



HAL
open science

Modeling Evolving Behaviors in Ant Colonies

Ugo Galassi, Guénaél Cabanes, Dominique Fresneau

► **To cite this version:**

Ugo Galassi, Guénaél Cabanes, Dominique Fresneau. Modeling Evolving Behaviors in Ant Colonies. International Journal of Intelligent Systems, 2009, 18, pp.186 - 195. 10.1515/JISYS.2009.18.4.353 . hal-01461403

HAL Id: hal-01461403

<https://hal.science/hal-01461403>

Submitted on 8 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modeling Evolving Behaviors in Ant Colonies

Ugo Galassi,¹ Gueanel Cabanes² and Dominique Fresneau³

¹*Dipartimento di Informatica, Università Amedeo Avogadro, Via Bellini 25G, Alessandria, Italy;* ²*LIPN-CNRS UMR 7030 and* ³*LEEC, Université Paris XIII, Villetaneuse, Paris, France*

ABSTRACT

Modeling societies of individuals is a challenging task increasingly attracting the interest of the Machine Learning community. In ant colonies, individuals have the same physical attitudes and, in principle, can assume any role the social environment requires. For the biologist understanding the complex dynamics ruling a colony is hard due to the difficulty of collecting and classifying long term ant activities. Here we present a novel approach for analyzing activity logs from an ant colony using an RFID environment. A semi-automated algorithm for segmenting traces and discovering the role played by each individual during the observation phase is described. A Structured Hidden Markov Model was used to build the model of single individual activities. Then, the global profile of the colony was traced during the emigration from one nest to another. The method provided significant information concerning the social dynamics of ant colonies.

KEYWORDS: S-HMM, RFID, ant, emigration, activity learning

1. INTRODUCTION

This paper addresses the problem of analyzing the behavior of an ant colony to discover their social rules and social roles, which are still not yet well identified and understood. *Machine Learning* can offer a great help to the biologist in discovering the complex interactions between individuals in an ant colony. The large number of subjects that must be examined and the complexity of the environment make this

Reprint requests to: Ugo Galassi, Dipartimento di Informatica, Università Amedeo Avogadro, Via Bellini 25G, Alessandria, Italy; galassi@mfn.unipmn.it

task infeasible without the help of some (semi)automated framework for collecting and analyzing data.

In this specific case, an ant colony was observed after creating an artificial climatic mutation, which caused the emigration from one nest to another. Colony migrations are dangerous events because each colony member—and particularly the queen and brood—becomes vulnerable to predation or runs the risk of getting lost. Therefore, the relative timing of queen emigration and brood transport are strategic issues for the colony as they may significantly lower the risks incurred by these essential members.

The queen moving and the brood transport do not occur at random during nest relocation but rather in the middle of the sequence of all emigration events. In addition, the different brood categories are transported in a specific sequence (i.e. cocoons first, larvae second, and eggs last) (Pezon et al. 2005). The results correspond to robust colony-level strategies since they are related neither to emigration distance nor to colony or brood size.

All individuals have the same physical attitudes and, in principle, can assume any role the social environment requires. Facing the emergency, the different elements of the colony change their current social activity and assume specific roles to accomplish the emigration in the new nest, depending to the necessities of the colony. Afterwards, they return to normal activities, not necessarily the same they were accomplishing before the emergency.

The collective and individual decision rules followed by workers as well as the specific mechanisms underlying this colony-level phenomenon are still poorly understood so it is essential to find the rules that govern ant individual behavior and its integration at the colony scale.

We monitored individual activities using passive radio-frequency identification (RFID) technology, a novel procedure as applied to ants. RFID tagging has no observable effects on the behavior of the ants (Robinson et al. 2009) and allows tracking the activity patterns of large numbers of individuals over a long period.

The fundamental issue investigated here is the construction of the model of the activities that an individual can undertake, such as: nursing, transporting, foraging, and so on, according to the emerging needs of the colony. Activity models are used to segment ant traces obtained by RFID logs for determining the different roles played by each ant during the emigration act. Starting from those segmentations, global parameters characterizing the global colony behavior can be inferred.

The activity models are based on Structured Hidden Markov Models (S-HMM), a variant of classical HMM (Rabiner, 1989), well suited to combine a priori structural information from the domain experts with statistical information inferred from data. The paper describes both the methodological approach and the results obtained. Those results are innovative and offer to biologists new insights into the behavior of ants.

2. THE PROBLEM OF NEST RELOCATIONS

The ecological success of insect societies is due in large part to their ability to build long-lasting and elaborate nests, which provide a secure environment for the successful rearing of brood and reduce predation risks for all colony members (Wilson, 1971). As constructing a complex nest requires a considerable amount of resources, for long time the ethologist studying animal behavior believed social insect colonies to be tendentially rooted in the same nest. In the past few decades, however, a strikingly new picture has slowly emerged regarding the stability of ant nests. Ant colonies migrate from one nest site to another one more frequently than previously assumed (Holldobler & Wilson, 1990). Indeed, in many species, colony emigration proves to be the rule rather than the exception (Möglich, 1978, Herbers, 1985, Partridge et al. 1997), and even highly sedentary species will migrate if sufficiently disturbed due to mechanical vibration, micro-climate change, predation, competition and so on (Smallwood, 1982, Droual, 1984, Franks et al. 2002).

Emigrations represent a costly and perilous phase in the colony life cycle, especially for small colonies. Each colony member becomes exposed to predation and incurs the risk of getting lost, thus reducing the population size and colony resilience (Byrne, 1994, Bouwma et al. 2003). During emigration, an ant colony must swiftly move the brood to a better location, and the timing of brood transport is a crucial strategic issue.

Despite their importance to colony survival, the issues of risk minimization during colony movement and the different role played by each individual are still poorly understood. Only very specific individuals or events (queen or cocoon relocation) have been analyzed. This limitation is mainly due to the difficulty of collecting information about large numbers of individuals acting in such a chaotic environment like an ants' nest.

The brood is sensitive to prevailing environmental conditions and is generally helpless, requiring the constant care and protection of workers. Pezon et al. (2005) demonstrated that the median of the temporal distribution of brood transport events should occur in the middle of the emigration phase, thus ensuring that the brood is swiftly transferred to the new nest while still benefiting from maximum worker protection at all times in both the new and old nests.

The stereotyped techniques of chemical communication and recruitment by tandem running employed during nest relocation have been extensively studied in *P. Obscuricornis* (Holldobler, 1980, Holldobler, 1984), as well as other ant species belonging to the *Ponerinae*, *Myrmicinae*, and *Formicinae* (Holldobler & Wilson, 1990).

Recently, Franks and Sendova-Franks (2000), demonstrated that in monogynous ants, the queen is transported in the middle of the emigration phase, thus ensuring that the queen moves from the safe environment provided by one half the colony in the old nest to the safe environment provided by the other half in the new nest. This adaptive colony-level phenomenon guarantees that the colony's only queen benefits from maximum worker protection at all times in both sites. During emigration, functionally polygynous colonies (characterized by the simultaneous presence of many queens) should not necessarily face the same selection pressures as their strictly monogynous counterparts.

3. EXPERIMENTAL SETTING

The dynamic of task allocation through worker ontogenesis has been widely described in prior theoretical investigations. However, the lack of adequate tools for tracing ant behavior has prevented an experimental test of the theoretical hypothesis. In this paper we aim at studying the dynamic of a nest change in an ant's colony. Our interest is in characterizing and analyzing the evolution of the whole colony, understanding how ants act and partition their roles during the entire phase of nest relocation. The kinetics characterizing the sequence of departure from the original nest, passage through the foraging area (outside), and arrival in the new protected nest will be analyzed. These actions require a high coordination among workers. It is likely to expect that foragers will start the exploration of the new nest and will initiate recruitment to lead more static ants (nurses and inactive ants) toward the new

nest. Which ants will transport the brood and more importantly which ants will take care of the queen movement toward the new nest is unknown. Similarly, we do not know whether the established distinction of activities in a stable situation will condition the role a specific individual will assume during the emigration phase.

3.1 RFID For Ants Activity Tracking

RFID technology has been used for social insects' analysis for a few years. The first results of Streit et al. (2003) convinced other authors to use this technology because of its flexibility and the richness of the recorded information (Sumner et al. 2007, Molet et al. 2008). Concerning ants, the challenge is represented by the miniaturization of the RFID equipment. This task has been done successfully by Robinson et al. (2009).

In all prior experiments, RFID tags have been supported without great difficulty by the insects. Nevertheless the devices did not allow monitoring points or multiple individuals to be monitored simultaneously. The originality of the device presented here is that this tool supports many observation points and allows a large number of individuals to be traced simultaneously. In this way, we are able to follow the entire colony in real time.

3.2 Measurement Device

For this investigation, we chose a big-sized tropical ant, *Pachycondyla tarsata*, which constructs subterranean nests distributed in several interconnected chambers spread over a distance of 10 meters. Colonies of these species are typically composed of ten to a few thousand ants. The RFID tag consists of a chip attached to an antenna weighing under 40 mg (i.e., 25% of an ant weight), glued on the animal thorax (Fig. 1(a)). The movement between nests of a colony of 55 workers was monitored by the RFID device for the whole duration of colony emigration (about 4 h).

The experimental device for this experiment consists of three rooms each and a foraging area, linearly connected by six tunnels (Fig. 1(b)(c)). At the beginning of the experiment, the queen (not tagged) and its brood (eggs, larva and cocoon) are located in Room 3 of the first nest, the farthest from the foraging area. Each tunnel is equipped with two RFID readers (number 1 to 12 from Room 3 in Nest 2 to Room 3 in Nest 1) that detect the passage and the direction of tagged individuals between

rooms. The position of an individual may be inferred unambiguously by the information provided by the readers in the tunnels. The lack of detection implies that the individual is out of the tunnel and thus in one of the seven rooms. The exact location of a tag (i.e., of an individual) can be deduced from the travel direction.

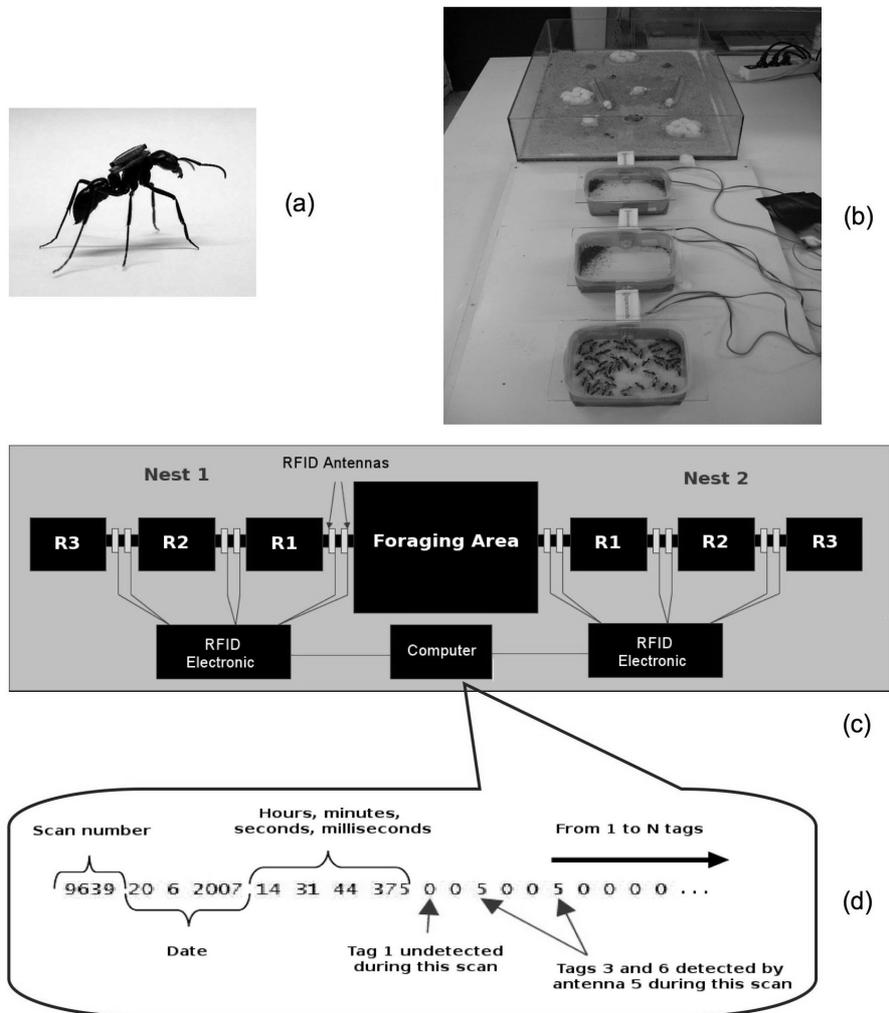


Fig. 1: The RFID environment: (a) Ant with RFID tag. (b) Snapshot of a part of the ant nest. One can appreciate the three rooms of the old nest and the foraging area. (c) The RFID experimental device. (d) Example of a recorded scan in the data file

The information recorded by readers is handled by an RFID electronic device, and then sent to a computer that creates and store the data files.

The data files are in text format and indicate for each antenna scan (about three scans per second), the progressive number of scan and the time of day, and for each individual (i.e., for each tag), which antenna is activated (Fig. 1(d)). If, during a scan, none is detected, then nothing appears in the data file. A simple treatment on these files makes it possible to obtain spatial information for each individual. Due to the limitation imposed by miniaturization, the RFID tags sometimes can be missed by the detecting apparatus. A missing detection rate ranging from 5% to 15% has been observed.

At time $t = 0$ we switch on a strong neon light (strongly repellent for ants) over the first nest and we open the entrance of the second nest, then we record the colony movement until the entire brood is moved into the second nest (4 h).

4. MODELING SOCIAL ACTIVITIES

The RFID apparatus only provides a partial observation of the individuals. No information is provided concerning what happens inside a room, but only the duration of the permanence of an ant inside it can be known. Moreover, sensors are not reliable having a missing detection rate ranging from 5% to 15%.

The goal is to reconstruct the evolution of the activity of the single individuals in the context of the social environment under the pressure of the simulated ecological mutation. More specifically, we want to discover which kinds of activities are undertaken during the emigration phase, how many individuals are dedicated to each activity, and when an individual may change activity.

Achieving this goal requires solving the following problems:

- To reconstruct the most likely paths made by ants considering that many transits in the tunnels are unobserved due to missing detections.
- To characterize the different activity patterns emerging from the data.
- To infer the activity models.
- To segment and label the paths according to the activity, which most likely produced the observed action sequence.

The major requirements that a modeling tool must satisfy are the capacity of handling partially observable status, and modeling the duration of the permanence in the different areas of the environment. To this purpose, the graphic model approach (Murphy, 2002) is the most promising.

The probability theory offers a framework for modeling the evolution of processes characterized by inherent randomness or operating in environments too complex for a precise analysis. The core idea is that the statistical distributions governing the evolution of a system can be estimated from a learning set of traces describing its past history. This approach allows us to infer a model from a relatively small set of strings. In particular, two tools emerge as candidates for the specific task of segmenting and labeling sequences in presence of hidden states: Hidden Markov Model (HMM) (Rabiner, 1989) and Conditional Random Fields (CRF) (Lafferty et al. 2001). Recent findings (Murphy, 2002) favor CRF, which in many cases outperformed HMM. Nevertheless, the requirement of modeling duration suggested that we adopt the HMM approach. Indeed, directed graphs are a natural choice when modeling successions of events that are characterized by some temporal causality is needed. Undirected models like CRF are appropriate in modeling those data in which such directionality does not exist. Well-assessed methods for extending HMMs to cope with durations are available, whereas CRFs have been little investigated in this sense (Natarajan & Nevatia, 2008).

Another reason that make the HMMs a natural choice is that they can be easily learned, also in real-time application. In fact the well-known forward-backward and Viterbi algorithms for HMMs (Rabiner, 1989) have direct analogues for CRFs, with the same asymptotic running times. But the training step, which determines a weight for each feature function, is somewhat more complex; generally, no closed-form solution exists for the optimal assignment of weights, so it must be found using numerical optimization techniques. The adopted tool is then a HMM variant called S-HMM (Galassi et al. 2007a), which offers specific features for modeling permanence inside rooms. In the following section, a brief explanation of the S-HMMs and their properties will be presented.

4.1 The Structured Hidden Markov Model

An S-HMM is a directed graph built up according to precise composition rules, with several independent sub-graphs (blocks) structured according to the paradigm

used in Object Oriented Programming. A block consists of a set of states, only two (the *initial* state I and the *end* state E) are allowed to be connected to other blocks. The blocks can be nested inside each other.

Two kinds of blocks are possible: *basic blocks* and *composite blocks*. The states of basic blocks produce observable emissions according to the classical HMM paradigm. The states of composite blocks correspond to basic or composite block defined at a lower abstraction level. When a transition into a state q occurs in a composite block, a call is made to the lower level block associated to q and the activity is suspended until the call returns.

The basic assumption underlying an S-HMM is that a sequence of observations $O = o_1, o_2, \dots, o_T$ could be segmented into a set of subsequences O_1, O_2, \dots, O_N , each generated by a sub-process with only weak interactions with its neighbors (Bouchaffra & Tan, 2006). This assumption is realistic in many practical applications, such as, for instance, speech recognition (Rabiner, 1989, Rabiner & Juang, 1993), and DNA analysis (Durbin et al. 1998).

A major feature of S-HMM is that it could be locally trained using the classical Baum-Welch algorithm, considering only a subset of the sub-models occurring in the compound one (Galassi et al. 2007b). A nice consequence of this property is that an S-HMM can be constructed and trained incrementally, by adding new sub-models or revising existing ones as new information comes in. Newly added sub-model could be produced by an independent learning process or may be provided by an expert as a-priori knowledge. Different models could be learned and successively re-trained independently one from the other, providing a natural sub-problems' decomposition.

Another major feature of S-HMM is that it allows different kinds of blocks to be defined for modeling specific tasks. This feature is fundamental for building models describing complex temporal dynamics like the permanence of ants inside different kind of rooms. As demonstrated in a recent work of Abou-Moustafa et al. (Abou-Moustafa et al. 2004), the use of an ad-hoc topology strongly increases HMM performances. S-HMM permits explicitly differentiating the models adopted for the various kind of rooms present in the nest.

All basic algorithms for computing probability distributions and estimating model parameters from sequences, such as forward-backward and Viterbi algorithm (Rabiner, 1989), immediately extend to S-HMM. A detailed description of S-HMM is provided in (Galassi et al. 2007a, Galassi, 2008).

4.2 Modelling Duration

The problem of modeling durations in the HMM framework has been principally faced in Signal processing. Two approaches to the problem emerge from the literature. The first produced an extend modeling tool called *Hidden Semi-Markov Model* (HSMM), which corresponds to HMM augmented with probability distributions over the state permanence (Levinson, 1986, Pytkkonen & Kurimo, 2004, Tweed et al. 2005). The alternative approach is the so called *Expanded HMM* (Josep, 1996). Each state, where it is required to model duration, is expanded into a network of states, properly interconnected. In this way, the duration of the permanence in the original state is modeled by a sequence of transitions through the new state network where the observation remains constant. The advantage of this method is that the Markovian nature of the HMM is preserved. Nevertheless, the complexity increases according to the number of new states generated by expansion.

In the framework of S-HMM, the approach of Expanded HMM is used and specific basic blocks are provided to model the probability distribution of the permanence inside a *macro state*. The specific HMM topology we adopted for the present application is reported in Figure 2(e). Basically, this model exhibits an Erlang's distribution, when the Forward-Backward algorithm is used to compute the probability distribution. Basically, an ant activity model assigns a probability distribution over the set of all possible paths an ant accomplishing a specific activity can go through the artificial environment. Let s be a sequence of observations. By comparing the different probability assigned to s by a set of different activity models, we can infer the activity that most likely generated s .

4.3 Ant Activity Model Architecture

The observation of a path is a sequence of pairs $\langle t_i, s_i \rangle$ collected from the RFID sensors, being t_i the time of the detection in msec, and s_i the id of the sensor. In the S-HMM framework discrete time is assumed. Then, a transformation from the numeric representation from the sensors to a discrete (symbolic) representation has been defined, which preserves the accuracy implicit in the original coding. The symbolic sequences are encoded using an alphabet $A = \{A, B, C, D, E, F, G, H, J, K, L, .\}$, where letters from A to L correspond to the RFID detectors from 1 to 12, respectively, and “.” denotes a time interval, in which no observation from the

sensorial apparatus is received. The transformation, from a numeric to a symbolic sequence is obtained by subdividing the time into discrete intervals of one *second*. A numerical sequence is scanned from the beginning to the end moving ahead of one interval at a time. If a signal from a detector is found, the corresponding symbol in the symbolic sequence is appended; otherwise a “.” is inserted. After the translation, the permanence in a room, is be represented as a string of “.”. Moreover, undetected transit in a tunnel will be report as a “.”, as well.

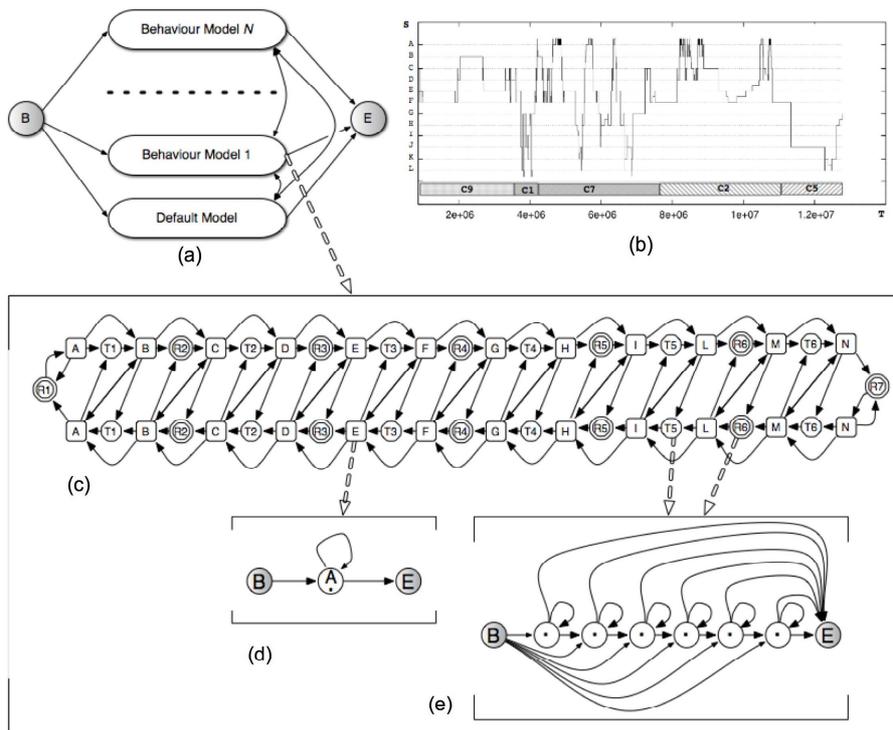


Fig. 2: Structured HMM used for modeling ant behavior. (a) The activity tagger, i.e. a three-level S-HMM used for modeling colony behavior. (b) Example of labeled sequence obtained using the activity tagger. (c) Second level model; For each behavior, an independent model has been learned. (b) basic block encoding the passage under a sensor; (c) basic block modeling the duration of the permanence in a room or a tunnel.

After experimenting with different model architectures, the one reported in Figure 2(c) has been chosen. It is a two level S-HMM, where the upper level models the long range path through the environment, while the lower level models the observations detected by the RFID sensors, and the duration of the permanence inside rooms and tunnels. States in the upper level defines a double chain sharing the ends. The states denoted with single circles represent the permanence in a tunnel, whereas those denoted with a double circle represent the permanence in the rooms of the nests or in the foraging area. States denoted with squares represent sensors. Each state is associated to a block at lower level, which models the probability distribution for the permanence in the associated location or the process of generating the observable emissions of the sensors. Referring to Figure 2(c), the upper chain models the action of going from the old nest to the new nest, while the other models the action of going from the new nest to the old nest. Changes of directions cause switching from one chain to the other.

5. LEARNING FROM THE ACTIVITY TRACES

From the model architecture described in Figure 2(c), the activity models have been estimated to construct an *activity tagger*. This feature is used to infer the most likely path of an ant and to segment and label it according to the activity, which most likely generated the signals reported by the RFID detectors. From the labeled paths, the global behavior of the colony along the emigration phase has been reconstructed.

5.1 Activity Tagger Architecture

An activity tagger is a three-level S-HMM, obtained by layering a new block on top of the activity models. As show in Figure 2(a), this new layer defines a fully connected graph among the blocks modeling the different activities. Then the activity tagger interprets the ant traces as a sequence of segments, each corresponding to a different activity phase. By exploiting the S-HMM compositional properties the activity tagger can be refined by training single blocks independently, as well as the entire structure using the classical Baum-Welch algorithm.

The standard method for interpreting a sequence using an HMM (S-HMM) makes use of the Viterbi algorithm (Forney, 1973, Rabiner, 1989) to find the maximum likelihood path in the model state space, which corresponds to the observed events. In our case, this method does not work because the HMM modeling the duration of the permanence in the rooms and in the tunnels requires forward-backward algorithm.

Then we adopted an alternative method, also described in (Rabiner, 1989), which consists in finding, at each time t , the maximum likelihood state q_t of the model, as defined by the following equation:

$$q_t = \operatorname{argmax}_i (\alpha_t(i) \beta_t(i)) \quad 1 \leq i \leq N \quad (1)$$

In expression (1) $\alpha_t(i)$ is the classical function that estimates the probability for model λ of being in state q_i after generating (in all possible ways) the sequence of observations from t_0 to time t . Symmetrically $\beta_t(i)$ is the probability for λ of generating the remaining part of the sequence from t to T starting from status q_i .

5.2 Learning Procedure

The complete learning procedure for learning the activity tagger integrates data-mining algorithms with the manual action of an expert of the domain. The domain expert is very good in detecting where the activity pattern changes and in providing an episodic interpretation of fragments of the paths, but it performs poorly in tasks requiring the systematic analysis of a large amount of data. On the other hand, the learning algorithm is very good in discovering regularities and similarities among different episodes discovered by the expert. From this cooperation, the groups of characteristic activities are progressively individuated and modeled. The procedure consists of the following steps, which are repeated until the convergence to stable models is achieved:

Let L be the set of sequences to be labeled. Let, moreover L_i a subset of L , used for iteration i .

1. label sequences using the current tagger version;
2. refine the assigned labels with the help of an expert of the domain;
3. segment every sequence according to the assigned label;
4. cluster segments according to the assigned label;
5. from every cluster C_k estimate a model λ_k ;

6. construct a new tagger using the models learned in the previous step, and optionally train it using the Baum-Welch algorithm;
7. add new sequences extracted from L to L_i obtaining a new learning set L_{i+1} .

Finally, after the procedure iteration stops, all sequences in L are labeled using the tagger constructed at the last step.

Notice that, the first time the procedure is executed, no tagger exists. In this case, the first step has been carried using an algorithm based on Kohonen maps (Cabanes et al. 2008, Cabanes & Bennani, 2008), which was able at providing a rough segmentation. Then the domain expert manually corrected the output of the program. As this task is time consuming, we started with a small learning set extracted from the a set L containing 57 sequences (one for each individual of the colony). The procedure has been iterated three times incrementing the learning set up to 40 sequences. Afterwards, all 57 sequences have been labeled using the final tagger. An example of labeled sequence is reported in Figure 2(b).

5.3 Tracing The Colony Profile

From the labeled sequences three groups of global parameters have been extracted tracing the profile of the ant colony during the emigration phase:

- The temporal evolution of the number N_j of individuals present in every room R_j of the old and of the new nest.
- The temporal evolution of the number N_{ij} of individuals involved in the activity characterized by model λ_i , in every room R_j .
- The global number $N_i = \sum_{R_j} n_{ij}$ of the individual involved in each activity λ_i

6. EXTRACTED KNOWLEDGE ANALYSIS

At the end of learning procedure, eight groups of behaviors $A_i (1 \leq i \leq 8)$ emerged. A last group (A_0) has been defined, which corresponds to activity segments where a clear pattern is not detectable. The activity logs of each ant have been labeled according to those behaviors. Starting from the labeled sequences, the global

parameters have been extracted tracing the profile of the ant colony during the emigration phase.

The first information extracted from the sensor logs, using the activity tagger, is the path made by every ant during the emigration phase. From it, the global parameters N_j ($1 \leq j \leq 8$) reporting the temporal evolution of the number of individuals in the different rooms, have been computed. The results are described by the diagram of Figure 3(a). Here every room is associated to a different color. Then, the area covered by a color is proportional to the number of ants in the corresponding room. Three facts emerge,

1. In the beginning, a small number of ants are also present in rooms R_1' , R_2' and R_3' of the new nest, meaning that the colony started to explore the new nest as soon as the door was opened. Anyhow, this does not produce remarkable effects until the neon light begins to trouble the colony.
2. After the emigration phase was concluded, no more ants were present in the internal rooms R_2 and R_3 of the old nest, but the presence of ants in room R_1 was remarkably higher than in the corresponding room R_1' of the new nest before the emigration. This can be explained considering the combined effect of the residual pheromone, which acts as an attractor for the ants in the old nest, and the neon light, which acts as a repellent.
3. The final repartition of individuals on rooms R_2' and R_1' is quite different than the one in the corresponding rooms of the old nest before the emigration phase. More specifically, the percentage of individuals in room R_2' is much higher. The explanation is that ants were busy with cocoons located in room R_2' . The distribution shows tendency to returning to the normal values at the end of the period of observation.

Figure 3(b) shows the evolution of parameters N_i corresponding to the number of individuals involved in activity A_0, \dots, A_8 . Figures 4, 5, 6 present the evolution of the number n_{ij} of individuals involved in the activity characterized by model λ_i , in every room R_j . For sake of clarity, in the figures we grouped together the rooms belonging to the same nest.

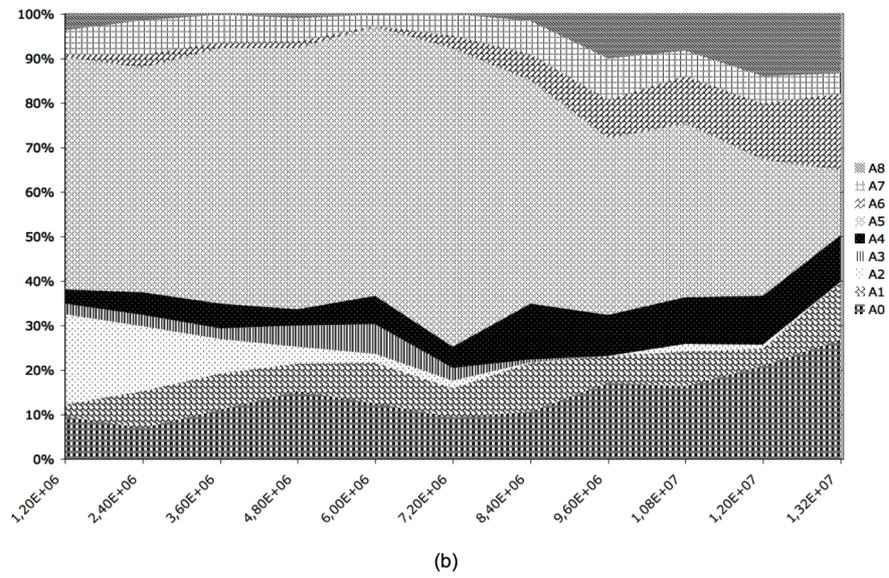
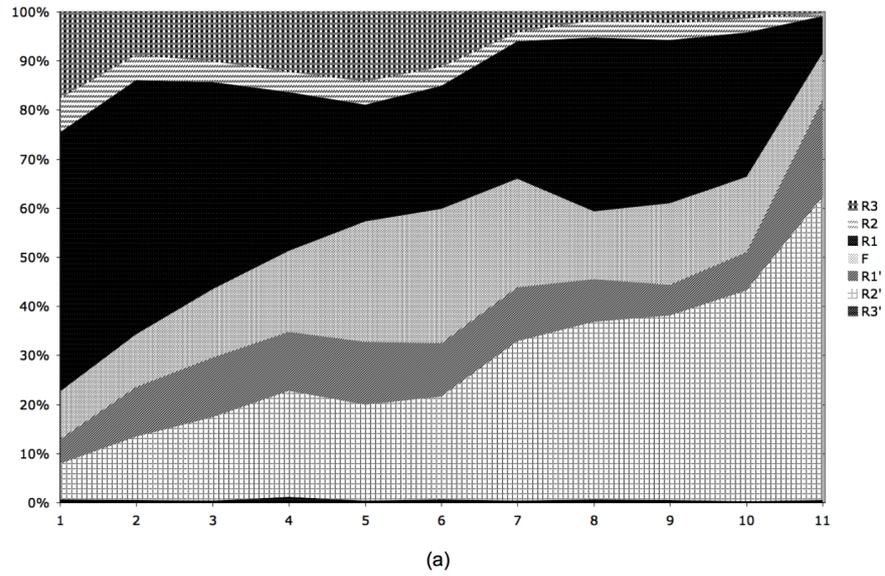


Fig. 3: Activity analysis: (a) Distribution of the individuals on the different rooms of the environment during the emigration phase. (b) Evolution of the number of individual involved in the different activities.

Those results have been used by ethologists for hypothesizing a plausible explanation for each behavior.

- A0 Unstructured patterns.** The ants are not participating into any activity and their path through the rooms is not structured. An ant can stay for a long time in a room, then change rooms, depending on random events: social interaction (sometime aggressive), noise in the environment, fear, thirst, hunger, etc...
- A1 Quick exploration of the new nest.** The ants have just discovered the new nest and start a fast and active exploration of the new site. They move quickly through all rooms for evaluating the safety of the new site and decide quickly if the emigration should start to this new nest.
- A2 Panic movements.** This behavior is expressed mostly by nurses: due to the new environmental conditions —strong light and increased temperature— the old nest is no longer suitable for a good brood care. The ants then start moving quickly inside the old nest to find a room with appropriate conditions. As such rooms no longer exist in the old nest, such ants are then ready to emigrate.
- A3 Panic movements in the old nest.** These movements are very similar to A2, but the ants are more fearful (this behavior seems to be expressed mainly by the youngest ants). With respect to behavior A2, the ants remain closer to the brood and the queen and never go outside.
- A4 General patrolling.** Patrolling is a general exploration of the environment. This behavior can be expressed by some foragers that are not really involved in the emigration activity, or it may be the beginning of a protection behavior, as has been observed in some other species. In this case, the ants move through the two nests and the foraging area to identify potential problems (predator, dangerous area, ...).
- A5 Transportation.** The ants are transporting something, i.e. the queen, a cocoon, a larva, or an egg. This behavior is characterized by many very regular round-trips between the two nests. This is a very characteristic pattern. Sometimes it can be expressed by ants that actually don't transport anything but rather act as a transporter (follow a transporter or move until finding something to transport).
- A6 Accommodation/preparation of the new nest.** The new nest is now known to be safe (it has no light, no disturbance, and most ants are installed in it,

then the colony is well structured again), but some work is needed to prepare the nest for an optimal brood care (find some water outside to increase the humidity inside the new nest in some room, find some construction materials, find some food, etc...).

A7 Patrolling in the old nest. This behavior is a defensive patrolling inside the old nest in reaction to the disturbance. Ants act for defending and protecting the queen and the brood until their relocation.

A8 Patrolling in the new nest and foraging area. This behavior is a defensive patrolling inside the new environment. The purpose is probably to defend and protect the queen and the brood while moving and installing inside the new nest.

From an ethological point of view, the previous results are of great help for understanding how tasks are distributed during a nest relocation. Indeed, we obtained a very accurate description of the dynamic of the whole colony during all the emigration phase allowing us to posit a strong hypothesis about the function of the different behavior during the nest relocation phase. Some results are in accordance with previous works, especially the behaviors that can be observed in the foraging area. For example, the dynamic of the transportation behavior detected by the system match the results presented in (Pezon et al. 2005). These hypotheses should now be validated by repeating the experiment with different colonies and different species. A complete understanding of the emigration process based on systematic experimentations would be an important step ahead for the research in social insects.

7. FINAL REMARKS

We have presented a novel application of graphical model methods to the interpretation of data collected from an ant colony. Very encouraging results were obtained using S-HMM, a variant of HMM, well suited for modeling duration. We obtained an exhaustive description of the dynamic of the whole colony during all the emigration phase, allowing the ethologist to better understand the role played by each individual. Noteworthy is that the presented application is very innovative with respect to the current state of the art. Indeed, even if several authors have addressed the problem of modeling insect colonies, then this has been done with different

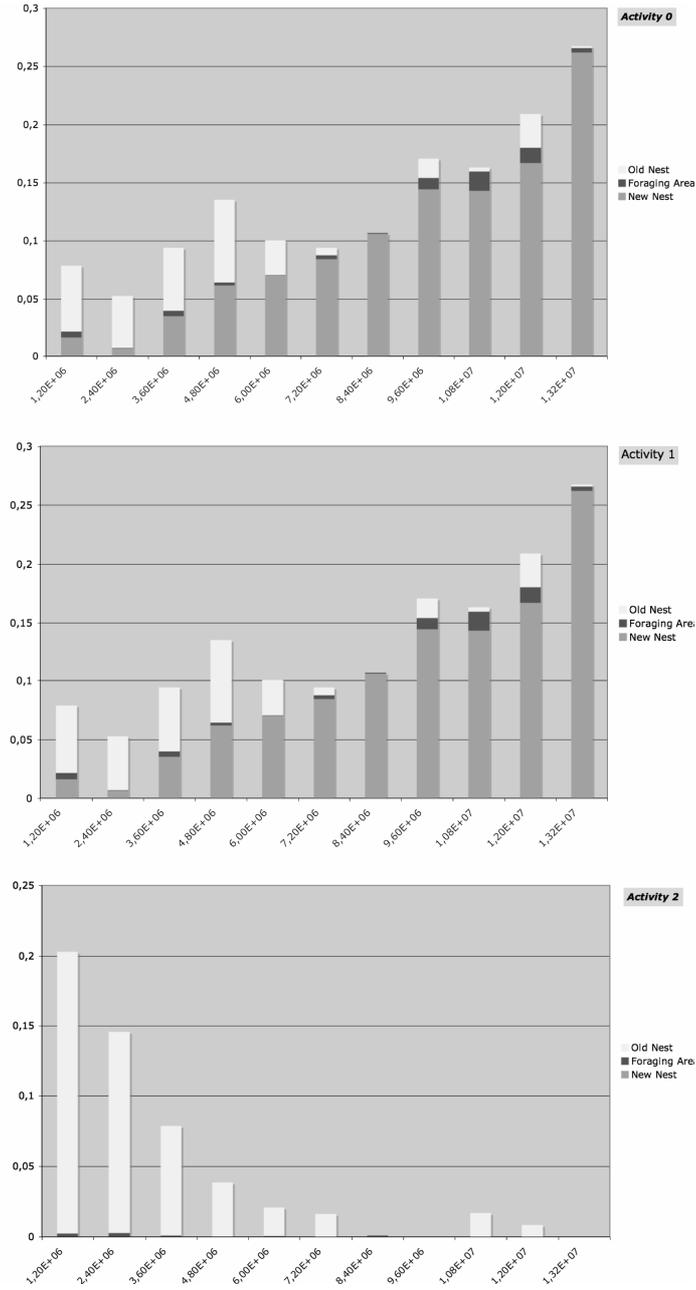


Fig. 4: Evolution of the number of individuals involved in the activities $A_0 \dots A_2$. Each histogram bin is proportional to the number of individuals acting such activity in a 20 minutes window. The bins are subdivided according to the number of individuals that are present in the old nest, the foraging area or the new nest.

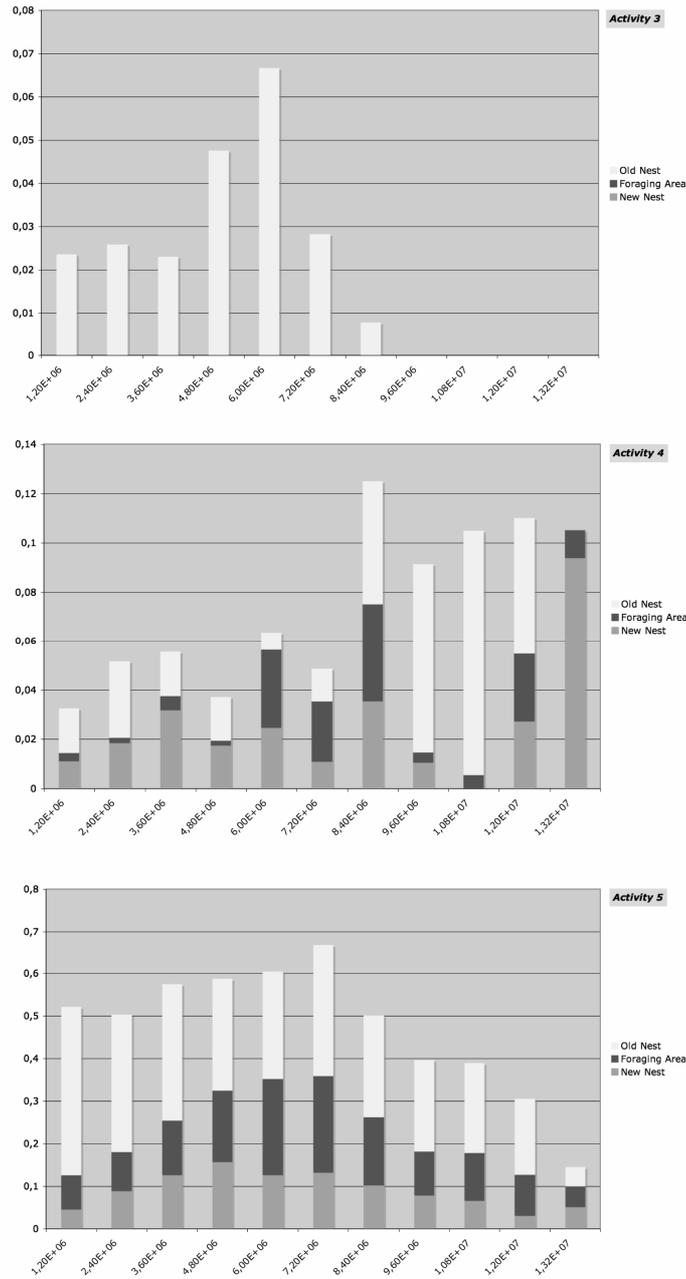


Fig. 5: Activity analysis: evolution of the number of individuals involved in the activities $A_3 \dots A_5$. Each histogram bin is proportional to the number of individuals acting such activity in a 20 minutes window. The bins are subdivided according to the number of individuals that are present in the old nest, the foraging area or the new nest.

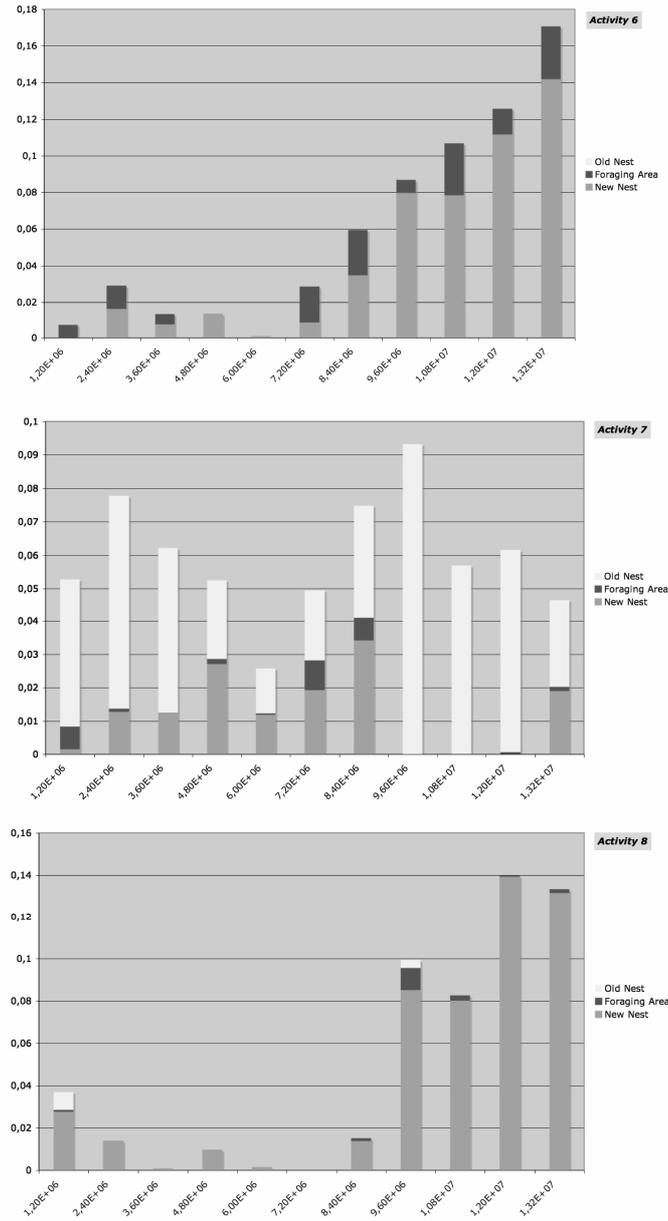


Fig. 6: Activity analysis: evolution of the number of individuals involved in the activities $A_6 \dots A_8$. Each histogram bin is proportional to the number of individuals acting such activity in a 20 minutes window. The bins are subdivided according to the number of individuals that are present in the old nest, the foraging area or the new nest.

goals. In general, the interest has been to study the emerging behavior of the colony, seen as a complex system, starting from a simple model of the individual. In our case, we started from the opposite point of view, proposing a method for observing and modeling the medium/long term behavior of real individuals, as induced by the conditioning of the social environment, in an emergency condition. Finally, RFID methods for tracking the positions of people or animals begin to be quite a diffused practice, which is attracting the interest of the data mining community. Nevertheless, a method for constructing a complex model of the behavior of an individual traced by an RFID has not been proposed until now.

REFERENCES

- Abou-Moustafa, K.T., Cheriet, M. and Suen, C.Y. 2004. On the structure of hidden Markov models, *Pattern Recognition Letters*, **25**(8), 923-31.
- Bouchaffra, D. and Tan, J. 2006. Structural hidden Markov Models using a relation of equivalence: application to automotive designs, *Data Mining and Knowledge Discovery*, **12**, 79-96.
- Bouwma, A.M., Bouwma, P.E., Nordheim, E.V. and Jeanne, R.L. 2003. Founding swarms in a tropical social wasp: adult mortality, emigration distance, and swarm size. *Journal of Insect Behavior*, **16**, 439-52.
- Byrne, M.M. 1994. Ecology of twig-dwelling ants in a wet lowland tropical forest. *Biotropica*, **26**, 61-72.
- Cabanes, G., Bennani, Y., Chartagnat, C. and Fresneau, D. 2008. Topographic connectionist unsupervised learning for RFID behavior data mining, in *Proceedings of the 2nd International Workshop on RFID Technology—Concepts, Applications, Challenges, IWRT 2008*, edited by Sheng, Q.Z., Maamar, Z., Zeadally, S. and Cameron, M., ICEIS 2008, Barcelona, Spain, June 2008; 63-72.
- Cabanes, G. and Bennani, Y. 2008. A local density-based simultaneous two-level algorithm for topographic clustering, in *Proceedings of the IEEE World Congress on Computational Intelligence*, 1176-82.
- Droual, R. 1984. Anti-predator behavior in the ant *Pheidole desertorum*: the importance of multiple nests. *Animal Behavior*, **32**, 1054-8.
- Durbin, R., Eddy, S., Krogh, A. and Mitchison, G. 1998. *Biological sequence analysis*, Cambridge University Press.
- Forney, G.D. 1973. The Viterbi algorithm. *Proceedings of IEEE*, **61**, 268-78.
- Franks, N.R. and Sendova-Franks, A.B. 2000. Queen transport during ant colony emigration: a group-level adaptive behavior. *Behavioral Ecology*, **11**, 315-8.
- Franks, N.R., Pratt, S.C., Mallon, E.B., Britton, N.F. and Sumpter, D.J.T. 2002. Information flow, opinion polling and collective intelligence in house-hunting social insects. *Philosophical Transactions of the Royal Society of London (B)*, **357**, 1567-83.

- Galassi, U. 2008. *Structured Hidden Markov Models: A General Tool for Modeling Process Behavior*. Ph.D. thesis, Università degli Studi di Torino, Dottorato di ricerca in Informatica.
- Galassi, U., Giordana, A. and Saitta, L. 2007a. Incremental construction of structured hidden Markov models, in *Proceedings of the IJCAI - IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, India, edited by Veloso, M.M., 798-803.
- Galassi, U., Giordana, A. and Saitta, L. 2007b. Structured hidden Markov Model: A general framework for modeling complex sequences, in *AI*IA 2007: Artificial intelligence and human-oriented computing, Proceedings of the 10th Congress of the Italian Association for Artificial Intelligence*, edited by Basili, R. and Pazienza, M.T., Rome, Italy, September 10-13, 2007, Lecture Notes in Computer Science 4733 Springer 2007, **4733**, 290-301.
- Herbers, J.M. 1985. Seasonal structuring of a north temperate ant community. *Insectes Sociaux*, **32**, 224-240.
- Holldobler, B. 1980. Canopy orientation: a new kind of orientation in ants. *Science*, **210**, 86-8.
- Holldobler, B. 1984. Communication during foraging and nest-relocation in the African stink ant, *Paltothyreus tarsatus*. *Z. Tierpsychol*, **65**, 40-52.
- Holldobler, B. and Wilson, E.O. 1990. *The ants*. Cambridge, MA: Harvard University Press.
- Bonafonte, A., Josep. V. and Nogueiras, A. 1996. *Duration modeling with expanded Hmm applied to speech recognition*. <http://www.asel.udel.edu/icslp/cdrom/vol2/809/a809.pdf>
- Lafferty, J., McCallum, A. and Pereira, F. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data, in *Proceedings of the 18th International Conference on Machine Learning*. San Francisco, CA: Morgan Kaufmann, 282-9.
- Levinson, S.E. 1986. Continuous variable duration hidden Markov models for automatic speech recognition. *Computer Speech and Language*, **1**, 29-45.
- Möglich, M. 1978. Social organization of nest emigration in *Leptothorax* (Hym. Form.). *Insectes Sociaux*, **25**, 205-25.
- Molet, M., Chittka, L., Stelzer, R.J., Streit, S. and Raine, N.E. 2008. Colony nutritional status modulates worker responses to foraging recruitment pheromone in the bumblebee *Bombus terrestris*. *Behavioral Ecology and Sociobiology*, **62**, 1919-26.
- Murphy, K.P. 2002. *Dynamic Bayesian Networks: Representation, Inference and Learning*, Ph.D thesis, UC, Berkeley, Department of Computer Science.
- Natarajan, P. and Nevatia, R. 2008. View and scale invariant action recognition using multiview shape-flow models, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008*, Anchorage, Alaska, 1-8.
- Partridge, L.W., Partridge, K.A. and Franks, N.R. 1997. Field survey of a monogynous leptothoracine ant (Hymenoptera, Formicidæ): evidence of seasonal polydomy? *Insectes Sociaux*, **44**, 75-83.

- Pezon, A., Denis, D., Cerdan, P., Valenzuela, J. and Fresneau, D. 2005. Queen movement during colony emigration in the facultatively polygynous ant *Pachycondyla obscuricornis*. *Naturwissenschaften*, **92**, 35-39.
- Pylkkonen, J. and Kurimo, M. 2004. Using Phone Durations in Finnish Large Vocabulary Continuous Speech Recognition, in *Proceedings of the 6th Nordic Signal Processing Symposium*, NORSIG 2004, <http://www.cis.hut.fi/jpylkkon/pylkkonen04norsig.pdf>.
- Rabiner, L. and Juang, B. 1993. *Fundamentals of Speech Recognition*. Englewood Cliffs, NY: Prentice Hall.
- Rabiner, L.R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of IEEE*, **77(2)**, 257-86.
- Robinson, E.J., Richardson, T.O., Sendova-Franks, A.B., Feinerman, O. and Franks, N.R. 2009. Radio tagging reveals the roles of corpulence, experience and social information in ant decision making. *Behavioral Ecology and Sociobiology*, **63**, (5), 779.
- Smallwood, J. 1982. Nest relocations in ants. *Insectes Sociaux*, **29**, 138-47.
- Streit, S., Bock, F., Pirk, C.W. and Tautz, J. 2003. Automatic life-long monitoring of individual insect behavior now possible. *Zoology*, **106**, 169-71.
- Sumner, S., Lucas, E., Barker, J. and Isaac, N. 2007. Radio-tagging technology reveals extreme nest-drifting behavior in a eusocial insect. *Current Biology*, **17**, 140-5.
- Tweed, D., Fisher, R., Bins, J. and List, T. 2005. Efficient hidden semi-Markov model inference for structured video sequences, in *Proceedings of the 2nd Joint IEEE Int. Workshop on VSPETS*, 247-254.
- Wilson, E.O. 1971. *The insect societies*. Cambridge, MA., Harvard University Press.