Embodied Moving-Target Seeking with Prediction and Planning

Noelia Oses[1], Matej Hoffmann[2], and Randal A. Koene[1]

[1] Fundación FATRONIK-Tecnalia,
Mikeletegi pasealekua 7, 20009 Donostia-San Sebastián, Spain
{noses,rkoene}@fatronik.com
[2] Artificial Intelligence Laboratory, Department of Informatics, University of Zurich
Andreasstrasse 15, 8050 Zurich, Switzerland
hoffmann@ifi.uzh.ch

Abstract. We present a bio-inspired control method for moving-target seeking with a mobile robot, which resembles a predator-prey scenario. The motor repertoire of a simulated Khepera robot was restricted to a discrete number of 'gaits'. After an exploration phase, the robot automatically synthesizes a model of its motor repertoire, acquiring a forward model. Two additional components were introduced for the task of catching a prey robot. First, an inverse model to the forward model, which is used to determine the action (gait) needed to reach a desired location. Second, while hunting the prey, a model of the prey's behavior is learned online by the hunter robot. All the models are learned *ab initio*, without assumptions, work in egocentric coordinates, and are probabilistic in nature. Our architecture can be applied to robots with any physical constraints (or embodiment), such as legged robots.

Keywords: bio-inspired control; forward model; inverse model; prediction; planning; egocentric coordinates.

1 Introduction

This paper deals with the problem of moving-target seeking by a mobile robot, a predator-prey scenario. This problem has been long solved in nature, hence we use bio-inspired control methods to approach it. In order to approximate the real-world conditions we use a Khepera robot model with specific physical constraints. We define a set of 10 gaits, each gait being a pair of velocities for left and right motors. This restricted repertoire of gaits helps us to approximate the context of animal behavior (our final goal is to address more complex platforms such as legged robots).

We implement a forward model, which enables the robot to learn to predict how a set of motor commands from its repertoire will influence its state in the environment [1,2]. The robot needs to learn its own dynamics model for navigation, in accordance with its limited set of gaits. We achieve this through autonomous exploration inspired by the motor-babbling observed in infants [3].

E.S. Corchado Rodriguez et al. (Eds.): HAIS 2010, Part II, LNAI 6077, pp. 478–485, 2010.

The inverse model of the forward model is used to determine the gait needed to reach a specific location in one time-step, such as the expected relative location of the prey. If a single time-step does not suffice then a sequence of gaits is planned. The number of possible combinations of gaits increases exponentially with the length of the sequence, so that efficient heuristics are needed. Finally, the hunter learns a model of the prey's behavior online and without any prior knowledge or assumptions. This prey model is used to predict future prey locations. All models operate in egocentric (robot-centered) coordinates, without any assumptions on the action space, and incorporate uncertainty.

The combination of dynamics and uncertainty (in the robot's forward model and in prediction of prey behavior) provide a useful approximation of real-world conditions. In these conditions extensive planning is unfeasible, because algorithms need to operate in real-time and a deep plan would need to be updated too rapidly to be useful (the frame problem [4]). Instead, we begin with a bottom-up approach: Find a solution that is as reactive as possible; then add lookahead prediction and planning that is required to catch the prey. Planning is therefore added only to the extent that the combination outperforms a simple reactive architecture.

2 Learning a Forward Model in an Egocentric Coordinate System

We use a relative reference system in polar coordinates centered in the hunter robot's center of mass (Fig. 1a). Angle is measured clockwise from the robot's posteroanterior vector (PA), i.e. the hunter's heading is zero degrees. Location and heading constitute a robot's pose. For the forward model, the hunter's reference system at time t is used to express the next pose after one time-step. We indicate the heading of the hunter at time t plus one time-step as the angle that the hunter robot's PA vector subtends measured clockwise with respect to the hunter robot's PA vector at time t (Fig. 1a).

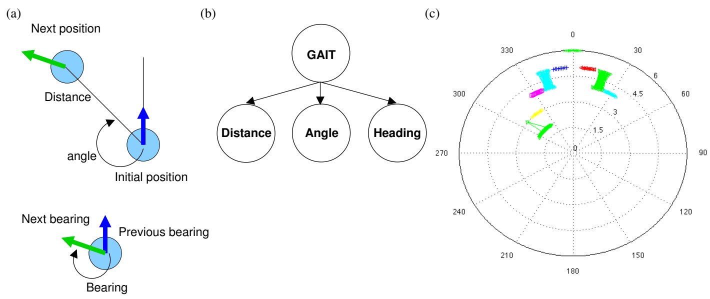

Fig. 1. (a) Egocentric coordinates. (b) Bayesian network for the forward model. (c) Plot showing the outcome of applying different gaits.

The robot needs to learn to predict the outcomes of its actions. The forward model enables this. We define the application of a specific gait for a specific amount of time as an action. The consequence of such an action is a new pose of the robot. We implement the forward model as a Bayesian network (BN), as in Demiris and Dearden [5], because BNs provide a powerful probabilistic framework in which to express the causal nature of a robot's control system. A motor command (Gait) and the observations Distance, Angle, and Heading are each represented as random variables in the BN (Fig. 1b). We use a naïve Bayes classifier, which is often quite effective even when the attribute values are not conditionally independent [6,7].

The BN parameters (the conditional probability distributions) are learned offline from data obtained during motor-babbling (randomly applied gaits, see Fig. 1c). The data is complete, the structure of the network is known and the prior probability distribution for the gaits is uniform (gaits were applied randomly with equal probability). Maximum a posteriori (MAP) learning therefore reduces to maximum-likelihood parameter learning.

3 Inverse Model and Prey Model

The inverse of the forward model describes which gait to take in order to achieve a desired location (distance, angle) in one time-step. We can obtain this inverse model, $P(Gait|Distance, Angle)$, through inference from $P(Distance|Gait)$ and $P(Angle|Gait)$, which are provided by the BN of the forward model. We approximate (distance, angle) tuples with the nearest learned polar coordinates encountered during learning.

The hunter learns a probabilistic transition model of the prey's movement online, independent of the models for its own movement. The hunter observes how the prey moves, new prey pose as a function of prey pose one time-step earlier (Fig. 2a). Currently, the hunter robot gets the GPS data corresponding to the location of the prey at each time-step; at time $t+\Delta t$ (Δt being the time-step) transforms that into the prey's egocentric coordinates with respect to the prey's reference system at time t; and incorporates the egocentric coordinates to the prey model. This prey model is used to predict the prey's future positions. This approach resembles Thrun *et al.* [8], except that we make no *a-priori* assumptions about the way in which the prey moves or about its possible actions (unlike [9]). We note the frequency of each pose transition observed in terms of distance, angle and heading. An illustrative plot of a prey transition model can be seen next (Fig. 2b).

4 Models and Experiments

We develop a reactive model, a prey prediction model, and a planning model, and we assess the performance of each with the same experiment, conducted both in a walled-in environment and in an open environment. The experiment has seven initial states. These consist in the prey being located at five bodies' distance from

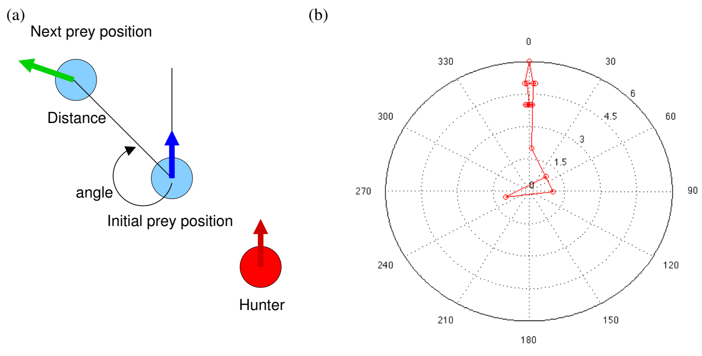

Fig. 2. (a) Illustration of one prey transition. (b) Example transition model of the prey.

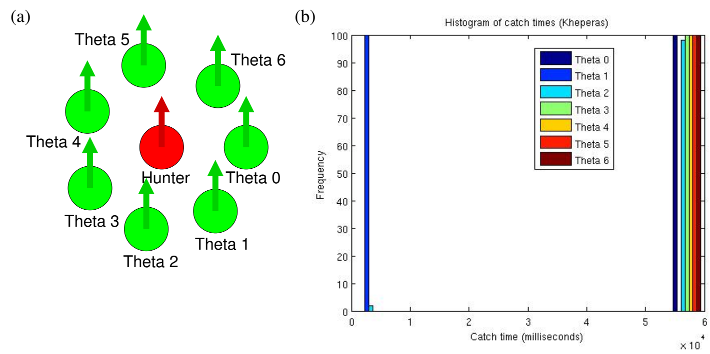

Fig. 3. (a) Experiment set-ups. (b) Results of the reactive model.

the hunter and with the prey at angles $\theta = 0$, 1, 2, 3, 4, 5 and 6 radians, with identical headings for hunter and prey (Fig. 3a). Cyberbotics' Webots'$^{\text{TM}}$ [10] Braitenberg controller runs the prey, so that it moves straight ahead until it senses an obstacle and turns. The hunter performs no obstacle avoidance. The simulated time elapsed until the hunter catches the prey is recorded. Simulations end when the prey is caught or after one simulated minute. An experiment consists of 100 simulations for each initial state.

4.1 Reactive Model

The hunter applies a gait determined by the inverse model in accordance with the current prey position. Resulting reactive behavior only enables the hunter to catch the prey in very concrete circumstances. In general, the hunter appears to follow the prey around (Fig. 3b). Out of 700 runs only 102 were successful (14.57% success rate). When the prey started off at $\theta = 1$ radians the hunter was always successful. The hunter also caught the prey on 2 occasions when the prey started off at $\theta = 2$ radians.

4.2 Prey Prediction Model

The hunter learns the prey model online and uses it to predict the prey's future position (Fig. 4a). At each time-step, the prey's predicted position is used as target position for the inverse model, which determines the hunter's gait. The prey's position can be predicted ahead for a number of time-steps (T), and the optimal number depends on the distance between hunter and prey. We set T to the nearest integer to half of that distance.

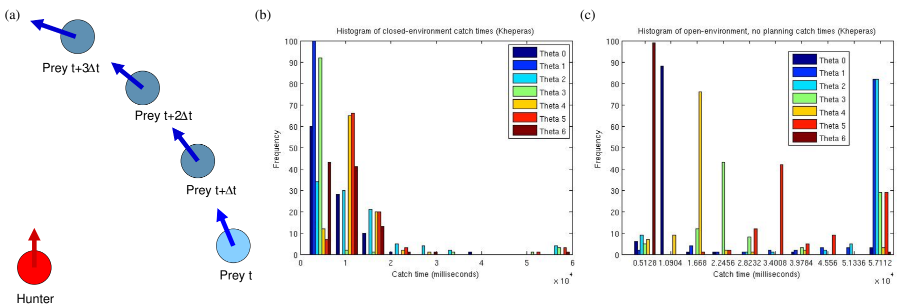

Fig. 4. (a) Prey prediction. (b) Results of the prey prediction model in the closed environment. (c) Results of the prey prediction model in the open environment.

Figure 4b shows results in a walled-in environment. The prey was caught in 655 of 700 runs (92.8% success rate), with average catch time (including misses): 15.287 seconds. Figure 4c shows results in an open environment. The prey was caught in 473 of 700 runs (67.6% success rate). The lower success rate was influenced by the prey controller, as the prey can continue to run straight when nothing forces it to turn.

4.3 Planning Model

For a better success rate in the open environment, the hunter needs to plan more than one gait ahead, composing gaits to catch the prey. We now predict the prey's position at successive time-steps and select the minimum at which a composition of gaits will minimize the distance between the hunter and the prey.

Heuristic Solution with Best-First Search: The theoretical solution would involve calculating the probability distribution for the distance between the hunter and the prey. In doing so we would encounter the "curse of dimensionality" due to the exponential increase in the size of the state space with each level of the search tree (Fig. 5). We can avoid this by using sampling to predict hunter

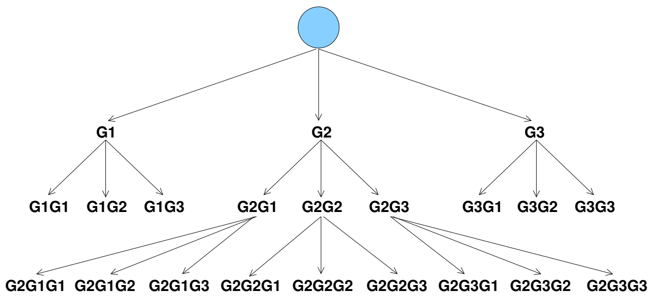

Fig. 5. Search tree for planning sequence of gaits (shown only for 3 gaits for the sake of clarity)

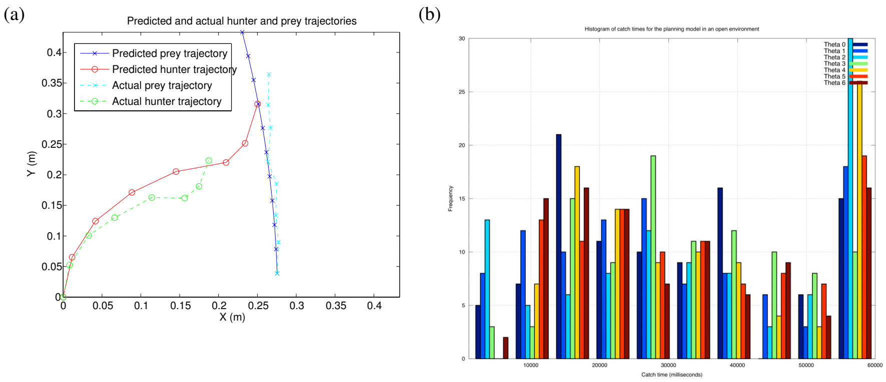

Fig. 6. (a) Example of a planning iteration. (b) Results of the planning model in an open environment.

position. We calculate samples for each time step and each different sequence of gaits. Each node in the tree has associated information: T (time-steps or depth the node plans), Gait[t] (sequence of gaits applied at time-steps $t < T$), cost in terms of number of gaits used in planning, cost in terms of number of gait transitions (transitions will be important in the legged robot scenario), Hunter[t] (predicted hunter coordinates at time-steps $t < T$ relative to hunter pose during planning), Value (final predicted distance between hunter and prey).

Choosing a sequence of gaits is a combinatorial optimization problem. A breadth-first search of the tree was too slow, so we proceeded to use a best-first search. The best-first search algorithm explores a graph by expanding the most promising node. In our case, the most promising node is the one that most reduces the distance to predicted prey position. The search tree with g^T nodes needs to be pruned further, for example by eliminating combinations with more than one gait transition. With the planning model (Fig. 6b), 591 of 700 runs were successful (84.4% success rate).

5 Discussion and Conclusions

We have presented a bio-inspired control architecture that allows a mobile robot to: (1) learn a model of its own action repertoire (a forward model); (2) learn a model of an object (prey) it is seeking; (3) combine the forward model and the prey model to seek the prey.

Braitenberg [11] and Brooks [12] showed that robots that rely on embodiment, purely reactive behaviors and that exploit interaction with the environment could address real-world dynamic problems that representations in classical A.I. could not adequately deal with. Such robots exhibit sophisticated behaviors and properties such as adaptivity, robustness, versatility and agility found in biological organisms, yet without emphasizing cognitive capabilities such as planning, abstract reasoning or language. Following this inspiration, we took a bottom-up approach by developing a reactive model first and only adding cognitive capabilities as and when necessary.

Our architecture has the following properties: (1) An egocentric coordinate system is used; (2) The model can deal with an arbitrary action repertoire of the hunter and the prey. There are no assumptions on the behavior of the hunter or prey; (3) The action space is discrete; (4) The models are learned *ab initio*. The hunter's forward model is is learned as a result of a motor-babbling phase. The prey's model is learned online and incrementally updated; (5) Our model accounts for and plans with uncertainty.

We see two possible uses for our architecture. First, it can be applied as a whole to moving-target seeking by an autonomous vehicle, for instance. Or, only individual components can be utilized. The forward model implementation would allow an arbitrary robot to learn its motor repertoire and plan with it. The prey model can be applied to any target object, such as in a person-following scenario [9]. Second, our scenario could serve to model biology. By adding details about particular behaviors we may test hypotheses for the way in which animals achieve similar behaviors, for example: the prey-catching behavior of the spider Portia [18] or hunting in vertebrates. At the same time, our scenario is a case for minimalistic model of cognition which is firmly grounded in body dynamics [13,14,15,16,17].

Future work planned includes extending our model to a legged platform which uses real gaits, adding real sensing of the prey (through a camera on the hunter, for instance), and studying various cost functions for the trajectory planning of the hunter. These can include energy consumption, or computational complexity/reaction time.

Acknowledgments. The authors would like to thank the AI Lab of the University of Zurich for welcoming the first author to the lab through an internship. We would also like to thank Juan Pablo Carbajal for many interesting and informative discussions about the project. This work was undertaken in the context of the "From locomotion to cognition" project funded by the Swiss National Science Foundation, Grant Nr. 200020-122279/1.

References

1. Webb, B.: Neural mechanisms for prediction: do insects have forward models? Trends in Neurosciences 27, 278–282 (2004)
2. Wolpert, D.M., Miall, R.C., Kawato, M.: Internal models in the cerebellum. Trends in Cognitive Sciences 2, 338–347 (1998)
3. Meltzoff, A.N., Moore, M.K.: Explaining facial imitation: a theoretical model. Early Development and parenting 6(2), 157, 1–14 (1997)
4. Pfeifer, R., Scheier, C.: Understanding intelligence. The MIT Press, Cambridge (2001)
5. Demiris, Y., Dearden, A.: From motor babbling to hierarchical learning by imitation: a robot developmental pathway. In: Proceedings of the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems. Lund University Cognitive Studies, vol. 123, pp. 31–37 (2005)
6. Mitchell, T.M.: Machine learning. McGraw-Hill, New York (1997)
7. Russel, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs (1995)
8. Thrun, S., Burgard, W., Fox, D.: Probabilistic robotics. The MIT Press, Cambridge (2005)
9. Vazquez, A.: Incremental learning for motion prediction of pedestrians and vehicles. PhD thesis, Institut National Polytechnique de Grenoble (2007)
10. Michel, O.: Webots: Professional Mobile Robot Simulation. International Journal of Advanced Robotic Systems 1(1), 39–42 (2004)
11. Braitenberg, V.: Vehicles Experiments in Synthetic Psychology. The MIT Press, Cambridge (1986)
12. Brooks, R.A.: Intelligence Without Representation. Artificial Intelligence Journal 47, 139–159 (1991)
13. Pezzulo, G.: Anticipation and Future-Oriented Capabilities in Natural and Artificial Cognition. In: Lungarella, M., Iida, F., Bongard, J.C., Pfeifer, R. (eds.) 50 Years of Aritficial Intelligence. LNCS (LNAI), vol. 4850, pp. 258–271. Springer, Heidelberg (2007)
14. Duro, R.J., Graña, M., de Lope, J.: On the potential contributions of hybrid intelligent approaches to multicomponent robotic system development. Information Sciences (in press, 2010)
15. Clark, A., Grush, R.: Towards a Cognitive Robotics. Adaptive Behavior 7(1), 5–16 (1999)
16. Grush, R.: The emulation theory of representation: motor control, imagery, and perception. Behavioral and Brain Sciences 27, 377–442 (2004)
17. Schomaker, L.: Anticipation in cybernetic systems: A case against mindless anti-representationalism. In: IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands (2004)
18. Tarsitano, M.: Route selection by a jumping spider (Portia labiata) during the locomotory phase of a detour. Animal Behavior 72, 1437–1442 (2006)