', *Robotics and Autonomous Systems*, Elsevier (to appear).
13. A. Martinoli, E. Franzi, and O. Mattey. "Towards a reliable set-up for bio-inspired collective experiments with real robots." In A. Casals and A. T. de Almeida, editors, *Proc. of the Fifth International Symposium on Experimental Robotics ISER-97*, pages 597-608, Barcelona, Spain, June 1997. Springer-Verlag.
14. P. Maechler, "Robot Odometry Correction Using Grid Lines on the Floor", Proceedings of MPCA, Pisa, Italy, 1997
15. P. Maechler, "Robot Positioning by Supervised and Unsupervised Odometry Correction", PhD Thesis No. 1858, Computer Science Department, EPFL, Lausanne, 1998.
16. O. Michel, (1998), 'Webots: a Powerful Realistic Mobile Robots Simulator', Proceeding of the Second International Workshop on RoboCup, LNAI, Springer-Verlag. http://www.cyberbotics.com/
17. A. Murciano & J. del R. Millan, (1997), 'Learning signalling behaviour and specialisation in Cooperative agents', *Adaptive Behaviour Journal*, **5:1**, pp. 5-28.
18. F. Mondada, E. Franzi & P. Ienne, (1993), 'Mobile Robot Miniaturisation: a Tool for Investigation in Control Algorithms', Proceedings of ISER'93, Kyoto, Japan, October 1993. Also at http://diwww.epfl.ch/lami/robots/K-family/.

19. L. E. Parker, (1994), 'An Experiment in Mobile Robotic Cooperation', Proceedings of the ASCE Speciality Conference on Robotics for Challenging Environments, February, pp. 131-139.
20. L. E. Parker, (1994), 'Heterogeneous Multi-Robot Cooperation', PhD-thesis, Massachusetts Institute of Technology, January, 1994.
21. Pfeifer, R. and Scheier, C., (1998), 'Sensory-motor coordination: the metaphor and beyond', R. Pfeifer, R. Brooks (eds.), *Robotics and Autonomous Systems*, special issue on practice and future of autonomous agents, **20:2-4**.
22. Premvuti, S. and Wang, J., "A medium access protocol, supporting wireless inter-robot communication in distributed robotic systems", proceedings of IEEE 1994, pp. 1050-4729.
23. I. M. Rekleitis, G. Dudeck, E. E. Milios, "Multi-robot exploration of an unknown environment, efficiently reducing the odometry error", Proc. of Int. Joint Conf. on Artificial Intelligence, Nagoya, Japan, August 1997. Morgan Kaufmann, pp. 1340-1345.
24. Thrun, S., (1996), 'Explanation-Based Neural Network Learning - A Lifelong Learning Approach', Kluwer Academic, Boston, MA.
25. E. Yoshida, T. Arai, M. Yamamoto and J. Ota, (1998), 'Local communication of multiple robots: Design of optimal communication area for cooperative tasks',*Journal of Robotics Systems* 15(7), pp. 407-419.

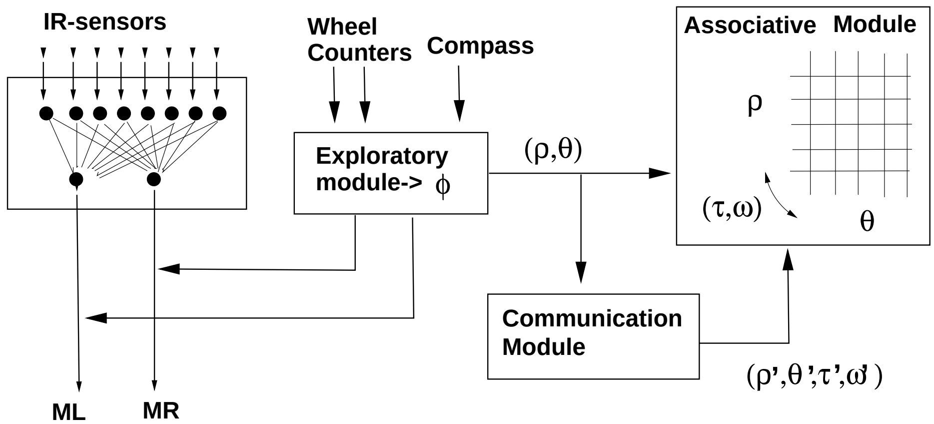

Fig. 3. The robots' controller.

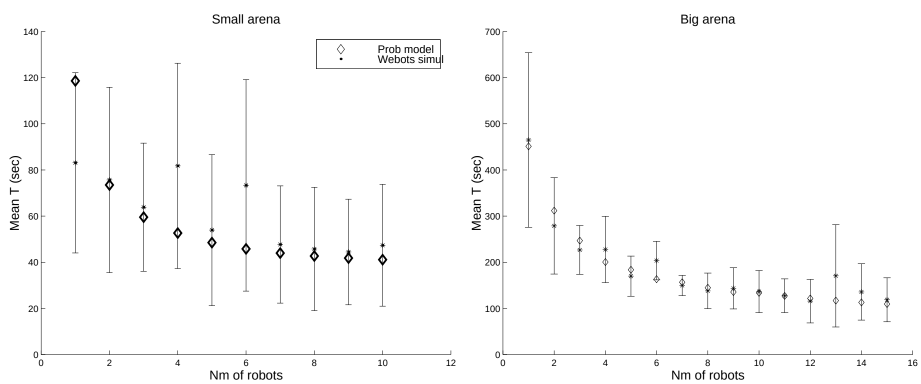

Fig. 4. The Y-axis represents the mean (over 10 runs) time delay T of the database robot to learn all object locations. The X-axis is the number of robots. Each figure compares the prediction of the probabilistic model (diamonds points) and of the Webots simulations ('$*$' point with error bars) in the small arena (left) and in the big arena (right).

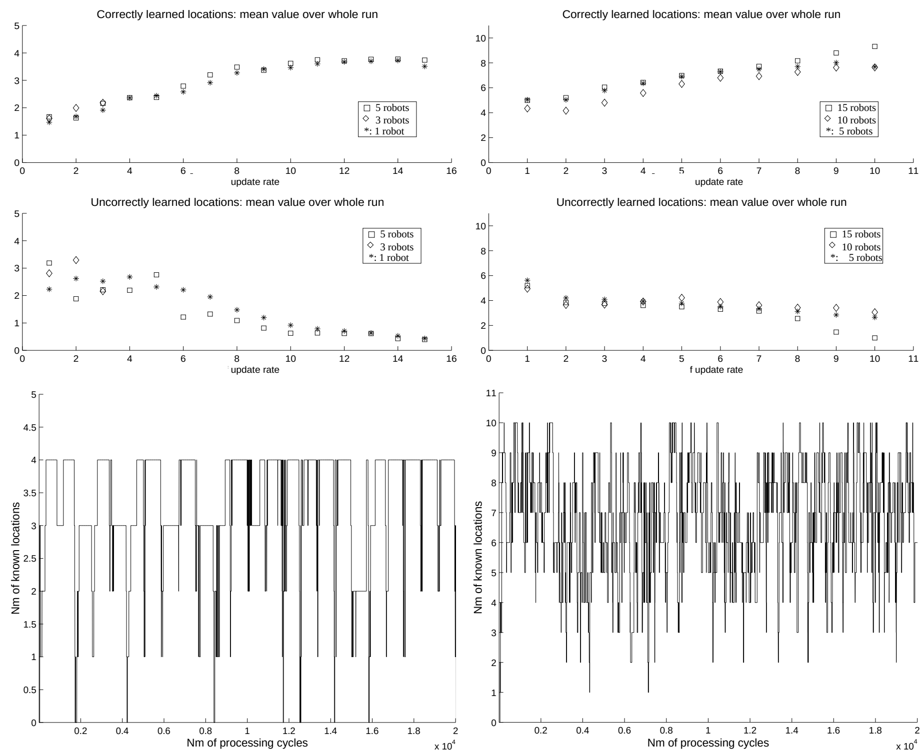

Fig. 5. **Top:** Mean number of correctly and incorrectly learned locations over the whole run for small (left) and big arenas (right); superposition of the results for three robots' configurations (left: 1,3,5 robots; right: 5,10,15 robots). X-axis is the update rate equal to $P/5$. **Bottom** State of the database's knowledge (number of known locations) along a run. 5 robots configuration in small arena (left) and 10 robots configuration in big arena (right). 1 processing cycle is 0.05 second.

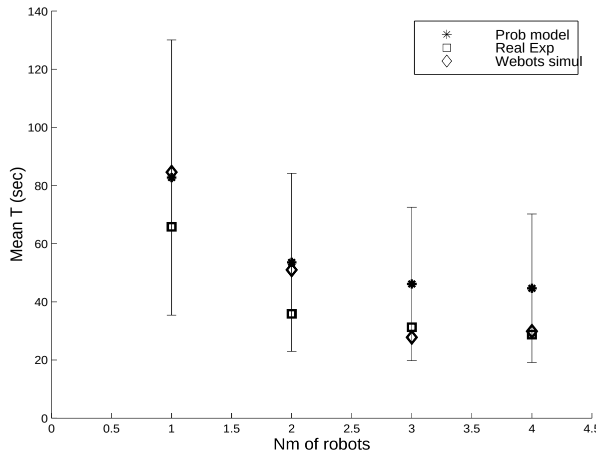

Fig. 6. The Y-axis represents the mean (over 10 runs) time delay T and the X-axis is the number of robots. The prediction of the probabilistic model is compared to the results of the Webots simulations and the physical experiments.

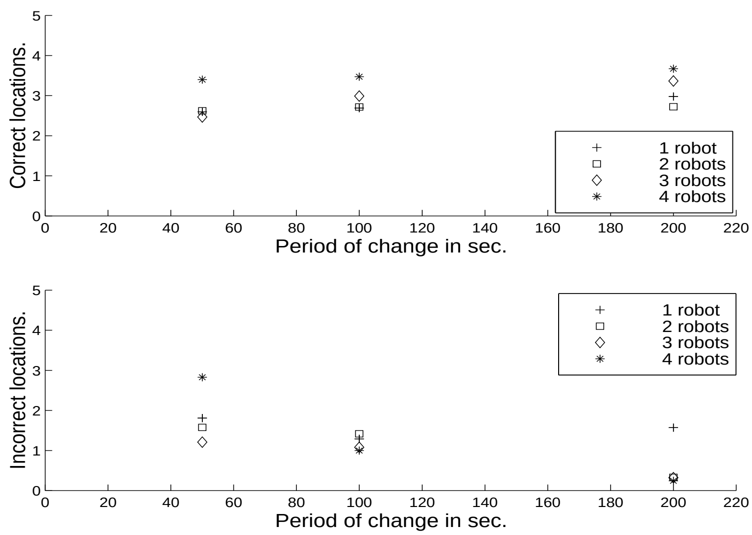

Fig. 7. Mean number of correctly and incorrectly learned locations over the whole run; superposition of the results for experiments with 1,2,3 and 4 robots.

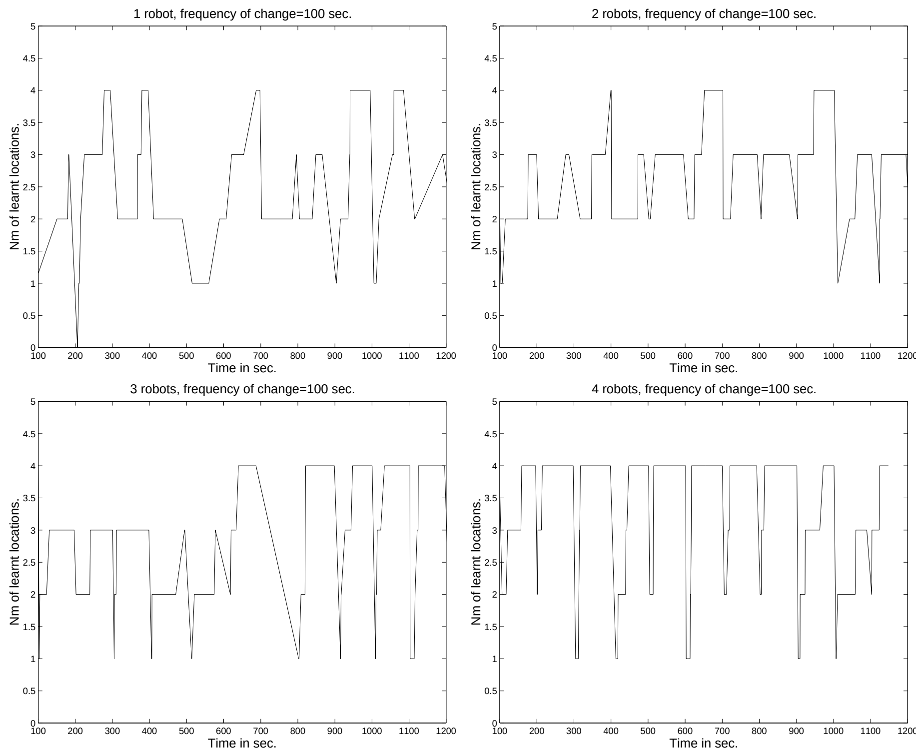

Fig. 8. State of the robots' knowledge (number of known locations) along a run. From left to right, top to bottom: experiments with 1,2,3 and 4 robots when the object locations change with a period of 100 seconds.