My Wonderful Article in IEEE Format for IROS2000

I.M.Author	**M.Y. Coauthor**
My Department	Coauthor Department
My Institute	Coauthor Institute
My Address	Coauthor Address
My e-mail	Coauthor e-mail

Abstract

Realistic applications of Autonomous Robotics need to face a lot of difficulties in real environments. To navigate, self-localize and cooperate in such contexts, a multi-robot system has to find a way to represent and communicate about its world. In this paper we discuss the necessity and the advantages of grounding representations and communications of a robots team in real perceptions. We describe a technical framework designed for this purpose, based on the classification of the perceptions of a robots population. Percepts must have specific properties to enable grounding, and we present particular percepts well suited to this process. Finally we discuss the possible usages of grounded classes and propose a grid representation for fusion and exchange.

Keywords: Multi-Robots Systems, collective representations, grounding, robot perceptions.

1 Introduction

Realistic applications of autonomous robotics need to face the difficulties that arise in most indoor environments. Real environments are at least partially unknown, they are not prepared for robots, and they are subject to dynamic changes. In those environments perceptions of robots are uncertain, robots build and follow imprecise maps, robots get lost, they are starving while looking for energy, their knowledge becomes blurred during long term operation. Beside this, most realistic applications require the deployment of several robots in the same environment, for cleaning, surveillance, exploration, guided tour, delivery, animation... This is needed at least to scale to normal environments size and often needed by the cooperative nature of the tasks.

We believe that concrete problems such as navigation, localization and map building benefit of being envisaged from the outset in the multi-robot perspective. If this is envisaged later the potential solutions may not be adaptable to a multi-robot context and will not take advantage of the robot team's distributed point of view and multi-sensor capacities. In order to face those problems, we argue in the first section of the paper, that a shared medium for representation and communication between robots has to be discovered in the available set of real perceptions. In the second section a technical framework for collective representations is described, based on the classification of the robots population perceptions. The third section presents particular perceptions, well suited to grounded representations. Preliminary results of percept classification are given to illustrate the approach. In the last chapter possible usages of grounded collective representations are discussed and Percept Distribution Grid are proposed to support cooperation in dynamic environments.

This work is part of the larger project Microbes [1] headed by Alexis Drogoul at LIP6 in Paris(Computer Science lab. of university P6). The main aim of the Microbes project are to enable the long-term adaptation of a robots colony in human standard environments and to study human/robots interactions and arising problems apart from a specific task context. The robots used in the project are Pioneer 2DX equipped with 16 ultrasonic sensors, a ccd color camera, compass and radio modem. All the computing is performed on-board on the embedded Linux-based PC. The project takes place in the AI-laboratory, a typical office area of 4000 m^2.

2 The need for grounded collective representations

A lot of activities performed by a robot such as mapping, learning pathways, involve the use of some kind of representations. Those representations do not necessarily refer to a human abstract model. Beside this, a multi robot system benefits to mix up and collate robots partial informations and therefore use some communication scheme, for instance when specifying a meeting point, or building a global map from partial ones. It is clear that cooperative activities require the use of a shared medium where the communication can take place. In simulated robotic systems a shared medium can be easily found using the underlying precise model of the world and robots can communicate precise locations, precise perceptions. However in a real multi-robots team, deployed in an unknown environment, this is not the case. Robot locations are imprecise and sometimes unknown, sensors are noisy, not necessarily following Gaussian models and due to the dynamicity of the environment unpredictable events happen. Also in a world that resists to an immediate symbolic transcription, the robots may have difficulties to communicate initially with symbolic tokens. So basically what will the robots use to communicate ?

Our hypothesis is that, in a real multi robot system involved in concrete tasks, there should be a tight relation between the perceived world, the internal representations and the communication language elements, and that this bottom-up chain should be anchored in the real perceptions. As an example, the problem of self-localization and the problem of sharing the description of a particular location between robots are similar, they both rely on finding sufficiently stable features in the environments. Generally we could say that talking 'per se' about the world (representing) and talking to others about the world are two facets of the same problem, both require an initial material, strongly related to what is possible to do and to perceive in the environments. Grounding in robotic systems is a promising direction, as pointed out by Brooks [2], and it is the natural emanation of today's ideas, that progressively blur the mind-body dualism. In Collective Robotics somehow similar orientations are proposed in papers such as [3, 4], however those papers do not address the problem of long term adaptation in large environments.

Of course the internal representations used by a robot can be made of human categorized landmarks (doors, corners, crosspoint, lanes ..) and the communication chunks be list, graphs,... constituted with those elements. But, particularly in an unknown, dynamic environment, trying to ground representation and communication in real perceptions has the following advantages:

- The robots elaborate and communicate about what they are 'used' to perceived and not what we guess they should perceive.
- This prevents us from developing specific feature recognition mechanisms which are too much environment-specific.
- This prevents us to follow an eventually misleading human 'a priori', therefore giving a better environment adaptation.

For this purpose we favor the emergence of a collective substratum that can be used as an elementary material for representation and communication.

3 Percept class language

The proposed substratum is a rudimentary language: a finite set of classes whose 'meaning' is well known in the robot team. The set of classes is obtained by classifying the percepts of the robot population with a self organizing classifier. Several authors [5, 6] have used unsupervised classifiers to classify perceptions of a single robot (ie: sets of sonar values) in a robust way. After sufficient learning, percepts are input in the classifier and the output classes are placed as landmarks indicators in a topological map. Here, in addition to this method we broadcast the percepts over the robot team to obtain homogeneity and use rather different percepts (see next chapter). We also choose to consider only percepts that are sufficiently singular: we do not take into account too frequent percepts (not informative) or too rare (spurious) ones. The class set is built in four steps (see fig 1):

- *Perception*: each individual produces percepts regularly while exploring its environment.
- *Broadcast*: the produced percepts are broadcasted (through radio modem) to all other members of the team, each of them constitutes a set of its own percepts plus the received ones.
- *Singularity Filtering*: the set of percepts is filtered so as to retain only percepts which are sufficiently singular regarding the whole set.

[width=.7]fig1.eps

Figure 1: Each robots broadcasts its percepts to others. Individual percepts and received ones are input to an unsupervised classifier

- *Classification*: each robot classifies the percepts set. After classification, the resulting class set is the same for each individual.

After a sufficient maturation phase each robot can use the classifier to categorize its own percepts. The finite class set constitutes an elementary language whose signification is well known among the robots and anchored in the possible perceptions. More elaborated representations (maps) can later be built using those basic bricks.

The filtering and the classification may be performed with a Kohonen Self-Organizing Map (SOM)[7]. A SOM classifier is constituted by a 2D lattice of nodes. Each node i corresponds to a class and stores a reference vector $\vec{m}_i$, a representative of the class. The class c of a new input vector $\vec{x}$ can be obtained by comparison to reference vectors using appropriate distance:

$$c = argmin_i \|\vec{x} - \vec{m}_i\|$$

During learning, the reference vectors are updated to reflect the input vectors distribution, for each new input vector, the winning node and neighbors are moved towards the input vector value. At the end of the learning process the reference vectors are topologically ordered.

A SOM classifier has the following interesting properties:

- *Topology preservation* : the similarity relations that exist between the percepts is maintained between their corresponding classes. It permits the use of a distance between classes.
- *Unsupervised learning* : autonomous exploration is possible.
- *Error quantification* : the distance between an input vector and its class reference vector : $\|\vec{x} - \vec{m}_c\|$, it can be used to reject irrelevant percepts (ie: caused by temporary obstacles).

In the above process, the filtering step is required because we want to focus on the singular aspects of the environment, we dont want to take into account percepts which are too frequent. The SOM classifier over-fits the most frequent percepts, and thus the less frequent ones have a higher quantification error. We use this property for filtering, with a first SOM classifier. The percepts are classified a first time and the average quantification error for all percepts is computed. Then the percepts with a quantification error inferior to the average error (over-fitted) are eliminated. In the classification step remaining percepts are simply classified again using SOM.

In order to be anchored in the complete (eventually changing) environment the building of the substratum has to be coupled with an appropriate exploration behavior and a regular update of the classifier.

4 Choosing appropriate percepts

The word ‘percept’, here, refers to the result of some transformation applied to the raw sensors. This can be horizontal lines extracted from video image, derivative of sonar values over the time to detect sudden variations, precise landmark detections such as doors.... The transformations can involve several sensors, sonar and odometry for instance, they can also incorporate internal states information, or they can come from an active perception behavior. While senses are most of the time immediate (ie : instantaneous sonar range), percepts can integrate information over time, they can contain informations that are currently out of reach (after passing a door).

In a grounded system the choice of an appropriate transformation from sense to percept is necessarily crucial because it orientates the whole system. As an indication, we can refer to the importance of this transformation in animal nervous systems [8].The track followed is that the transformation is not necessarily to obtain more accurate and precise information about the environment, but at the opposite to obtain an ambient feeling that captures invariant structure of the environment on sufficiently wide areas. Ambient here means invariant relatively to specific poses, invariant relatively to particular trajectories and possibly invariant to local sensing sequences. To illustrate, one may think of sauntering in a town and being successively under the influence of a narrow street ambiance, then an open place ambiance, and so on. This kind of property is desirable in a mapping system, it should helps stable disambiguation of sufficiently wide areas. This can be stated as follow:

- Percept that are not instantaneous, integrated [width=.3]fig3.eps

Figure 2: Occupancy grid obtained in a part of our office environment (here a large corridor and a narrow lane - env. 20 m long). Grey, black and white cells correspond respectivelly to unknown, occupied and free space.

[width=.2]figloc.eps

Figure 3: Local occupancy grid (using 16 sonars and odometer, no position correction) - size 6x3 meters

[width=.4]fig5.eps

Figure 4: The percept, obtained from the local occupancy grid matured during a short period, is a vector of n distances measured from the current mass center of the grid to the nearest obstacles

[width=.4]fig6.eps

Figure 5: Successive percepts obtained along a path in the office environment (real robot experiment). They are rotated according to the internal compass

over a period of time, containing informations not accessible to immediate senses (out of reach).

- Percepts that are ambient, as much as possible invariant to particular poses and that do not rely on detection of details.

And to fulfill the requirements of the SOM learning :

- Synthetic percepts, specially to shorten SOM processing time and broadcast cost.
- Comparable percepts because SOM requires a distance between input vectors. Also we want to discriminate if two percepts are sufficiently different .

4.1 Percepts obtained from an occupancy grid

Sonar sensors are widely used in the robotic systems and their disadvantages are also well known, they are subject to noise and specular reflections. Moreover they give a poorly informed and imprecise vision of obstacles surrounding a robot (in our case 16 sonar ranges around the robot), and this at a rather low rate (1/3 of a second). H.P. Moravec and A. Elfes [9, 10] have proposed a statistic method to improve sensors information, and eventually to fuse heterogenous sensors. The occupancy grid have been succesfully used in lot of works [11, 12]. In the occupancy grids, the space around the robot is tessellated into small square cells (env. 15x15 cm). A probability value associated to each cell represents the occupancy probability of the corresponding location. Initially the probabilities are unknown (0.5), as the robot progresses the probabilities are refined (using Bayes' rule), the cells receiving more sonar hits see their value increased. The moves of the robot in the grid reflect the moves of the robot according to the internal odometer. The figure 2 shows an occupancy grid obtained in our office environment. The sonars have been modeled by single rays. Here the odometry drift and slippage have been partially compensated by using a mechanism based on wall orientation detection. However the odometry local precision is sufficient to build local grids (several meter wide) without any position correction, thus without presupposition on the environment conformation (Fig.3).

A percept with desired properties (Fig.4) can be obtained from this local grid: The mass center of the grid constitutes a virtual point of view, independent of the trajectory. Drawing n rays (n=36) from this point, we can build a vector of n distances to the nearest obstacles. The orientation of the first ray is given by the north of the filtered magnetic compass. Such vectors can be compared easily using an Euclidean distance. This percept can be input in the SOM classifier. Percept with similar shape and orientation have similar resulting classes.

4.2 Example of Percept classification

The figure 6 page 4 shows the class reference vectors obtained by a single robot during exploration of a part of the environment. This gives a short vocabulary of 6x10 classes. The classes around class (5,4) correspond to narrow North-West/South-East corridors with opening at NE. Around (5,6) are larger corridors with larger opening. Classes around (2,4) corresponds to transition between large and small corridor. Corridors with different orientation have classes around (6,10). Wide areas are around (1,8). This classification is solely intended as an example of the proposed approach and wider experiments are under way.

[width=.8]figmap.eps

Figure 6: SOM reference vectors obtained during exploration of a part of our office by a single robot.

5 Usage of the Percept classes

5.1 Support for traces and maps

During its traveling the robot can produce single traces having the following conformation: a trace is a succession of linked nodes; each node is labelled with the corresponding percept class, each link is labelled with geometric distance and compass sense orientation. A new node can be added to the trace when sufficient distance has been traveled through and when the distance between successive classes exceeds a given threshold (exploiting the topology preservation property of SOM classifier).
The traces can be combined into complex topological maps. The difficult problem of maintaining a consistent map during long term operation, as well as localization/re-localization problems can be eased by using ambient percept and percept class comparison. In a populated environment such as ours, invalid percepts (ie: caused by a group of people standing near the robot) can be rejected using quantification error.

5.2 Fusion and exchange- Percept Distribution Grids

The fact that a representation can use classes significant for all the robots helps the design of various exchange schemes. Primarily robots can exchange useful pathways that they have discovered, for instance exploration pathways starting at the power stations.

Topological representations are rather difficult to match and are not easily tractable. Alternatively we propose that the robots exchange discretized representations of the environment. Each robot can project its traces, and topological maps into a tessellated representation of the environment, a grid made of large cells (ie: 2x2 meter wide) and where each cell can store the probability distribution (histogram) of the encountered percept classes (fig 7). The cells can be labelled with utility informations (ie : visited or not). Arbitrarily the fixed point constituted by the power station is set in the center of the world. The histogram can reflect variations in the environment such as doors opened, closed , partially opened... The possibility of a grid description of the world with large cells directly rely on the ability to distinguish, stable ambient percepts in the environment, each cell being under the influence of a small set of percepts. Of course the reliability of such grids depends on correctness of the topological maps. However even imprecise they can provide a basis for cooperative behaviors.

[width=.9]fig9.eps

Figure 7: Percept Distribution Grids: The environment (here hypothetical) space is tessellated into 2x2m cells. The grid keeps memory of the percept classes that have been observed in each cell. For a given location, several percepts may be observed (door close, door semi-opened ..), therefore each cell keeps the histogram of observed classes.

The robots can exchange/fuse their maps cyclically to obtain a collective global map. The detection of overlapping zones is eased by the small amount of grid cells and by the fact that partial maps follow the same orientation (compass sense), and have a common fixed point (power station) therefore the size of the search space is reduced. A territorial division behavior, useful in most office environment applications, can be obtained by minimizing the overlapping areas between individual maps.

The grid provides the robot with a cooperation medium where synchronization, resources sharing, etc.. can take place more easily than in the physical world. Robots encounters detection, difficult to realize in the physical world, can be made easier by traces and grids confrontation. By recent traces comparison robots can confirm if they have encounter a colleague.

5.3 Meeting human models and representations

At the end, in most of multi-robot applications, it is necessary to meet the models and representations commonly used by humans. This is required for monitoring purpose or control purpose, for the design of behaviors that are somehow related to human activities. A simplistic but realistic approach can be to label 'a posteriori' the set of SOM classes. A user can visualize the percept reference vectors, and associate a symbolic description, such as 'large corridor' , 'narrow lane', 'opened door at south ' etc... The labelling can also applied to the grids, by specifying interesting points.

6 Conclusions and future work

In this paper we have presented a method for grounding a robot team's representations and communications in real perceptions. We have argued that this is an essential feature to deploy a robot group in real environments. A technical framework has been described, relying on broadcast and classification of adequate perceptions. Preliminary results have been presented to illustrate the approach. In future work we plan to perform larger experimentations with a six robots team, studying particularly : the classification process, the relevance of discovered classes, the building of reliable maps with those classes and the confrontation of individuals' maps. Also we want to extend the ambient percept with vision informations (ie: ambient color histogram).

We believe that the classification of the population's percepts provides a material for various learning experiments. Implicitly, the sharing of percepts in a robot population is an attempt to increase the initial amount of common experience, suggesting that the possibility of any communication act depends on such common experience, a 'common ground' as we may say.

References

[1] S.Picault, A. Drogoul. (2000) Towards Situated Robots Collectivities in a Human Environment : the MICRobES Project. submitted to ECAI.

[2] R. A. Brooks. (1991) Intelligence Without Reason, MIT AI Lab Memo N0 1293.

[3] A. Billard, K. Dautenhahn. (1997) Grounding communication in situated, social robots, proceedings 1997 of Towards Intelligent Mobile Robots TIMR.

[4] D. Jung, A. Zelinsky. (2000) Grounded Symbolic Communication between Heterogeneous Cooperating Robots, Autonomous Robots Journal, special issue on Heterogeneous Multi-robot Systems, Kluwer Acadamic Publishers, Balch, Tucker, Parker Lynne (eds), Vol 8, No 3.

[5] A. Kurz. (1996) Constructing Maps for Robots Navigation Based on Ultrasonic Range Data, Transactions on Systems, Man and Cybernetics, Vol 26, No 2, 233-242.

[6] U. Nehmzow, T. Smithers. (1991) Mapbuilding using Self-organizing Networks, From Animals to Animats, J.A. Meyer and S. Wilson (eds.), MIT Press, 1991, 152-159.

[7] T. Kohonen. (1997) Self-Organizing Maps. Springer-Verlag, New York.

[8] D. Hubel. (1995) Eye, Brain and Vision. Scientific American Library No 22, W.H. Freeman, New York.

[9] H .P. Moravec. (1988) Sensor Fusion in Certainty Grids for Mobile Robots, AI Magazine, Summer 1988 ,61-74.

[10] A. Elfes. (1989) Using Occupancy Grids for Mobile Robot Perception and Navigation. IEEE Computer, june 1989 , 46-57.

[11] B. Yamauchi, P. Langley. (1997) Place Recognition in Dynamics Environnements, Journal of Robotic Systems, Vol 14 , No 2 , 107-120.

[12] S.Thrun, A. Bucken, W. Burgard, D. Fox, T. Frohlinghaus, D. Hennig, T. Hofmann, M. Krell, T. Schmidt. (1997) Map Learning and High-Speed Navigation in RHINO, AI-based Mobile Robots : Case studies of successful robot systems. D. Kortenkamp, R.P. Bonnasso, R.R Murphy (eds) MIT Press.