A Framework for Exogenous and Endogenous Reflexive Behavior in Virtual Characters

Ulysses Bernardet, Steve DiPaola

iVizLab, Simon Fraser University

{ubernard, sdipaola}@sfu.ca

Abstract. Plausible virtual characters should not only be capable of planful, task oriented behavior, but also exhibit reflexive behavior. We present ongoing work on the development of an architecture for virtual characters that should serve as a base layer, equipping them with a minimal set of plausible reactive behavior. The architecture conceptualizes reflexive behavior as a response to external (exogenous) and internal (endogenous) stimuli.

We firstly present a model of exogenous behavior that defines the causal relationship between processing stages of attention, action selection, and action execution, and secondly a model for the endogenous behaviors of self-touching and posture switching. Lastly, we present the implementation of an "artificial attention module" serves as placeholder for an actual attention component, and that allows the bootstrapping of the development of the behavior regulation component of the model.

Keywords: Reactive behavior · reflexive behavior · cognitive architecture idle · attention · virtual character

1 Introduction

Humans do exhibit volitional, conscious behavior, but a great proportion of all behavior can be classified as non-conscious, autonomic. In this manuscript we present ongoing work on the development of a "pre-cognitive" behavior regulation framework, meaning we are mainly interested in the behaviors that lie "below" cognition (Fig. 1). In the domain of artificial intelligence this behavior without explicit planning is referred to by some authors as "reactive" [1] or reflexive. Yet, reflexive should not be taken as meaning not useful. On the contrary, reflexive behaviors are goal oriented in the sense that they fulfill a function such as protecting the organism from harm. We can, though, distinguish them from higher cognition task oriented behavior where with an arbitrary task content (as general intelligence is defined by some). From the perspective of behavior execution these reflexive behaviors are characterized by the fact that they are largely feed-forward, meaning that there is no continuous control during the execution, and that they are difficult to interrupt once initiated.

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-21996-7_15

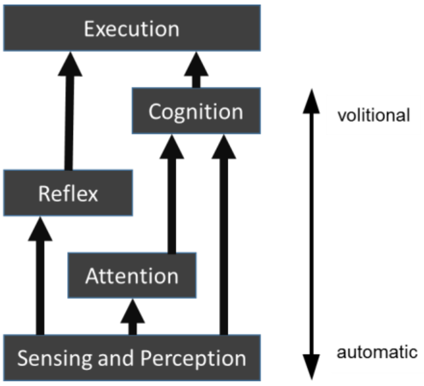

Fig. 1: Proposed hierarchy from sensing to action.

The key distinction we will be making is the one between reflexive behavior that has external (exogenous) and internal (endogenous) motivators. For exogenous behavior (reflexive and cognitive), attention as a mechanism that filters and prioritizes stimuli perceived by an organism, plays a key role. Correspondingly, a number of attention models have been proposed for robots (e.g. [2]) and virtual characters (e.g. [3]). Interestingly, most of these models do not elaborate on the behavioral consequences of the attention process. Endogenous mechanisms, at the behavioral level, partially overlap with what is referred to as "idle" behavior in the domain of virtual characters.
In our model we conceptualize both, endogenous and exogenous actions, as a form of reflexive behavior, and aim to provide a mechanistic model that is grounded in psychodynamic processes. The development of the model is motivated by the observation that there are few integrated frameworks that bring these two types of reflexive behaviors together. Ultimately, our goal is to provide a reusable architecture that can be used as a base layer of behavior in e.g. conversational agents.

1.1 Organization of reflexive behavior

As a working hypothesis we propose the following organization of reflexive behavior that divides distinguishes behavior based on the source of the stimulus that triggers it.

- *Exogenous*: driven by external stimuli such as auditory, visual, and haptic (not all human sensory modalities made sense for a virtual character)
  - Orienting behavior (overt attention towards a stimulus)
  - Aversive behavior
  - Protective behavior
- *Endogenous*: driven by internal stimuli such as somatic proprioception
  - Self-touching
  - Posture (weight shifting, discomfort)

The hierarchy presented here should allow us to make *predictive* inferences, i.e. do foresee **w**hat behavior a given stimulus (e.g. a bright light) will cause. Complementarily, *diagnostic* inferences can be made that determine the cause of the behavior of a given effector (e.g. gaze).

At the behavior level we can distinguish dimension such as much dynamic the behavior is (execution of an action sequence, or a configuration e.g. a static posture), and the “extent” i.e. which any how many body parts are affected (full body, arms, head, eyes etc.).

2 Model of exogenous reflexive behavior

We define exogenous reflexive behaviors these that are a direct consequence of an event external to the agent. Functionally the reflexive behavior mostly subserves the avoidance of harm, and hence is generally more concerned with negative stimuli than with appetitive ones. The reflexive behavior will in many cases be accompanied with a brief, autonomic expression of affect.

Given the context of virtual characters, we will focus on vision and audition, and set aside other modalities such as touch, smell, and taste that are relevant in biological systems. We will begin with a qualitative description of what the behaviors are, and subsequently present a model that relates behavior to triggers in a quantitative fashion.

Orienting behavior is the reaction to a salient stimulus in the environment. The overt action is an orienting towards the stimulus, and hence takes as input the saliency, valence, and location at which the stimulus occurred in the environment. For both, the visual and auditory modality, orienting comprises a turning of head and body towards the stimulus. Additionally, we are likely to observe an expression of surprise.

Aversive behavior is triggered by a negative, potentially painful, sensory stimulation above a given threshold. In the visual domain the actions associate with aversive behavior are a shutting of the eyes and a turning away of the head from the direction where the stimulus occurred. The latter is equally the response in the auditory domain. At the affective level we expect the experience and expression of irritation, disgust, and/or pain.

Protective behavior serves to protect the agent from the negative effect of t a stimulus. For a visual stimulus the protective action constitutes shielding the eyes with the hand, and of an auditory stimulus the protection of the ears. Protective actions are very likely associate with the expression of pain.

Evasive behavior is triggered by a prolonged exposure to aversive stimuli, and comprises a more complex set of action that result in the agent effectively evading the stimulus.

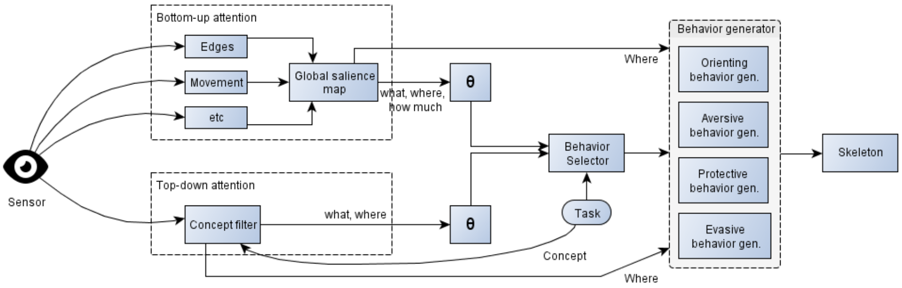

Fig 2: Integrated qualitative model of exogenous reflexive behavior

The proposed integrated qualitative model for exogenous reflexive behavior (Fig 2) comprises three main components for attention, behavior selection, and behavior execution. The *attention* component in turn is composed of a bottom-up and top-down processes. In the bottom-up steam a finite number of "generic" filters is applied to input. Generic meaning that they are primitive (edges, speed, color contrast). These filters are always active, and (largely) not influenced by the task at hand and knowledge about context. Conversely, in the top-down steam, an arbitrary concept can be "content" of the filter (e.g. face, hands, whole body, house, car etc.). These filters are not always active but activated depending on the task, expectation, and context. The integrated model of visual attention postulates that there are several bottom-up filters that are integrated into a single global saliency map. Parallel to this "concept" filters are activated. Both streams are outputting coordinates and some form of identification. The attention module provides to the *behavior selection* component information about the stimuli in the environment such as their location and valence. Based duration on the characteristic and the time course of the stimulus, the behavior selection module will trigger one of the aforementioned behaviors. The role of the behavior generator modules is the actual execution of the orienting, aversive, protective, and evasive behavior. For that purpose the generator need to receive information about location of the stimulus in the environment.

While the qualitative model shows the causal relationships, the complementary quantitative model specifies how stimulus strength and duration are related to the ensuing behavior (Table 1).

Table 1. Proposed quantitative relationship between stimulus duration and amplitude on the one side, and ensuing behavior on the other side.

		Amplitude / strength of stimulus		
		Weak	Medium	Strong
Duration of stimulus	Short (<= second)	Orienting	Weak aversive reaction	Aversive behavior
	Medium (seconds)	Orienting followed by habituation	Aversive behavior	Protective behavior
	Long (minutes)	Orienting followed by habituation	Irritation, evasive action	Evasive actions

3 Model of endogenous reflexive behavior

With the term endogenous behaviors we denote behaviors that are driven by internal e.g. proprioceptive, signals. This class of behaviors comprises self-touching (e.g. scratching) and posture. Both of these classes of behaviors do have functions that go beyond mere reflexive action e.g. in communication. We do, however, explicitly not include these factors in the architecture presented here.

3.1 Self-touching behavior

Self-touching is the touching of one's hand to one's face or body to scratch, rub, groom, or caress it [4]. Self-touching behavior can be motivated externally, or arise from internal motivations such as psychological discomfort, or as a displacement activity. Our functional view of self-touching assumes that there are specific triggers and associated action that are being performed. We identify four core parameters of self-touching: 1) the location at which the manipulation is performed (e.g. head region, torso, forearm, etc.), 2) the action that is executed (rubbing, scratching, stroking, etc.), 3) frequency of the manipulation, 4) duration of the behavior. We assume that both, location and frequency, will be the under control of “somatic” cues, affective state, implicit communicative signaling, and habitual factors. We will elaborate on two mechanism for generating self-touching behavior.

Model 1 We generate one separate Poisson process for each combination of location and action (e.g. scratch-head, rub-hand, caress-forearm). To avoid overlapping of actions, and ensure a minimal waiting time between behaviors we will need to add a lateral inhibition, and refractory period to the overall system. Parameters of this model are the lambda for each Poisson process, and the minimal interval between actions. One downside of this approach is that we need to specify a large number of combinations.

Model 2: In this model we use a single Poisson process. At each event of the process, we select a location and an action from predefined probability tables. The probability tables specify the individual probabilities for each type of action, and location, with the sum of probabilities for each of the tables being one. This model has the advantage that mutual exclusivity of action is guaranteed. It should be noted that as statistical models, proposed mechanism are under specified in terms of the exact external or internal motivators of the behavior, e.g. how the frequency is specified. Additionally, above models assume a logical independence of action and location, which might not be entirely plausible, since humans might have a tendency to e.g. scratch more their chin, and rub more their forearms.

3.2 Posture and posture switching

Functionally, posture refers to the stabilizing action of the musculoskeletal system that allows to maintain a stable pose such as standing or sitting. A crucial variable of the stabilization of posture is to hold the body's center of gravity steady [5]. Next to this functional aspect, human posture can provide a significant amount of cues about emotional state, and serves as information channel in nonverbal communication.

In the model presented here we are primarily interested in the somatic aspect of posture, and more specifically the motivation for switching from one posture to another. The actual posture of the virtual character is either based on predefined joint-angle lists, or generated in real-time.

In our model the switching of postures is a function of the somatic cue of fatigue caused by exertion. As an example the agent will shift the weight from one leg to the other based on how long he/she had the weight on one leg. In the present case we mainly focus on the lower body apparatus, but the model could equally be applied to other body parts. To calculate we apply a simple heuristic, as opposed to a realistic computation of actual strain on the joint [6].

The current posture is constantly analyzed in terms of the strain that it produces on the target muscles. This strain is in turn integrated for each of the two body sides (weight integrator left right/left leg). Once the integrator has reached threshold θ, a new posture is selected from a repository of postures, and send to the skeleton as list of joint-angles. Simultaneously to this, the weight integrators are reset. This reset can also triggered by any other movement that are performed by the body, such as walking.

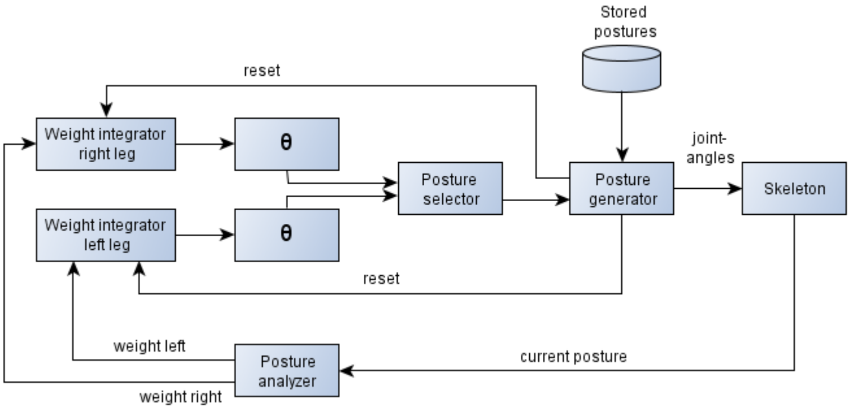

Fig. 3 Model for poster shifting based on fatigue.

4 Implementation

4.1 Artificial attention module

To bootstrap the development of the reflexive behavior architecture, we implement an artificial attention module that allow to feed into the behavior regulation system what would normally be sent as the output of an attention module (AAM).

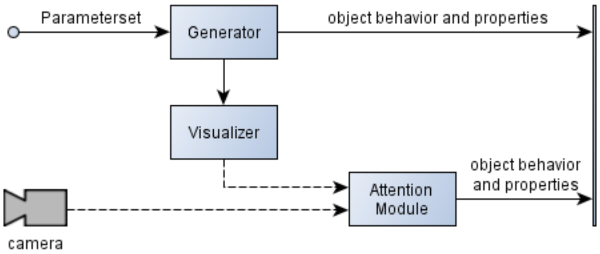

Fig. 4: Implementation of an "artificial attention" module.

The implementation of the AAM (Fig. **4**) takes four parameters (velocity, distance increment, valence, and saliency), and produces two outputs: 1) A temporal sequence describing the spatial behavior of an object with Brownian motion, 2) a graphical rendering of the output (valence coded as hue of the blob, saliency as value in HSV color

space). The graphical output allows for a direct visual inspection, and, most importantly, will in a later development stage allow to feed the graphical output through a module that implements an actual attention module.
The AAM module allows to generate a set of canonical stimuli that the model should be able to sensibly deal with. Give some descriptions of parameter sets and the expected behavior of the model. Examples of such sets are: A harmless (neutral valence) object that approaches; harmful object (negative valence) object that recedes; a fast-moving positively valenced object at a constant distance. The AAM hence provides a shortcut, not only in terms of a need to have an attention module, but also by providing a highly parameterized animations tool, that bypasses the need to videograph a library of stimuli. The AAM is implemented in Python, with the rendering done using the Python OpenCV binding. The AAM is sending to temporal sequence of blob behaviors to the YARP middleware. From here the information can be integrate e.g. with the hybrid discrete-continuous control system implemented using MathWorks Simulink and Stateflow (mathworks.com/products/stateflow) as described in [7].

5 Summary and Outlook

In this manuscript we present an organization of pre-cognitive, reflexive behavior, distinguishing between exogenous behavior motivated by external, and endogenous behavior triggered by internal stimuli. Ultimately, the proposed architecture should serve as a behavioral base layer for virtual characters, equipping them with a minimal set of plausible reactive behaviors. Future steps are the implementation of the architecture in Matlab and Simulink, model intrinsic testing e.g. for stability, and ultimately the testing of the generate behavior in an interactive scenario.

Acknowledgments. This work was partially supported by "Moving Stories" and "Moving+Meaning" Canadian SSHRC and CANARIE grants respectively.

6 References

1. Ferber, J.: Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence. (1999).
2. Ruesch, J., Lopes, M., Bernardino, A., Hornstein, J., Santos-Victor, J., Pfeifer, R.: Multimodal saliency-based bottom-up attention a framework for the humanoid robot iCub. 2008 IEEE International Conference on Robotics and Automation. pp. 962–967. IEEE (2008).
3. Hillaire, S., Lécuyer, A., Regia-Corte, T., Cozot, R., Royan, J., Breton, G.: A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments. Proc. 17th ACM Symp. Virtual Real. Softw. Technol. 191–198 (2010).
4. Heaven, L., McBrayer, D.: External motivators of self-touching behavior. Percept. Mot. Skills. 90, 338–342 (2000).
5. Horstmann, G. a, Dietz, V.: A basic posture control mechanism: the stabilization of the centre of gravity. Electroencephalogr. Clin. Neurophysiol. 76, 165–176 (1990).
6. Rodríguez, I.: A Joint-level Model of Fatigue for the Postural Control of Virtual Humans. Proc. of the 5th Int. Conference “Human and Computer” HC02 (2002).

7. Saberi, M., Bernardet, U., DiPaola, S.: An Architecture for Personality-based, Nonverbal Behavior in Affective Virtual Humanoid Character. Procedia Comput. Sci. 41, 204–211 (2014).