Robots and Bananas: Exploring Deliberation in Cognitive Robots*

Jasmin Grosinger, Federico Pecora, Alessandro Saffiotti
Center for Applied Autonomous Systems (AASS)
Örebro University, 70182 Örebro, Sweden
{jngr,fpa,asaffio}@aass.oru.se

Abstract

Under what conditions should a cognitive robot act? How do we define "opportunities" for robot action? How can we characterize their properties? In this position paper, we offer an initial apparatus to formalize opportunities and to frame this discussion.

A Robot-Banana Problem

There is a physical object in our world, a banana, which can be either *fresh*, *ripe*, *overripe* or *rotten*. The banana changes state over time, from the first to the last. Maintaining a *desirable* world state includes that physical objects in the world, the banana, are in states that are desirable. It is desirable that the banana is fresh, ripe, or overripe, while a rotten banana is undesirable. Also, the banana should be eaten before it becomes rotten. Assume there is a mobile robot, capable of bringing a banana to a human for consumption. Assume also that the robot has a model representing the banana's states and how long it takes to transition over them. How should the robot choose, among all possible intermediate states of the banana, when to act? Is it desirable that the robot immediately takes action as soon as there is a banana in the world, no matter what the states of the banana and the user are?

The act of bringing the banana to the user for consumption achieves a desired state from the banana's perspective — but what does this imply in terms of the world state? Not only should the banana be eaten in a favorable state, but the robot should not act intrusively against the user. For example, it would not be appropriate for the robot to force-feed the sleeping user just because the banana will soon be rotten! There are states in the world which are more suitable for taking a specific action than others — they offer *opportunities* for acting. For instance, the robot may offer the banana the next morning for breakfast. Consuming the banana before it is rotten and not intruding are desirable, and thereby contribute to the maintenance of a desirable world state.

How do the desirable states of the banana affect whether we classify a state of the user as being suitable for robot action? When the banana is fresh, it is not necessary to act. This may even result in an undesirable state of the user, therefore an undesirable world state. A ripe banana increases the necessity to act (we can predict that it will eventually become rotten), but we can afford to choose among few, well tailored states that are suitable for taking action. An overripe banana is closer to being rotten. This influences which states we now classify as suitable for taking action: this is a larger set, and potentially less perfectly tailored to acting.

For all the desirable states of the banana in this simple example, the actions of the robot are always the same: bring the banana to the user for consumption. The banana's desirable states do not differ in their influence on *what* the robot should do, rather *when* it should act. If the banana becomes rotten, the user's state has no influence on which context the robot uses as a "trigger" to act. Also, the robot will act differently: instead of bringing the banana to the user, the robot will decide to dispose of it.

In this paper we focus on characterizing the overall problem entailed by the example above. The attentive reader has certainly spotted that our hypothetical robot should be able to perform a wide range of cognitive tasks, which include perceiving, planning and acting. Studies in cognitive architectures, e.g., ACT-R (Anderson et al., 2004), lend support to the argument that diverse cognitive capabilities must be studied jointly. Similarly, the view proposed by Ghallab, Nau, and Traverso (2014) defines the deliberative capabilities that enable a robot to act appropriately. The view put forth in this paper agrees with these holistic perspectives, and is inspired by both cognitive architectures and planning as acting. Our starting point is an extremely simple example — the story of a robot, a banana and a thrifty human who does not like to waste food. The result is a preliminary formulation of opportunities that highlights the multi-faceted nature of this concept, and provides a language to state questions like the ones above.

Formalizing Opportunity

We consider a system $\Sigma = \langle S, U, f \rangle$, where S is a finite set of states, U is a finite set of external inputs (the robot's actions), and $f \subseteq S \times U \times S$ is a state transition relation. We assume discrete time, and that at each time t the system is in one state, denoted by s_t. The f relation models the system's dynamics: $f(s_t, u_t, s_{t+1})$ holds iff Σ can go from state s_t to s_{t+1} when the input u_t is applied. The free-run behavior

*This work was funded by the EC Seventh Framework Programme (FP7/2007-2013) grant agreement no. 288899 Robot-Era.
Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

F of Σ is defined in terms of the set of states that can be reached from s in k steps when applying the null input $\emptyset$:

$$F^0(s) = \{s\}$$
$$F^k(s) = \{s' \in S \mid \exists s'' : f(s, \emptyset, s'') \wedge s' \in F^{k-1}(s'')\}$$

We consider a set $\mathit{Des} \subseteq S$ and a set $\mathit{Undes} \subseteq S$ meant to represent the *desirable* and *undesirable* states in S. For instance, a state in which the banana is rotten is in *Undes*, whereas any state in which the banana is gone is in *Des* (whether because it was eaten, or disposed of, or was never there). For the time being, we assume that *Des* and *Undes* form a partition of S.

We now want to capture the notion that Σ can be brought from some states to other states by applying appropriate actions in the appropriate context. We define an *action scheme* to be any partial function

$$\alpha : \mathcal{P}(S) \rightarrow \mathcal{P}^+(S),$$

where $\mathcal{P}^+(S)$ is the powerset of S minus the empty set. An action scheme α abstracts all details of action: $\alpha(A) = B$ states that there is a way to go from any state in A to some state in B. We denote with $\mathit{dom}(\alpha)$ the domain where α is defined.

We define what it means for α to be *beneficial* in a state s:

$$\mathit{Bnf}(\alpha, s) \text{ iff } \exists A \in \mathit{dom}(\alpha) \text{ s.t. } s \in A \wedge \alpha(A) \subseteq \mathit{Des}$$

For example, the scheme α_{feed}, which delivers a banana to the user, can be applied in any state s where the user is having breakfast and the banana is either ripe of overripe: these conditions characterize $\mathit{dom}(\alpha_{\text{feed}})$. This scheme is beneficial in any such state s, since the resulting states are desirable because the banana has been eaten.

We can extend the notion of being beneficial to take a time horizon k into account:

$$\mathit{Bnf}^k(\alpha, s) \text{ iff } \exists A \in \mathit{dom}(\alpha) \text{ s.t. } s \in A \wedge F^k(\alpha(A)) \subseteq \mathit{Des},$$

where $F^k(X) = \cup_{s \in X} F^k(s)$. Intuitively, a beneficial^k scheme is a way to bring the system (now) to a state that will be desirable after k time steps. One may also define a durative version, where all future states up to k are desirable.

We can use the above apparatus to characterize the different types of opportunities for action discussed in our example. Let t be the current time and k be a finite time horizon. There are at least six properties that determine whether α is an opportunity for acting:

$$\mathit{Opp}_1(\alpha) \text{ iff } s_t \in \mathit{Undes} \wedge \big(\exists s \in F^k(s_t) : \mathit{Bnf}(\alpha, s)\big)$$
$$\mathit{Opp}_2(\alpha) \text{ iff } s_t \in \mathit{Undes} \wedge \big(\forall s \in F^k(s_t) : \mathit{Bnf}(\alpha, s)\big)$$
$$\mathit{Opp}_3(\alpha) \text{ iff } \exists s \in F^k(s_t) : (s \in \mathit{Undes} \wedge \mathit{Bnf}(\alpha, s))$$
$$\mathit{Opp}_4(\alpha) \text{ iff } \forall s \in F^k(s_t) : (s \in \mathit{Undes} \rightarrow \mathit{Bnf}(\alpha, s))$$
$$\mathit{Opp}_5(\alpha) \text{ iff } \big(\exists s \in F^k(s_t) : s \in \mathit{Undes}\big) \wedge \mathit{Bnf}^k(\alpha, s_t)$$
$$\mathit{Opp}_6(\alpha) \text{ iff } \big(\forall s \in F^k(s_t) : s \in \mathit{Undes}\big) \wedge \mathit{Bnf}^k(\alpha, s_t)$$

The first four properties characterize schemes that can be *applied in the future* in response to either a *current* ($\mathit{Opp}_1, \mathit{Opp}_2$) or a *foreseen* ($\mathit{Opp}_3, \mathit{Opp}_4$) undesired situation. The last two characterize schemes that can be *applied now* in order to prevent *future* undesired situations. Note that for $k = 0$ all the above properties collapse to

$$\mathit{Opp}(\alpha) \text{ iff } s_t \in \mathit{Undes} \wedge \mathit{Bnf}(\alpha, s_t),$$

that is, α can be used *now* to resolve a *current* threat.

A few examples help to appreciate the differences among these properties. Consider a system whose free run behavior goes through the sequence of states s_0 (user sleeping, ripe banana); s_1 (user having breakfast, overripe banana); and s_2 (user at work, rotten banana). Let the current state be s_1 and let $k = 1$. Then, scheme α_{feed} is an opportunity of type Opp_6 because, if applied now, it will avoid reaching the undesired state s_2. Scheme α_{dump} (dump the banana into the trash can) is Opp_4 because it can be applied later, once we get in undesired state s_2, and bring the system back to a desired one. Imagine now a GM banana which may take longer to become rotten, i.e., $F^1(s_1)$ includes both s_2 as above and s_2' in which the banana is still overripe. Then, scheme α_{feed} is Opp_5 in s_1, but not Opp_6. Finally, suppose in state s_3 a garbage-bot will stop at the door; then, the scheme α_{ho} (hand-over the banana to the garbage-bot) is Opp_2 in s_2, since we need to wait until the garbage-bot passes by.

Discussion and Outlook

Although preliminary, the above formulation points to several under-addressed issues connected with acting in robotic systems. Characterizing types of opportunities helps to discover and discriminate between qualitatively different contexts in which robot action is called for. Ghallab, Nau, and Traverso (2014) notice that a model capturing motivation still is unexplored, while work by Hawes (2011) identifies the need to investigate the issue of goal management. Both works address problems related the present discussion.

Our ambition is to develop a general framework of opportunity and to investigate how it relates to deliberation. We believe that this will require the introduction of degrees of desirability of states and urgency of goal achievement. These may be used to capture the trade-off among the points of view of different entities and agents — in our example, the user, the banana, and the robot.

Current techniques for planning, acting, context awareness and other abilities that a cognitive robot should possess, are usually ignorant of the *reason* for affecting change. Part of this reason is *opportunity*. We believe that it is important to characterize this formally, if only to discover which existing techniques are applicable in a proactive robot, which have to be adapted, and which are missing entirely.

References

Anderson, J. R.; Bothell, D.; Byrne, M. D.; Douglass, S.; Lebiere, C.; and Qin, Y. 2004. An integrated theory of the mind. *Psychol Rev* 111(4):1036–1060.

Ghallab, M.; Nau, D.; and Traverso, P. 2014. The actor's view of automated planning and acting: A position paper. *Artif Intell* 208:1–17.

Hawes, N. 2011. A survey of motivation frameworks for intelligent systems. *Artif Intell* 175(5):1020–1036.