

Emergence of Memory-like Behavior in Reactive Agents Using External Markers

Ji Ryang Chung and Yoonsuck Choe

Department of Computer Science and Engineering
Texas A&M University, College Station, TX, USA
jchung@cs.tamu.edu, choe@tamu.edu

Abstract

Early primitive animals with simple feed-forward neuronal circuits were limited to reactive behavior. Through evolution, they were gradually equipped with memory and became able to utilize information from the past. Such memory is usually implemented with recurrent connections and certain behavioral changes are thought to precede the reconstitution of the neuronal circuit's topology. If so, what could have been the behavior to drive such a rewiring? Our hypothesis is that the secretion and detection of chemical markers in the environment could be a precursor of internal memory. We will show how memory-like behavior can be expressed in memoryless reactive agents by taking advantage of the external chemical markers. Our results show that given chemical marker use, reactive agents are able to develop intelligent strategies in solving a biologically plausible food foraging task requiring spatial memory. We also found interesting analogy between the evaporation of the chemical markers and the recency effect in memory and how it affects the foraging strategy. These results are expected to help us better understand the possible evolutionary route from reactive to cognitive agents.

1. Introduction

From Lamarck and Darwin to contemporary evolutionary biologists, a consensus seems to be on the premise that behavioral changes make the agent-environment relationship anew, which in turn generates a novel evolutionary pressure and leads to morphological changes [16, 9]. Under the same premise, we have already shown elsewhere that the emergence of memory could be achieved through the internalization of the external marker manipulating behavior [6]. Our argument was that producing and detecting external markers could have served as an intermediate stage between reactive and contemplative behavior (cf. [12] on the use of “inert” matter for augmenting cognition). Here, we focused particularly on the use of external chemical mark-

ers because as mentioned in [6], various evidence suggest the relation between olfaction and the evolution of memory.

We deliberate that the volatility of chemical markers could have fostered the individual intelligence. We tested our hypothesis by comparing reactive agents (feed forward neural network controllers) using either non-evaporating or evaporating chemical markers in a simple food foraging task. Our results show interesting memory-like behavior that would normally require a recurrent architecture, through strategic uses of the external chemical markers. Moreover, better strategy and finer food foraging trails were developed when we let the chemical markers evaporate. This does not only support our previous hypothesis by showing a more detailed and more biologically plausible experiment and analysis, but also suggests that it could be particularly the use of external chemical markers that induced a fully internalized memory system.

The rest of the paper is organized as follows. In Section 2, we detail the food foraging task with two types of chemical markers. Section 3 includes experiments and results, followed by discussion and conclusion (Section 4 and 5).

2. Methods

2.1. Task: Food Foraging

Food foraging is the most essential ability for survival in animals. Many birds, ants, and mammals do not always consume food on site but move them for various purposes including to feed the young, to cache, or to reduce competition from others [13, 14]. Inspired by this, we designed a simple food foraging task which requires memory. An agent controlled by a neural network freely moves in an open environment, a grid of 300×300 . The agent is initially located at its nest and three food are placed in the environment (Figure 1A). The task for the agent is to consume all the food. We impose a memory requirement for the task by limiting the range of the agent's sensors. From the nest, only food#1 is detected. When the agent moves toward food#1, food#2 becomes within its sensor range (gray area marked as ‘prox-

imity' in the figures) and so on (Figure 1B-D). Here, we put additional (and biologically plausible) constraints that the nest is the only place where the agent can consume the food, and the agent has limited food delivery capacity, that it can carry only one food at a time. Therefore, it needs to travel back and forth between the food and the nest at least three times to consume all the food. The life span of the agent is limited to 80 steps and whenever the agent successfully consume a food, it is increased by 100 steps. Memory is necessary to be successful in this task because the agent needs to remember seeing other food(s) while carrying the one found earlier to the nest. Moreover, the nest does not generate any sensory cue at all, thus the nest location also needs to be remembered.

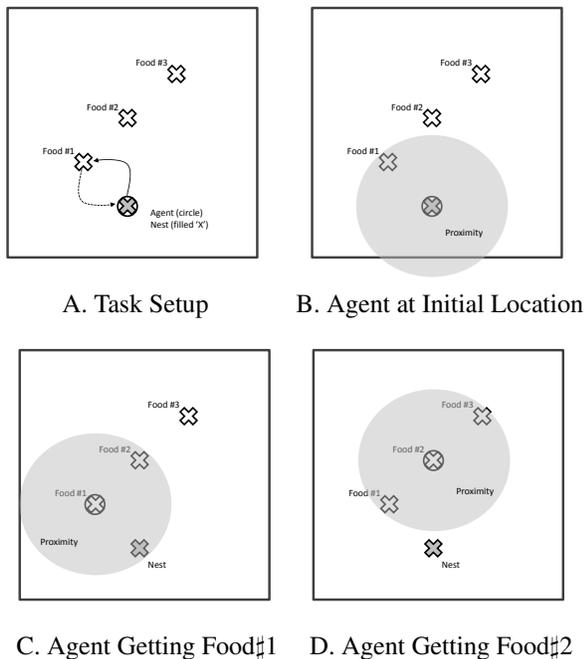


Figure 1: Task Definition.

2.2. Chemical Dropper Agent

The agent model is shown in Figure 2. The agent interacts with the environment with 8 pairs of sensors distributed uniformly around the center at an interval of 45° . 8 of them are chemoreceptors sensitive to the chemical markers the agent drops. The other 8 can be whatever distance sensors (photoreceptors or another chemoreceptors) sensitive to the food. Both types of sensors are limited in range subject to the constraint mentioned in the above task description, and the magnitude of the sensor signal is inversely proportional to the distance to the detected object. The agent can freely navigate within the environment while bouncing off when it

comes into contact with the wall.

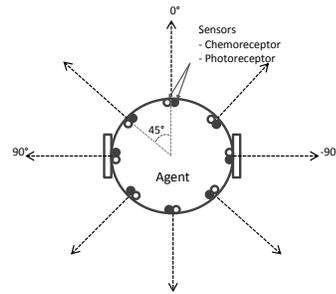


Figure 2: Agent Model.

The underlying neural network of the agent is similar to the Dropper Network in [6], which is simply a feed forward network with an output unit to trigger marker dropping. The equations for this baseline agent take the same standard form as equation (1) in [6] (weighted sum followed by a sigmoid [8]).

Default movements are blocked by removing the bias units to maximize the memory requirement of the task, and one additional output unit is included to a typical feed-forward network to allow chemical dropping behavior (Figure 3). Below is how this additional output works:

$$DropChemical = \begin{cases} \text{True} & \text{if } O_4 > \theta \\ \text{False} & \text{otherwise} \end{cases} \quad (1)$$

If the value of O_4 is greater than the threshold θ , the agent drops a marker in its current position before it moves. This threshold is not fixed but is also evolved through genetic search. The outputs O_1 , O_2 , and O_3 decide x-axis (-3.0 to 3.0), y-axis (-3.0 to 3.0), and angular offset (-180° to 180°) relative to the current orientation of the agent, respectively.

3. Experiments and Results

The learning of connection weights of the agents was achieved through genetic search. The fitness for an agent was set inversely proportional to the sum of the distances between itself and each remaining food, and between each remaining food to the nest. Agents were divided into two groups (NE and E), one using non-evaporating chemical markers (NE) and the other using evaporating markers with the evaporation rate of 0.99 per step (E). The task was repeated 30 times for each group. The population size of each trial was 100 and 10% top-performers of the population survived to the next generation. One-point crossover with probability 0.9 and mutation with rate 0.04 were applied to these top-performing agents. If the number of generation reached the preset maximum ($=300$) without consuming all three food, the trial was recorded as a failure. Quantitative analysis of the results is plotted in Figure 4.

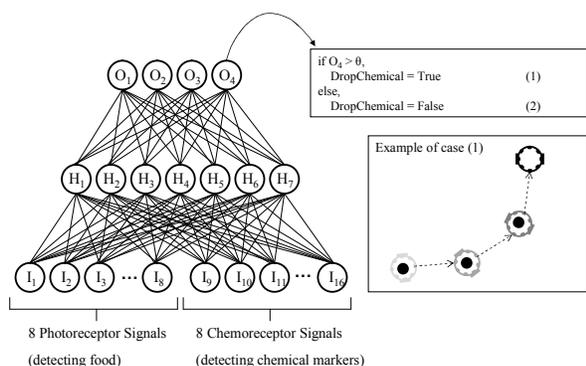


Figure 3: Neural Network of the Agent. An output unit (O_4) is added to the typical feed-forward network to allow chemical marker dropping behavior. The box shows an example of leaving chemical markers when the agent moves to a next location.

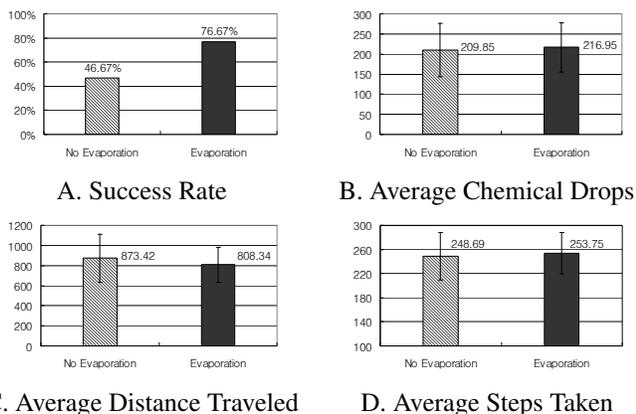
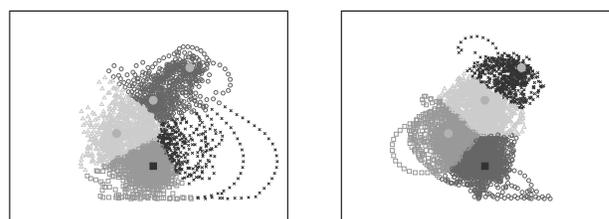


Figure 4: Quantitative Data of successful agents from 'No Evaporation'(NE) and 'Evaporation'(E) groups ($N_{NE} = 14$, $N_E = 23$). A. Large difference in success rate ($NE = 14/30$, $E = 23/30$). B. No significant difference in the number of chemical drops between NE and E ($NE_{stdev} = 66.25$, $E_{stdev} = 61.41$). C&D. Agents in NE group traveled longer than those in E group, while the number of steps taken by NE group are shorter. This suggests that the speed of the agents in E group were faster (C: $NE_{stdev} = 239.43$, $E_{stdev} = 174.02$, D: $NE_{stdev} = 39.65$, $E_{stdev} = 34.34$).

Both the groups show fairly high success rates (Figure 4A). We have also performed the task with feed-forward or recurrent neural network agents but none of them was successful because of the high memory requirement of the task. These results are not reported here because they are trivial. Interestingly, the success rate of the group using evaporating chemical markers was significantly higher. The average chemical dropping rates (drops per steps) are almost the same (0.854 vs. 0.844). The difference in the

number of steps taken by the agents in each group is relatively small (5.97% difference), but the difference in the traveled distance is more significant (7.97% difference). Considering the average speed, the NE group agents moved more than 15% faster than the E group agents ($speed_{NE} = 3.5294$, $speed_E = 3.0647$). Investigation on the reasons for this difference will give clues to understand the difference in the success rate. To make qualitative analysis for further investigation, we plotted the stack of chemicals dropped by all the successful agents in each group ($N_{NE} = 14$, $N_E = 23$) and applied k-means clustering to them (Figure 5).



A. Non-Evaporating Chemical B. Evaporating Chemical

Figure 5: Cluster of Chemical Markers. K-means clustering ($k=4$) is applied to the chemical markers to show the relationship between the markers and points necessary to remember for the task (food&nest).

Clusters of the chemicals dropped by the NE group do not match the food and nest locations, whereas the E group dropped chemicals in correspondence with them. This suggests that evaporating chemicals are used more correctly for the reference of the spatial information critical for solving the task. To determine the difference in the chemical dropping strategy, we performed a detailed comparison between the best agents in each group. We picked the agents with the shortest step size in both groups as the best ones because they must have the most economical strategies among their group members in solving identical tasks.

Figure 6-A describes the strategy of the best agent using non-evaporating markers. Foraging of each food is separated by rows. Each row shows (1) agent trajectory, (2) chemical trail dropped while moving toward food, and (3) chemical trail dropped after the agent picks up the food until reaching the nest. This agent demonstrates an interesting rules developed which can be summarized as follows:

1. Access food on its left side.
2. Follow chemicals on its left side.
3. If a food is detected, the movement is "food-driven".
4. Otherwise, the movement is "chemical-driven".
5. Release chemicals when its movement is "food-driven" but rarely in "chemical-driven" movements.

When foraging the food #1, rules 1, 3, and 5 apply to (2), and rules 2, 4, and 5 apply to (3). Note that the agent only drops a single chemical marker on its way back to the nest. For the food #2, rules 1, 2, 3, 4, and 5 apply to (2), and rules 2, 4, and 5 apply to (3). Again, only several chemical markers were released while the agent is coming back. The same rules applied for the food #2 are also used for food #3.

This agent could not use the simple chemical-following strategy because of the excessive past information about the food already consumed. Therefore, the agent had to develop strange rules such as 1 and 2 and it generated circular trace with increasing radius. The outermost chemical trails in Figure 5 (lower right part) clearly shows this. This agent traveling in circles needed to move faster to successfully gather all 3 food within the limited life span.

Figure 6-B illustrates the strategy of the best agent using evaporating markers, which has less complex rules:

1. Throw chemicals most of the time.
2. If detected chemical signal is strong enough, the movement is “chemical-driven”.
3. Otherwise, the movement is “food-driven”.

All three rules are generally applied all the time. Because this agent is precisely following the food and chemical signals without any ad hoc constraints, it shows finer foraging trajectories by using a single route to and from the food.

This finer trails are due to the evaporating property of the chemical markers. Because the chemicals contain temporal information, this agent does not have to develop a special rule to avoid conflicts among the chemical markers. The resulting trajectory looks more efficient because the agent does not visit all the previous locations of the food that have already been consumed. The toward-food trajectory for food #2 looks very interesting. Initially, the agent makes a vain effort to visit the location of food #1 even though it has been consumed just before. Food #2 is once detected on the way, but ignored because the attractions of the chemicals are greater than that of food #1. After this futile effort, the agent makes zigzagging movements around the nest. Internally for the agent, this behavior was perhaps because of the conflicting chemicals near the nest and this conflict is resolved as the chemicals evaporate. However, for an external observer, this behavior can be interpreted as highly intelligent. It looks as if the agent, after an unsuccessful outing, sits in the nest area both to mark the nest strongly with accumulating chemicals and to wait for the failed chemical trail to evaporate. As a result, the agent reinitiates its journey following the weakened chemical trail. This time, the attraction of food #2 becomes greater and the agent can successfully maneuver its way toward the food.

4. Discussion

The results show an interesting role chemical sense could have played in early animals with a limited reactive nervous system. In addition to the credits from the evidence from other research domains indicating that the highly probable relation between olfactory and memory systems (see the Discussion section of our earlier paper [6]), the present results provide even stronger suggestion that the chemical sense can indeed help evolve highly intelligent behavior in a biologically plausible task necessary for survival.

The use of pheromone and its volatile property has gained a great degree of interest [5, 17, 7]. Ant Colony Optimization (ACO), a metaheuristic used for various optimization problems, may be the most cited nature-inspired application employing evaporating pheromone [7]. However, just as most other investigations on animals or artificial agents using pheromone, the ACO uses volatile pheromone as a medium of communication and is limited to the context of social learning. We perceive that this fading-out characteristic of chemical markers can be an external analog of *forgetting* in internal memory (cf. [2] on the recency effect). Our work stands out from these earlier works in that it shows that the chemical markers could have been used in a self-serving manner assisting to develop intelligence critical for individual survival, prior to their use in a social context.

Our results do not only provide viable hints on the evolution of internalized memory, but also suggest operable option to the behavior-based robotics field. Unlike [1] or [3] which required *indexed* representations of the environment, the strategies in our approach are learned automatically from evolution without an internal representation. To think reversely, the interaction between agent and external marker can be regarded as a distributed cognition minimizing the cognitive complexity of the agent. This way, our results may shed light on scaling up the ability of behavior-based systems to solve more complex problems, and that in a highly biologically inspired way (cf. [11]).

We would also like to mention the relationship between external marker use and general tool use. Considering the wide variety in animal species, only few vertebrates make a monopoly of tool use [4, 15, 10]. The use of external material may have laid the road to a more complex use of external matter, i.e., tool use.

5. Conclusion

The main contribution of our work is to have shown experimentally the use of external chemical markers enables reactive agents to perform successfully in a biologically inspired memory task. Our results reinforce our hypothesis that memory may have its evolutionary origin in the chemical sense and that an important step toward the evolution of

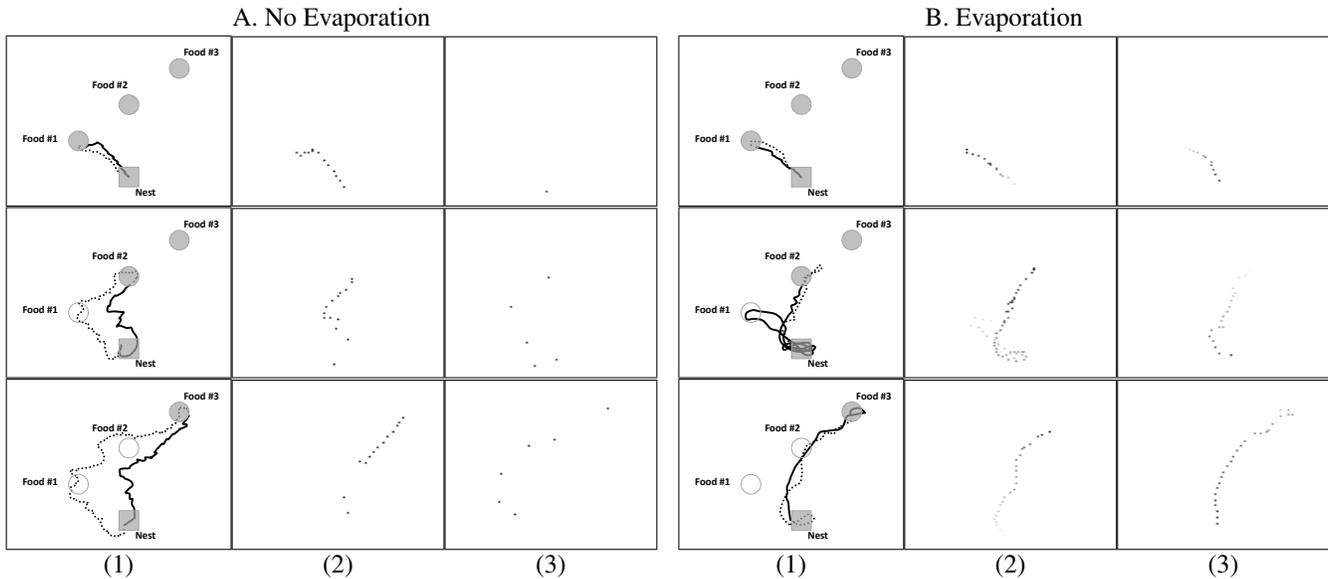


Figure 6: Chemical Trails. A&B(1) Agent trajectories while foraging food in the environment. Solid/dashed lines denote the nest-to-food/food-to-nest trajectories, respectively. Foraged food are marked as blank circles. (2)&(3) Chemical trails of the agent dropping/detecting chemicals. A(2)-(3) The agent drops chemicals when its movement is driven by food but it seldom does so when it is following the chemical trail to come back to the nest. Because the chemical markers do not evaporate, all the markers have the same sensitivity to the agent (shown by the same grayscale value). B(2)-(3) Agent drops chemicals almost always. As the chemicals evaporate, their sensitivity fades (fading color in figures).

cognition and intelligence could have been the use of external chemical markers. Future investigations will be focused on the topological modification of the neural network controller to internalize the behavioral changes driven by the reactive material interaction.

References

- [1] P. E. Agre and D. Chapman. Pengi: An implementation of a theory of activity. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)*, pages 268–272, Seattle, WA, 1987.
- [2] A. D. Baddeley and G. J. Hitch. The recency effect: Implicit learning with explicit retrieval? *Memory and Cognition*, 21:146–155, 1993.
- [3] T. Balch and R. Arkin. Avoiding the past: a simple but effective strategy for reactive navigation. In *Proceedings of the 1993 IEEE International Conference on Robotics and Automation*, pages 678–685. IEEE, 1993.
- [4] B. Beck. *Animal Tool Behavior: The Use and Manufacture of Tools by Animals*. Garland, 1980.
- [5] S. Chandrasekharan and T. Stewart. Reactive agents learn to add epistemic structures to the world. In K. D. Forbus, D. Gentner, and T. Regier, editors, *CogSci2004*, Hillsdale, NJ, 2004. Lawrence Erlbaum.
- [6] J. R. Chung, J. Kwon, and Y. Choe. Evolution of recollection and prediction in neural networks. In *Proceedings of the International Joint Conference on Neural Networks*, pages 571–577, Piscataway, NJ, 2009. IEEE Press.
- [7] M. Dorigo and C. Blum. Ant colony optimization theory: A survey. *Theoretical Computer Science*, 344(2-3):243–278, 2005.
- [8] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley, 1991.
- [9] R. B. Huey, P. E. Hertz, and B. Sinervo. Behavioral drive versus behavioral inertia in evolution: A null model approach. *American Naturalist*, 161:357–366, 2003.
- [10] M. Krützen, J. Mann, M. R. Helthaus, R. C. Connor, L. Bejder, and W. B. Sherwin. Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences, USA*, 102:8939–8943, 2005.
- [11] M. J. Mataric. Behavior-based robotics as a tool for synthesis of artificial behavior and analysis of natural behavior. *Trends in Cognitive Science*, 2(3):82–87, 1998.
- [12] L. M. Rocha. Eigenbehavior and symbols. *Systems Research*, 13:371–384, 1996.
- [13] A. F. Skutch. Do tropical birds rear as many young as they can nourish? *Ibis*, 91:430–455, 1949.
- [14] C. C. Smith and O. J. Reichman. The evolution of food caching by birds and mammals. *Annual Review of Ecology and Systematics*, 15:329–351, 1984.
- [15] S. Tebbich, M. Taborsky, B. Fessl, and M. Dvorak. The ecology of tool-use in the woodpecker finch (*cactospiza pallida*). *Ecology Letters*, 5:656–664, 2002.
- [16] W. T. Wcislo. Behavioral environments and evolutionary change. *Annual Review of Ecology*, 20:137–169, 1989.
- [17] T. Ziemke, N. Bergfeldt, G. Buason, T. Susi, and H. Svensson. Evolving cognitive scaffolding and environment adaptation: a new research direction for evolutionary robotics. *Connection Science*, 16:339–350, 2004.