

An Insect-Based Method for Learning Landmark Reliability Using Expectation Reinforcement in Dynamic Environments

Zenon Mathews, Paul F.M.J. Verschure and Sergi Bermúdez i Badia

Abstract—Navigation in unknown dynamic environments still remains a major challenge in robotics. Whereas insects like the desert ant with very limited computing and memory capacities solve this task with great efficiency. Thus, the understanding of the underlying neural mechanisms of insect navigation can inform us on how to build simpler yet robust autonomous robots. Based on recent developments in insect neuroethology and cognitive psychology, we propose a method for landmark navigation in dynamic environments. Our method enables the navigator to learn the reliability of landmarks using an expectation reinforcement method. For that end, we implemented a real-time neuronal model based on the Distributed Adaptive Control framework. The results demonstrate that our model is capable of learning the stability of landmarks by reinforcing its expectations. Also, the proposed mechanism allows the navigator to optimally restore its confidence when its expectations are violated. We also perform navigational experiments with real ants to compare with the results of our model. The behavior of the proposed autonomous navigator closely resembles real ant navigational behavior. Moreover, our model explains navigation in dynamic environments as a memory consolidation process, harnessing expectations and their violations.

I. INTRODUCTION

In recent years, a great deal of progress has been made in the field of autonomous navigation in unknown environments. Moreover, autonomous navigation has seen a great interest with the advance of embedded systems in cars, aerial vehicles, blimps, humanoid robots etc. Nevertheless, even with the use of global positioning information systems, autonomous navigation remains a challenge in spite of the enormous advances in computing power and classical branches of robotics [1], [2], [3], [4].

On the other hand, biological systems solve complex navigational tasks at levels unparalleled by artificial systems. For example, a wide range of foraging animals from mammals to insects demonstrate remarkably stable navigational skills in unknown dynamic environments [5], [6], [7]. Biomimetic robotics aims at capturing such capabilities of biological systems to construct more advanced artificial systems. Indeed, in recent years, the control design of artificial autonomous systems have seen a shift from mere symbolic artificial

intelligence (sense-plan-act) to newer concepts like situatedness and contextual intelligence as exhibited by biological systems [8]. In this context, the problem of information representation and its usage is of crucial importance to both artificial and biological systems [9].

In biological systems there has been the controversy of information representation in the context of navigation, between the map-based and mapless systems. On the one hand, the concept of *cognitive map* for navigation, carried out mainly by Tolman [10], was fuelled by the discovery of the so-called place cells in the hippocampus of the rat and has widely increased our understanding of cognitive navigation mechanisms [11], [12]. It spawned early research on navigational strategies in cognitive neuroscience based on hippocampal representations of space [13], [14]. Until today a number of neurobiologically plausible models of navigation paradigms have been proposed for mobile robots [15], [16]. Also, newer neurobiological studies of place-cells have recently inspired biomimetic robotic models of map-based navigation [16], [17], [18], [19], [20].

On the other hand, while mammals are assumed to learn a place/map-like representation for foraging [11], [12], this does not seem the case in insects. Insect navigation has been studied for more than a century [21], [22], [23]. Interestingly, a wide range of findings suggest that insects do not rely on a map for solving foraging tasks [24], [22]. There is evidence that rather than using map-like representations, insects make optimal use of proprioception, landmark recognition and memory to navigate [7]. In particular, desert ants use sun position and visual panorama for heading direction computation [6], [25]. Complex allocentric navigational behaviors using mainly ego-centric cues can be seen both in mammals like rodents and in insects. Remarkably, insects like ants and bees have considerably low computational resources with only a few hundreds of thousands of neurons. Therefore, insect navigation studies are highly useful for revealing essential components for a computationally cheap navigation strategy. Only through building artificial models of ant navigation and by directly comparing the performance of the artificial system and the real ant, we can demonstrate that we capture the underlying navigation principles. This will then lead to the development of real-world robots that can navigate autonomously based fully on ant navigation strategies.

In our previous recent work we described a comprehensive mapless model, including chemical search, PI and landmark navigation, of insect navigation strategies [26]. That navigational strategy was based on the so called heading direction accumulation cells, which were recently hypothesized to

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216916 for NEUROCHEM and EU SYNTHETIC FORAGER (FP7-217148) projects.

All authors are with the SPECS laboratory at Institut Universitari de l'Audiovisual (IUA) of the technology department of Universitat Pompeu Fabra (UPF) in Barcelona, Spain (telephone: +34 93 542 22 01, zenon.mathews@upf.edu

P.F.M.J.V. is also with Institució Catalana de Recerca i Estudis Avançats (ICREA) Barcelona, Spain.

be present in the brain [27]. Nevertheless the problem of information gathering, representation and usage was not addressed. This becomes highly relevant as the autonomous navigator has to act in a dynamic world with non-static landmarks.

In this work we propose an autonomous navigation method in dynamic environments based on recent views from cognitive psychology [28] and insect neuroethology and physiology, which suggest that insects expect future events based on past experiences [29]. We use an expectation reinforcement paradigm to adapt the confidence of the artificial forager in encountered landmarks. We also propose a variant of an insect-like probabilistic search strategy suggested earlier to be used by insects upon expectation violation [30], [31]. Our model is based on the Distributed Adaptive Control (DAC) framework that organizes behavior in three levels of adaptive control [32], [33]. The implementation of the model is achieved using the large-scale neural simulator IQR [34] and tested on a simulated robot. In our experiments we first evaluate the capability of our model to solve a complex navigational task in an unknown dynamic environment. We further carry our experiments with real ants to compare their behavior with the robot model in the same navigational task. Our results show a successful autonomous navigation strategy in dynamic unknown environments and striking similarity to insect behavior. Moreover, our model explains autonomous navigation as a dynamic process of memory reconsolidation, which harnesses expectations that are readily available from past explorations.

II. METHODS

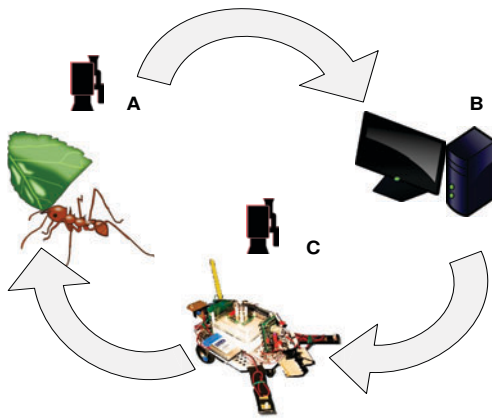


Fig. 1. **Insect behavioral analysis, modeling and testing on robots:** A) Real-world ant experimental studies are performed and the behavior of the ant is recorded using a tracking camera. Controlled manipulations of visual landmarks in the ant arena are made in order to analyze the ant behavior. B) We first model the ant behavior based on the understanding of the underlying neural principles. The ant experimental data is analyzed and fed into the *SyntheticAnt* simulation of the navigational model. This simulation is used to tune the parameters of our navigational model on a simulated robot. C) The real-world robot *SyntheticAnt* is tested with analogous manipulations of visual landmarks in the arena, allowing direct comparisons of the behaviors of real ant and the robot. The results of this is again used to design new ant experiments.

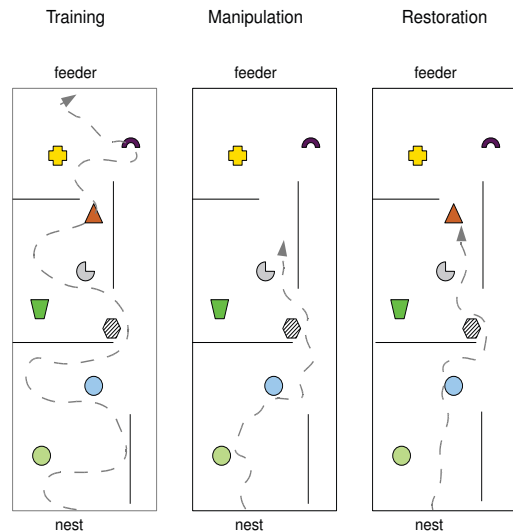


Fig. 2. **Autonomous Navigation Task:** **Left)** In the training session (also called foraging) the navigator leaves its nest to forage for food, traversing an arena with visual landmarks. These visual cues can be used by the navigator to memorize its routes. **Middle)** After several foraging runs, a manipulation is made in the arena (i.e. a landmark is displaced or removed as the red triangle in the figure), and the behavior of the navigator is observed. **Right)** The original constellation of the landmarks is restored and the behavior of the navigator is observed.

A. Navigational Task and the Test Environment

1) *Task:* Both the real ant and the artificial forager called *SyntheticAnt* are made to forage for the feeder (food location) in an initially unknown environment starting from its nest. The environment contains many visual landmarks which could be used by the forager for navigation. After many foraging runs, some landmarks are displaced or removed. The forager is then made to forage in this manipulated environment. In the next foraging runs, the original landmark constellation is restored. This task allows to study and compare three key issues of autonomous navigation:

- learning landmark navigation in a dynamic environment
- behavior upon detection of landmark manipulation
- confidence adaptation depending on the reliability of the landmarks after restoration.

See figure 2 for an illustration.

2) *Real Ant Experiments:* We perform real ant foraging experiments as described above using the desert ant (*genus Cataglyphis*). Landmark manipulation tasks are performed using several colored ants and the ant trajectories of the foraging runs are recorded using the AnTS overhead camera tracking system. It is important to note that in the real ant navigation experiments, the paper on the floor of the arena was replaced after each foraging run, in order to avoid the use of chemical cues by the ants.

3) *Modeling and Simulation Environment:* The modeling environment allows the input of the real ant trajectory for the simulated ant called the *SyntheticAnt*, allowing it to have the same perceptual input during foraging. Also free foraging behavior is available. The simulation replicates the real arena and the robot. It allows testing navigational paradigms before

they are tested on the real robot. In this work we only consider the simulated robot.

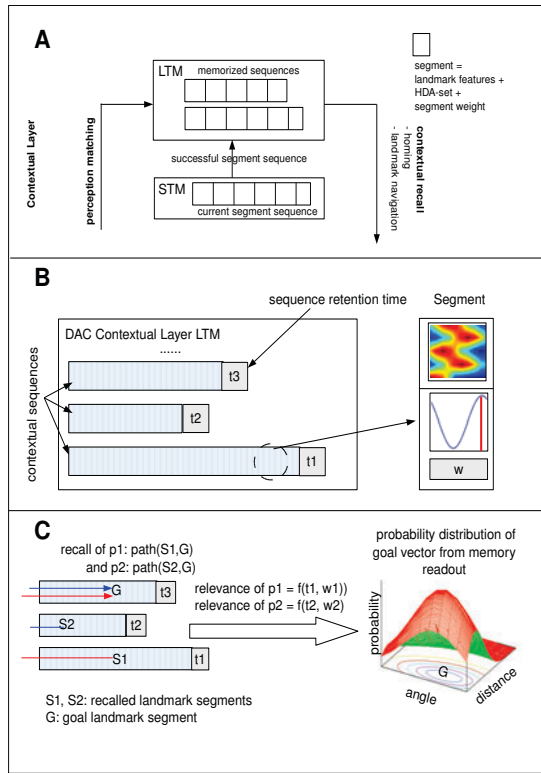


Fig. 3. (A) **DAC Contextual Layer** A segment contains the extracted landmark features, an HDA-set and the segment weight. These segments are sequenced temporarily in the short-term memory (STM) of the contextual layer whenever a landmark is encountered. Upon feeder detection, the contents of the STM are transferred into the long-term-memory (LTM). During recall phase (homing, landmark navigation), the LTM is matched against the current sensory events and an optimal trajectory is computed from recalled LTM segments. (B) **Sequencing**: During the foraging runs from the nest to the feeder, encountered landmarks are chained in the DAC contextual layer short-term memory (STM) together with the HDA set. Upon feeder detection the contents of the STM are transferred into the LTM and the HDA-set is reset. After several foraging runs, the LTM contains segment sequences of different lengths since in each foraging run only a subset of available landmarks are visited. Each LTM sequence has a retention time t owing to the *transiency* of memory and each segment has a weight w . (C) **Recall**: During the recall phase, the HDA-sets starting from the recalled segment to the goal segment are combined to compute the optimal homing vector. When the recalled segment and the goal segment are on different LTM sequences, the segments from the recalled segment to the nest on one sequence, and the nest to the goal segment on the other are combined (such a combination is called *path*). *Paths* are weighted according to the retention time of the sequence and the mean relevance weights of the segments of the recalled LTM sequence.

In this work we test the proposed navigational model on the simulated robot. As the neural implementation of the model is the very same for the simulated and the real *SyntheticAnts*, the transition from simulation to real world is readily manageable. More details of the experimental setup and the real robot version of the *SyntheticAnt* are discussed in [26].

B. DAC Contextual Layer

The Distributed Adaptive Control (DAC) architecture has been shown to be capable of optimizing behavior using

behavioral and perceptual learning in robots [32], [35]. DAC consists of three, tightly coupled layers for behavioral control; the reactive, adaptive and contextual layers. The reactive control layer provides the behaving system with a pre-wired repertoire of reflexes such as collision avoidance, chemosearch, homing etc. The adaptive layer provides the mechanisms for the processing and classification of sensory events. The contextual layer acquires, retains, and expresses sequential representations by means of short-term and long-term memories, figure 3, A. (In this work, as we are concerned with information representation and usage, we only deal with the contextual layer.) These representations are used to plan ongoing behavior, and have been shown to be compatible with formal Bayesian models of decision making [35]. We use the DAC contextual layer to learn sequences of landmarks together with their Heading Direction Accumulator (HDA) set. Using this the navigator can memorize several trajectories from the nest to feeder, leading through different landmarks. The DAC contextual layer recall mechanism is used when in the test scenario, a particular landmark is encountered, to recall the vectors to other landmarks. The details of this mapless navigational strategy is discussed in our previous work [26]. In this work, we consider the use of DAC contextual layer and its recall mechanism for learning to navigate in unknown and dynamic environments.

C. Dynamic Memory Consolidation Using Expectations

Cognitive psychology has recently begun to embrace a new position recognizing memory as a highly dynamic process [28]. In this new view, remembering and forgetting are not merely transient processes; moreover they are achieved through a highly dynamic (re)consolidation process. Strong support for such a dynamic memory comes from animal neuroscience studies as reviewed in [28].

At the same time, insects have also been shown to expect events using memories acquired in the past [29], [36]. Building upon these two recent advances in neuroscience and cognitive psychology, we propose a dynamic memory model for insect-like mapless navigation in uncertain environments, which uses expectations to consolidate or forget already acquired memories, figure 4. We propose a dynamic contextual layer LTM of Distributed Adaptive Control for mapless navigation using heading direction accumulation cells.

1) Transiency and Confidence in Memory Sequences:

We elaborate how transiency and confidence in memory traces are formulated for mapless navigation in uncertain environments. First we consider the situation, where the *SyntheticAnt* is kidnapped after many foraging runs and placed at an arbitrary visual landmark. The details of the memory recall using the DAC contextual layer for mapless navigation can be consulted in our previous work [26]. Memory recall of DAC contextual layer is used to compute an optimal path to another given landmark; we further refer to this as the *memory answer*. We define a vector \vec{V} as a two dimensional vector indicating the angle γ and distance ψ of a given memory answer: $\vec{V} = [\gamma, \psi]^T$. Assuming that n LTM

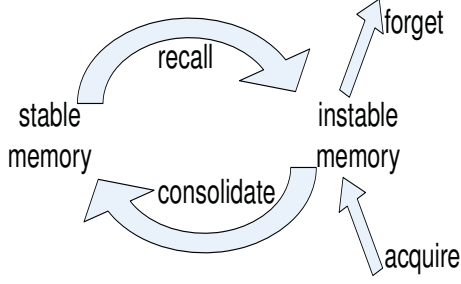


Fig. 4. **Dynamic Memory Reconsolidation Schema:** As a memory is acquired it enters the instable state and then is consolidated into the stable state. Nevertheless, memories in the stable state can reenter the instable state upon recall. Reconsolidation again stabilizes this memory or otherwise get forgotten [28]. We propose such a dynamic memory consolidation process using the Distributed Adaptive Control contextual layer for insect-like mapless navigation in uncertain environments.

sequences were recalled, we are interested in computing the probability distribution:

$$P(\vec{V}|t, \vec{V}_1^t, t_1^{ret}, r_1^{seg} \dots \vec{V}_n^t, t_n^{ret}, r_n^{seg}) \quad (1)$$

where \vec{V} is the vector indicating the goal, t is time, \vec{V}_i^t is the vector suggested by LTM sequence i , t_i^{ret} is *retention time* of LTM sequence i and r_i^{seg} is the segment relevance of the recalled segment of LTM sequence i .

We look at the contribution of one LTM sequence i to the overall memory answer:

$$P(\gamma\psi|t, \vec{V}_i^t, t_i^{ret}, r_i^{seg}) \quad (2)$$

First, we consider equation 2 without the retention times:

$$P(\gamma\psi|t, \vec{V}_i^t, r_i^{seg}) \quad (3)$$

Using conditional independence of angle and distance we get

$$\begin{aligned} P(\gamma\psi|t, \vec{V}_i^t, r_i^{seg}) &= P(\gamma|t, \vec{V}_i^t, r_i^{seg}) \\ &\quad \cdot P(\psi|t, \vec{V}_i^t, r_i^{seg}) \\ &\quad \cdot P(t, \vec{V}_i^t, r_i^{seg}) \end{aligned} \quad (4)$$

where $\vec{V}_i^t = [\gamma_i^t, \psi_i^t]^T$. We formulate $P(\gamma|t, \vec{V}_i^t, r_i^{seg})$ and $P(\psi|t, \vec{V}_i^t, r_i^{seg})$ as Gaussian distributions centered around the means γ_i and ψ_i respectively with time dependent variances, signifying that higher weight segments have a higher influence on the total memory answer as they have a smaller variance. This dynamically adapting variance is used to weight successful segments more and unsuccessful ones less.

$$P(\gamma|t, \vec{V}_i^t, r_i^{seg}) = \mathcal{N}(\gamma_i, f_1(r_i^{seg})) \quad (5)$$

$$P(\psi|t, \vec{V}_i^t, r_i^{seg}) = \mathcal{N}(\psi_i, f_2(r_i^{seg})) \quad (6)$$

where f_1 and f_2 are exponential growth functions of variance:

$$f_1(x) = a_1 e^{k_1(x)} \quad (7)$$

$$f_2(x) = a_2 e^{k_2(x)} \quad (8)$$

As all memory answers should be seen as equally probable and no correlations of distance, angle and retention-times of LTM sequences are known, we assume the uniform distribution:

$$P(t, \vec{V}_i^t, r_i^{seg}) = \mathcal{U} \quad (9)$$

Now we consider the contributions of all selected LTM sequences weighted by segment retention times t_i^{ret} :

$$P(\vec{V}|t, \vec{V}_1^t, t_1^{ret}, r_1^{seg} \dots \vec{V}_n^t, t_n^{ret}, r_n^{seg}) = \sum_i \frac{t_i^{ret}}{t_{tot}} P(\gamma|t, \vec{V}_i^t, r_i^{seg}) P(\psi|t, \vec{V}_i^t, r_i^{seg}) P(t, \vec{V}_i^t, r_i^{seg}) \quad (10)$$

where $t_{tot} = \sum_i t_i^{ret}$. Equation 10 signifies that the sequence retention times are used to weight the shares of LTM sequences to the final memory answer. In other words, sequences are weighted so that memory acquired further back in the past is weighted less than more recent memory.

2) Building up Confidence in Uncertain Environments:

When a landmark (or feeder) at a known location ceases to be available or its position is manipulated, insects have been shown to exhibit search patterns that optimize rediscovery of the landmark (or feeder) [30]. We further refer to landmark manipulations of all kinds as *manipulation*. It has also been proposed that such search patterns in insects could be modeled using Lévy walks [31]. Lévy walks are characterized by a distribution function

$$P(l_j) \sim l_j^{-\mu} \quad (11)$$

with $1 \leq \mu \leq 3$, where l_j is the walk length. The direction of the walk is drawn from a uniform distribution. The natural parameter μ has been shown to be optimal at the value 2 for modeling insect *manipulation* behavior, yielding an inverse square power-law distribution [30].

Here we propose a modified version of Lévy walk, to model the confidence building behavior exhibited by ants upon *manipulation*. The key idea is to propagate the initialization point of the Lévy walk towards the nest, so that the probability of encountering known landmarks is increased. Thus, the forager can slowly build up *confidence* about the distance and direction to the navigational goal. The proposed version of Lévy walk is summarized in the pseudo-algorithm below:

```

while |con| ≤ ε do
  t ← 0
  while t ≤ θ do
    perform Lévy walk from i
    upon landmark detection : update con
  end while

```

```

if  $i = nest$  then
  reset  $i$  to manipulation point
else
  propagate  $i$  towards nest
end if
end while

```

where con is the mean confidence about the goal direction and distance, i the current position at which the Lévy walk is initialized, and θ and ϵ two appropriate thresholds. The artificial forager is made to switch to the above described search mode, when either an expected landmark is not encountered or when a landmark with a very low confidence is found. The duration of the Lévy walk is much higher in the first case and linearly decreases with the confidence in the landmark, controlled by setting the above parameters ϵ and θ . Our new proposal of the Lévy walk helps the reinforcement of expectations as it lets the autonomous navigator validate its expectations by going back to known terrain, instead of searching randomly in the whole arena. This is strongly suggested by observed ant behavior in similar situations and is evaluated in our experiments.

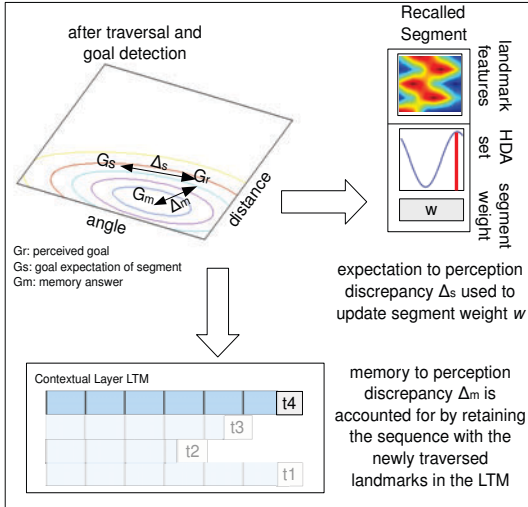


Fig. 5. **Expectation reinforcement for memory consolidation:** Discrepancy between expected paths and computed paths are used to consolidate memory by means of setting LTM segment weights and writing new LTM sequences.

3) *Adapting Landmark Reliability Measure Using Expectations:* The forager's belief about the reliability of individual landmarks should be updated each time a landmark is encountered. For this we use an expectation reinforcement mechanism, which gives rise to a dynamic memory reconsolidation process [28]. Expectations for positions of individual landmarks can be read out from the DAC contextual layer memory. The discrepancy between perceived position of a landmark and the expected position is used to update the reliability of the landmark, i.e. the weight of a DAC contextual layer segment, as shown in figure 5.

To update the DAC LTM segment weight using the dis-

crepancy, we use the exponential decay function

$$r_{t+1}^{seg} = r_t^{seg} e^{-\lambda d} \quad (12)$$

where d is the discrepancy and t is time step. This causes the segment weight to fall exponentially from its current value if the discrepancy is high. Discrepancies are normalized ($0 \leq d \leq 1$) so that $d = 1$ for the highest possible discrepancy (the length of the diagonal of the arena). The natural parameter λ was shown to optimally fit observed insect behavior at the value 2 [30]. The above decay is applied only if the discrepancy is above a fixed threshold. Otherwise, the segment weight is allowed to grow according to a linear function.

III. RESULTS

First we test the ability of the proposed model to learn the reliability of landmarks. For this purpose we test the simulated *SyntheticAnt* in an arena with 10 visual landmarks. During the foraging no landmark manipulation were carried out. After that, during the test 9 of the 10 visual landmarks were displaced randomly in each run, i.e. there was only a single stable landmark. After the foraging runs, we test the feeder direction and angle from the nest, computed as discussed earlier, using each of the 10 landmarks. Note that for this the distance and angle to the feeder from an arbitrary landmark is computed first using DAC contextual recall (as in [26]), which is then added to the vector leading from the nest to this landmark. This allows to visualize the belief of the forager in the reliability of each landmark as a probability distribution, as plotted in figure 6. The *SyntheticAnt* learns through expectation reinforcement that only landmark number 10 is stable. The evolution of the confidence in the landmarks are shown in figure 6 A (before the test runs), B (after 5 test runs) and C (after 10 test runs).

Note that a single simulation trial approximately runs for the same time as an ant foraging run and we perform qualitatively the same kind of landmark manipulations in the real ant experiment and in the simulation. We also use similar landmarks in the simulation as used in the ant experiments. Nevertheless, for simplicity we do not consider complex landmarks such as shadows, light direction etc. known to be used by real ants in the simulation. Our simulations run at about 30Hz on an Intel(R) Core(TM)2 Duo CPU 2.66GHz machine under openSUSE 10.3.

Table I shows the learned parameters of confidence in the individual landmarks, which corresponds to figure 6, C. The growing skewness of the distributions from left to right shows growing asymmetry in the distribution, indicating that the navigator increases confidence in some landmarks and decreases in some others.

The segment retention time t^{ret} is the number of the latest foraging run in which the corresponding landmark was encountered, i.e. giving a high value to recent runs. The segment weight r^{seg} is initialized to 1. Note that landmark 1 was seen more recently than landmark 10, but as the segment weight of landmark 1 is very low it has low influence on the overall answer. As shown in figure

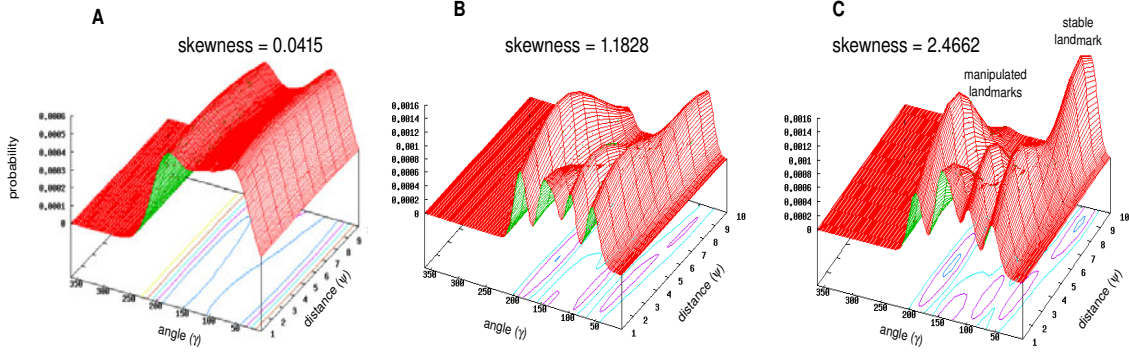


Fig. 6. **Expectation reinforcement for learning stable landmarks:** **A)** Before the test runs (but after several foraging runs) *SyntheticAnt* has encountered all 10 landmarks, but has the same confidence in all of them. Some variance is seen as the LTM sequence acquisition time is different. **B)** After 5 test runs, during which 9 out of 10 landmarks were displaced, the probability distribution for the confidence changes. This shows that the *SyntheticAnt* is slowly adapting the confidence in the landmarks. **C)** The confidence probability distribution after 10 test runs. The plot shows the computed home distances and angles (as a probability distribution) using individual landmarks using memory recall for each landmark. The stable landmark has a greater influence than instable ones on the overall memory readout (higher probability). The forager thus learns the stability of the individual landmarks. The indicated skewness values are the third central moments of sample values, divided by the cube of their standard deviations. The growing skewness from left to right indicates growing asymmetry in the distribution.

6, the confidence for stable landmarks is reinforced and is higher than instable landmarks; i.e. the forager learns the stability of the landmarks. In short, higher frequency of landmark position manipulations lead to lower segment weights, meaning lower confidence represented by higher covariances of gaussians. Therefore higher frequency manipulations of landmarks are reflected directly in the joint probability distribution with lower probabilities. And further the higher the search intensity the lower the confidence.

Given the above result, we now investigate how the *SyntheticAnt* can regain its confidence, once its expectation is violated. Expectation is violated when an expected landmark is missing or if a low confidence landmark is encountered. As described before, the *SyntheticAnt* falls into the Lévy search mode and the intensity of the search is higher if the landmark is missing. We evaluate how confidence is regained with time and how the search intensity influences (figure 7). The θ value indicates the search intensity ($\theta_3 > \theta_2 > \theta_1$). The results show that the confidence is regained with time, where high intensity searches allow to reach a higher confidence threshold in a shorter amount of time.

Given the above results, we predict that our model should behaviorally be similar to a real ant. This prediction is

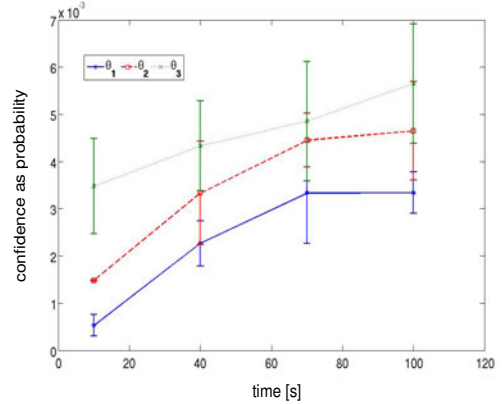


Fig. 7. **Confidence recovery and search intensities:** The recovery of confidence with time (after initialization of the Lévy search) is indicated. High intensity searches (indicated by the θ values) reach a certain confidence threshold quicker. A high intensity Lévy search is initialized when a landmark is missing and a low intensity one is launched when a low confidence landmark is encountered ($\theta_3 > \theta_2 > \theta_1$).

TABLE I

LANDMARK STABILITY LEARNING EXPERIMENT

Landmark	Distance	Angle	r^{seg}	t^{ret}
1	5.1	181.5	0.019569	10
2	2.5	220.3	0.09090	5
3	3.3	94.6	0.03333	3
4	1.1	144.4	0.04	4
5	2.4	124.6	0.06666	5
6	3.5	91.0	0.04348	6
7	3.5	55.9	0.05263	7
8	4.2	64.1	0.04762	1
9	5.7	44.1	0.05263	3
10	9.1	35.2	0.9	9

tested using real ants in the same landmark manipulation and restoration task. To perform the same landmark manipulation task, we first test the real ant and feed the recorded trajectory to the *SyntheticAnt*, so that both navigators have the same perceptual input. Further we consider the *confidence* of the real ant or the *SyntheticAnt* in a landmark as inversely correlated with the computed trajectory density. The ant is let to forage from feeder to nest for 21 runs, during which no landmark manipulations are made. In the 22nd run, two landmarks are removed at the indicated position, figure 8 B. The ant does a high intensity search and tracks back almost to its nest. In the next run, the landmarks are again restored in the original constellation. Nevertheless, the ant does not *trust* these manipulated landmarks in the following runs, as indicated by the high densities around these landmarks. Such behavior was prototypical in most experimented ants. We

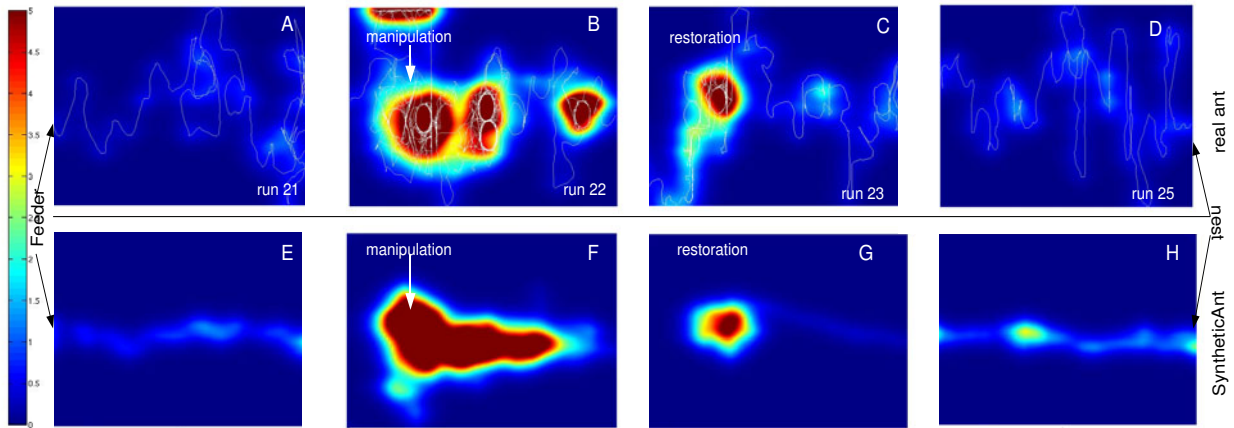


Fig. 8. **Real ant benchmark in landmark manipulation task:** The upper panel and lower panels show real and *Synthetic* ant data in the same landmark manipulation task. The ant performs 25 foraging runs from nest (located at the right end of the arena) to the feeder (at the left end of the arena). The arena contains displaceable visual landmarks and also non-displaceable obstacles, both of which the ant cannot walk over. Ant trajectories are indicated by white lines. The density maps are computed from the trajectories; we define here density as inversely proportional to the *confidence* in the landmarks. **A)** Ant trajectory in the 21st foraging run. This trajectory shows low densities all over the arena, indicating high confidence of the animal in the visual landmarks. **B)** Ant trajectory in the 22nd run, where some landmarks were manipulated at the indicated positions. The animal finds high discrepancy between its expectation and perception. This results in very low confidence and it performs correction manoeuvres by going back towards the nest. **C)** In the next run (23) all the visual landmarks are again placed in the positions as in runs 1 to 21. The density plots indicate that the ant does not trust the landmarks in the area that violated its expectations in the past. **D)** The confidence of the animal in the landmarks increases again through expectation reinforcement, as indicated by the low densities throughout the arena. The lower panel shows the performance of the *SyntheticAnt* in the same experiment. **E)** Before the manipulation, the *SyntheticAnt* has the same high confidence value in all landmarks and therefore does the traversal on almost a straight line. **F)** Upon manipulation, a high intensity Lévy search is initiated and propagated back home. **G)** Upon restoration, a low intensity search at the low confidence landmark is performed, as captured by the density plot. **H)** The confidence in the manipulated landmark recovers slowly as indicated by the density plot. The task resolution times for the same experiment is shown in figure 9.

compare the behavior of the real ant, with the behavior of the *SyntheticAnt* in the same task as above. The artificial forager switches to the above described random search using Lévy walk, when an expected landmark is not encountered (figure 8, F) or when a landmark with a very low confidence is encountered (figure 8, G). Our results show striking similarity in the behavior of the real and the *SyntheticAnt* in similar situations of landmark manipulation tasks. The task resolution time before and after manipulation shows the time taken by the forager to reach the feeder. This is plotted in figure 9. Both the real ant and the *SyntheticAnt* follow the same trend in the task resolution times. The observed tracking back to known terrain upon expectation violation in the real ant, is captured in our model using the proposed new variant of the Lévy walk. This new mechanism fits the context of expectation reinforcement, as it allows the navigator to track back to known terrain to test the validity of its expectations, while at the same time exploring new terrain.

IV. CONCLUSIONS

We presented an expectation reinforcement mechanism based on recent advances in cognitive psychology and insect studies to allow an artificial forager to navigate in unknown dynamic environments. Our model is based on the Distributed Adaptive Control and implemented using the neural network simulator IQR. The proposed navigational strategy enables the navigator to learn the reliability in the landmarks by using expectation reinforcement. The proposed version

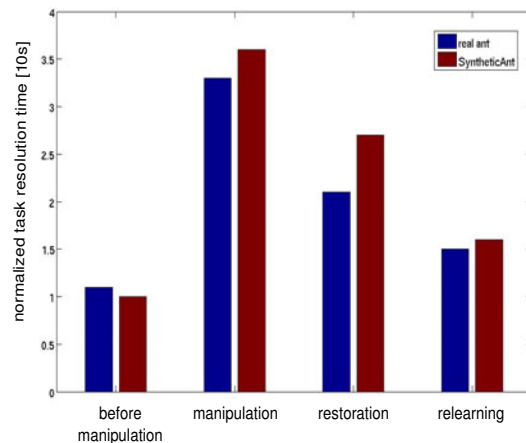


Fig. 9. **Task resolution time:** Here we look at the normalized task resolution times before, at and after manipulation of the landmarks. The real ant and the *SyntheticAnt* exhibit similar behavior in the time domain. The task resolution time rises in both cases when a landmark is manipulated, and falls again upon restoration of the landmark constellation. But it takes some more runs before the task resolution time gets back at the baseline before the manipulation. This data corresponds to the experiment in figure 8.

of Lévy walk supports the idea of expectation reinforcement by enabling the navigator to trace back to known terrain whenever expectations are violated. We tested our navigational model on the simulated robot called *SyntheticAnt*, and compared its behavior to the real ant. The striking

behavioral similarity of our model to the real ant suggests the strength of the proposed model, as real ants are still far better navigators in unknown dynamic environments than most artificial autonomous navigators. We demonstrated how the robot could learn the stability of the given landmarks without any available prior information. Moreover, our model formulates navigation as a dynamic memory reconsolidation process, which makes use of the expectation reinforcement mechanism. The primary aim hereby is not just to compete with existing artificial robotic navigation models, but even more to understand the principles underlying the robust navigation of insects. Our belief is that by modeling ant navigation and then directly comparing the results of our navigational model to real ant navigational behavior, we could better understand the capabilities and limits of the model. This again should lead to the new generation of insect-like navigation paradigms deployed on real-world robots. In future work we plan first to test the model on the real *SyntheticAnt* robot in the test arena (as in [26]) and then on other outdoor wheeled robots.

V. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216916 for NEUROCHEM and EU SYNTHETIC FORAGER (FP7-217148) projects.

REFERENCES

- [1] R. Benenson, S. Petti, T. Fraichard, and M. Parent, "Towards urban driverless vehicles," *International Journal of Vehicle Autonomous Systems*, vol. 6, no. 1/2, pp. 4–23, 2006.
- [2] H. Andreasson, A. Treptow, and T. Duckett, "Localization for mobile robots using panoramic vision, local features and particle filter," *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pp. 3348–3353, April 2005.
- [3] E. Royer, J. Bom, M. Dhome, B. Thuilot, M. Lhuillier, and F. Marmoiton, "Outdoor autonomous navigation using monocular vision," *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pp. 1253–1258, Aug. 2005.
- [4] A. J. Davison and N. Kita, "Sequential localisation and map-building for real-time computer vision and robotics," *Robotics and Autonomous Systems*, vol. 36, no. 4, pp. 171 – 183, 2001.
- [5] D. Stephens, J. Brown, and R. Ydenberg, *Foraging: Behavior and Ecology*. Chicago: University of Chicago Press, 2007.
- [6] R. Wehner, "Desert ant navigation: how miniature brains solve complex tasks," *Journal of comparative physiology. A. Neuroethology, sensory, neural, and behavioral physiology*, vol. 189, no. 8, pp. 579–588, Aug 2003.
- [7] T. S. Collett and M. Collett, "Memory use in insect visual navigation," *Nature reviews. Neuroscience*, vol. 3, no. 7, pp. 542–552, Jul 2002.
- [8] L. Paletta, E. Rome, and H. Buxton, "Attention architectures for machine vision and mobile robots," *Neurobiology of Attention*, p. 642..648, 2005.
- [9] M. Witkowski and D. Randell, "A model of modes of attention and inattention for artificial perception," *Bioinspiration and biomimetics*, vol. 2, no. 3, pp. S94–S115, Sep 2007.
- [10] E. C. Tolman, "Cognitive maps in rats and men." *Psychological review*, vol. 55, no. 4, pp. 189–208, July 1948.
- [11] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat," *Brain research*, vol. 34, no. 1, pp. 171–175, Nov 1971.
- [12] J. O'Keefe and L. Nadel, *The Hippocampus as a Cognitive Map*. Oxford University Press, USA, 1978-12-07 1978.
- [13] J. O'Keefe, *The hippocampal cognitive map and navigational strategies*. Oxford University Press, USA, 1991.
- [14] R. U. Muller, M. Stead, and J. Pach, "The hippocampus as a cognitive graph," *The Journal of general physiology*, vol. 107, no. 6, pp. 663–694, Jun 1996.
- [15] M. O. Franz and H. A. Mallot, "Biomimetic robot navigation," *Robotics and autonomous Systems*, vol. 30, pp. 133–153, 2000.
- [16] A. Arleo and W. Gerstner, "Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity," *Biological Cybernetics*, vol. 83, no. 3, pp. 287–299, 2000, article.
- [17] A. Guanella and P. F. M. J. Verschure, "Prediction of the position of an animal based on populations of grid and place cells: a comparative simulation study," *Journal of Integrative Neuroscience*, vol. 6, no. 3, pp. 433–446, 2007.
- [18] R. Wyss, P. Knig, and P. F. M. J. Verschure, "Invariant representations of visual patterns in a temporal population code," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 1, pp. 324–329, 01/2003 2003.
- [19] C. Giovannangeli, P. Gaussier, and G. Desilles, "Robust mapless outdoor vision-based navigation," *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 3293–3300, Oct. 2006.
- [20] C. Giovannangeli and P. Gaussier, "Autonomous vision-based navigation: Goal-oriented action planning by transient states prediction, cognitive map building, and sensory-motor learning," *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 676–683, Sept. 2008.
- [21] K. von Frisch, "Decoding the language of the bee," *Science (New York, N.Y.)*, vol. 185, no. 4152, pp. 663–668, Aug 23 1974.
- [22] T. S. Collett, "Insect navigation: visual panoramas and the sky compass," *Current biology : CB*, vol. 18, no. 22, pp. R1058–61, Nov 25 2008.
- [23] A. Si, M. V. Srinivasan, and S. Zhang, "Honeybee navigation: properties of the visually driven 'odometer'," *The Journal of Experimental Biology*, vol. 206, no. Pt 8, pp. 1265–1273, 2003.
- [24] R. Wehner, M. Boyer, F. Loertscher, S. Sommer, and U. Menzi, "Ant navigation: one-way routes rather than maps," *Current biology : CB*, vol. 16, no. 1, pp. 75–79, Jan 10 2006.
- [25] S. Akesson and R. Wehner, "Visual navigation in desert ants *cataglyphis fortis*: are snapshots coupled to a celestial system of reference?" *The Journal of experimental biology*, vol. 205, no. Pt 14, pp. 1971–1978, Jul 2002.
- [26] Z. Mathews, M. Lechón, J. M. Blanco Calvo, A. Dhir, A. Duff, S. Bermúdez i Badia, and P. F.M.J. Verschure, "Insect-like mapless navigation based on head direction cells and contextual learning using chemo-visual sensors," *Intelligent Robots and Systems (IROS 2009), 2009 IEEE/RSJ International Conference on*, 2009.
- [27] J. L. Kubie and A. A. Fenton, "Heading-vector navigation based on head-direction cells and path integration," *Hippocampus*, Dec 12 2008.
- [28] K. Nader and O. Hardt, "A single standard for memory: the case for reconsolidation," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 224–234, mar 2009.
- [29] M. Gil, R. Menzel, and R. J. De Marco, "Does an insect's unconditioned response to sucrose reveal expectations of reward?" *PLoS ONE*, vol. 3, no. 7, p. e2810, 07 2008.
- [30] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, "Optimizing the success of random searches," *Nature*, vol. 401, no. 6756, pp. 911–914, oct 1999.
- [31] A. M. Reynolds, A. D. Smith, D. R. Reynolds, N. L. Carreck, and J. L. Osborne, "Honeybees perform optimal scale-free searching flights when attempting to locate a food source," *Journal of Experimental Biology*, vol. 210, no. 21, pp. 3763–3770, nov 2007.
- [32] P. F. Verschure, T. Voegtlin, and R. J. Douglas, "Environmentally mediated synergy between perception and behaviour in mobile robots," *Nature*, vol. 425, no. 6958, pp. 620–624, Oct 9 2003.
- [33] P. F. M. J. Verschure and P. Althaus, "A real-world rational agent: unifying old and new ai," *Cognitive Science A Multidisciplinary Journal*, vol. 27, no. 4, pp. 561–590, 2003.
- [34] U. Bernardet, M. Blanchard, and P. F. M. J. Verschure, "Iqr: a distributed system for real-time real-world neuronal simulation," *Neurocomputing*, vol. 44-46, pp. 1043–1048, 2002.
- [35] P. F. M. J. Verschure and P. Althaus, "A real-world rational agent: Unifying old and new ai," *Cognitive Science*, vol. 27, pp. 561–590, 2003.
- [36] M. Gil, R. J. D. Marco, and R. Menzel, "Learning reward expectations in honeybees," *Learning & Memory (Cold Spring Harbor, N.Y.)*, vol. 14, no. 7, pp. 491–496, July 2007.