Lifelike Interactive Characters with Behavior Trees for Social Territorial Intelligence

SIGGRAPH2012

Claudio Pedica, Hannes Högni Vilhjálmsson
Icelandic Institute for Intelligent Machines and CADIA, Reykjavík University

Introduction

State-of-the-art technology allows for photo-realistic graphics but this is not always enough. The gaming industry is slowly evolving the art of story telling but no matter how compelling the graphics or thrilling a story, awkward character behavior often breaks player immersion. With knowledge of psychology and sociology, we can build lifelike interactive characters that grasp situations, react to contingencies and look aware of their surroundings.

A New Perspective

Body language, personal space, gaze attraction and territoriality, are all part of our daily social life; part of what makes us appear mindful and attentive. However interactive characters are often not showing any of those. Standing in line without rushing into others, facing a group of peers you are chatting with, sensing the comfort zone of another, looking into each other's eyes require a **different approach than what is used to handle tactical combat scenarios.**

Easier — We know how to model it

Harder — We don't know how to model it

Figure 1. Tactical combat is by far the prevalent applicative scenario of character A.I. in today's game production. Clearly, story telling goes beyond mere gunfire and there is a demand for more versatile A.I. Less has been done to model characters in social scenarios where we expect them to appear more lifelike.

In previous seminal work [2] we showed how the social theories of human territoriality [3] and face-to-face interaction [1] can serve as a solid base to model reactions expected by users when interacting with virtual characters.

From Space to Context

Path-finding and local avoidance are established techniques to simulate character space understanding. However, to achieve lifelike character behavior, **the understanding of space is insufficient without an understanding of context.** Certain locations may have an intended use that goes beyond their mere topological shape. **Space has affordance.** An easy solution could be to apply semantic tags to mark those locations but what if they change dynamically? Theories on human territoriality can help building a system that reasons about space affordance.

F-formation

Individuals in conversation tend to arrange in such a way to have equal, direct and exclusive access to a common shared space. **This positional and orientational arrangement is called an F-formation** [1] and the set of behavioral relationships among the participants defines a behavioral system called the F-formation system. This system is part of the participant's interactional context.

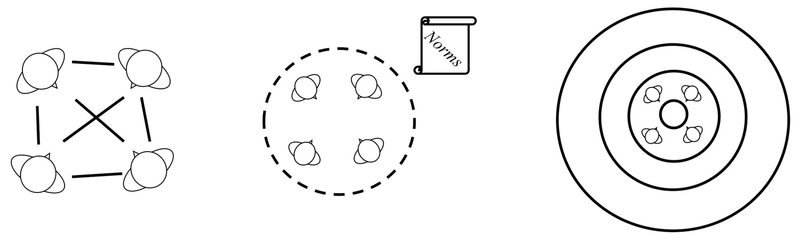


Figure 2. A schematic representation of the three main concepts of a conversation as an instance of F-formation. First, the participants behavioral relationship responsible to sustain a dynamic formation. Second, a conversation as a unit of interaction with emerging norms silently accepted by the participants. Third, norms and behavior relationship contribute to the organization of a human territory where space has structure.

Human Territories

Conversations are just one type of face-to-face interaction but there are many more. **Human territoriality generalizes the F-formation theory and sees conversations as part of a larger hierarchy of human territories** [3]. Within a human territory the space is organized in regions of intended meaning and usage. Such a structure of space is dynamic and sustained by territorial behaviors.

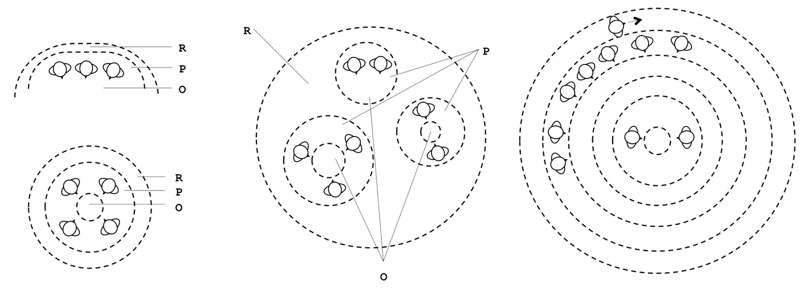


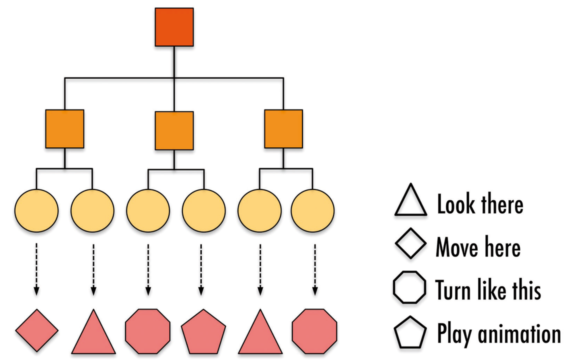
Figure 3. A classification of human territories. From left to right: element, F-formation, gathering and hub. Each of them is an increment of the previous in terms of amount of space and complexity. In addition, more complex territory can encapsulate simpler ones like nested Chinese boxes.

Approach

We have now integrated our reactive approach for social territoriality with Behavior Trees (BTs), an emerging game A.I. technique that is fast becoming a standard in the industry. This integration led to a variant of BTs where multiple branches can run simultaneously and blend. A middle-layer of custom-made arbitration strategies performs the blending before actuation, resembling **command fusion architectures**. We also gave behavior nodes a priority. High priority behavior branches can subsume lower priority ones to respond immediately to critical contingencies, akin to **subsumption architectures**.

Multiple Action Requests

Leaf nodes generate action requests. A request demands a certain action such as "look there", "move here", "play an animation", etc., without actually implementing it. Some BTs implementations stop the decision logic after an action has been selected, while in ours **multiple decision branches can run simultaneously** each leading to a different action request.



Action Blending

After generation, action requests are gathered into groups and each group **blended through an arbitration strategy which resolves potential conflicts**. The result is a set of final combined requests forming the attributes of what we call a motivation. A motivation models the psychological drive to react and results in a compound collection of motion requirements to be issued to the actuation layer for action execution.

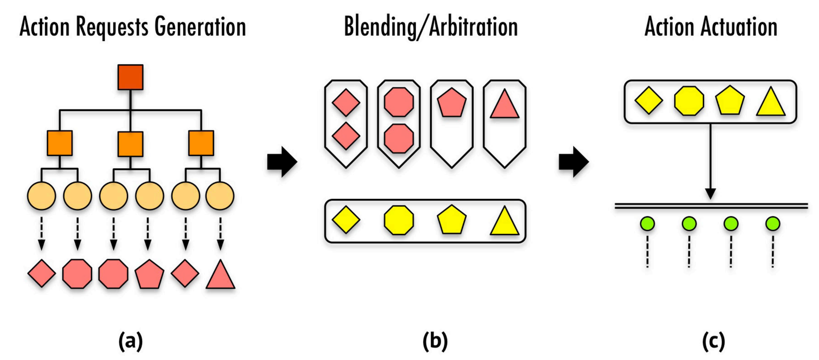


Figure 4. Action blending in a nutshell. (a) Behaviors generate different types of action requests to control different body parts. (b) Action requests are gathered in groups of same type, combined, and packed into a motivation. (c) The motivation is sent to an actuation interface for action rendition.

Behavior Tree Subsumption

Different branches may pursue conflicting goals with different degrees of urgency. Every behavior node has a priority and, if given an action request, it will suppress the execution of lower priority nodes. **Subsumption helps organizing complex BTs in horizontal layers of goals at different levels of abstraction**.

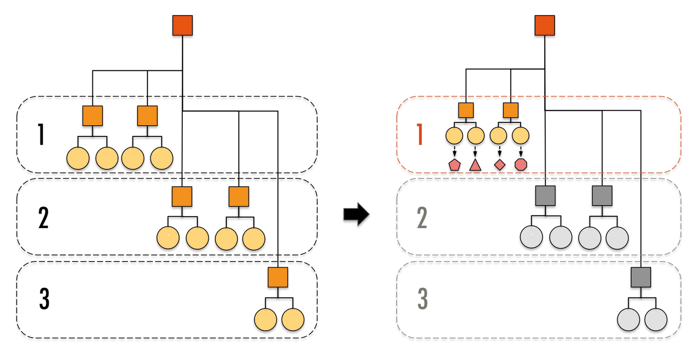


Figure 5. Subsumption can help designing BTs. High priority nodes may achieve critical goals to be resolved immediately when certain conditions hold, while lower priority nodes may achieve latent goals not demanding any urgent resolution.

Conversation Group Dynamics

We have simulated conversation group dynamics in a typical scenario. Some characters are conversing while a newcomer will try to join. For visual reference, around the group we have drawn a schematic of the territory. Characters have mixed autonomous and scripted behavior. Gaze attraction and dynamic group arrangement are fully automated.

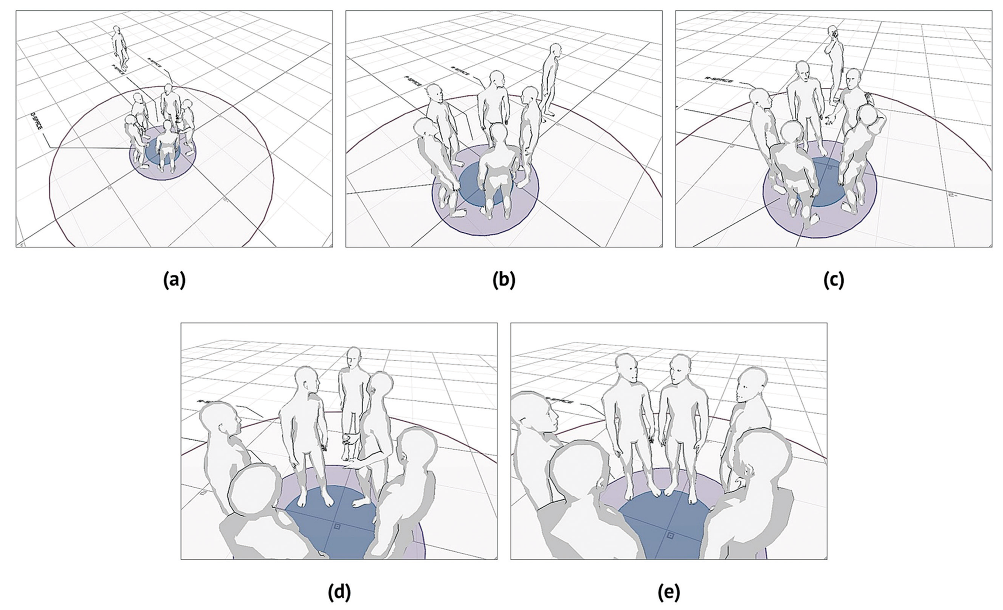
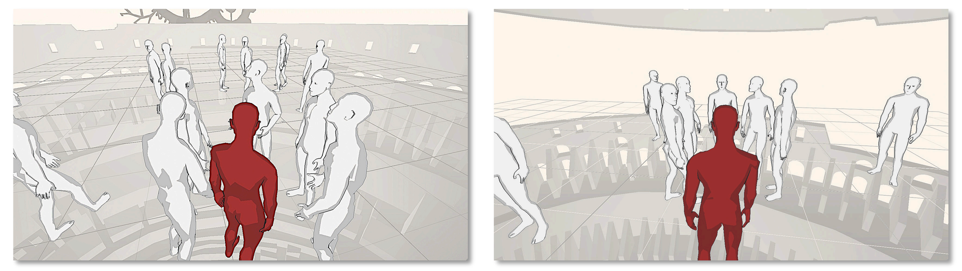


Figure 6. Five steps from our social group dynamics simulation. (a) A newcomer is approaching the group. (b) The passerby is noticed when stepping inside the territory. Through common attention, somebody looks back after watching a member looking at the passerby. (c) The passerby is engaged and invited to join. Other members notice the salutation. (d) The newcomer is welcomed with a short glance. (e) The group opens up to make room for the new member that finally joins.

We also used character social territoriality to improve NPCs responsiveness in an interactive scenario. The user controls a character who can join or leave conversing groups. NPCs react to the player's actions showing awareness of his presence.



Conclusions

Our variant of BTs helps the design of gaze and territorial behaviors for social scenarios. It achieves responsiveness, smoothness and continuity of motion when the decision logic simultaneously controls multiple bodily parts. It makes behaviors easier to extend and reuse.

Video at http://www.youtube.com/impulsionengine

References

[1] KENDON, A. (1990), Conducting Interaction: Patterns of behavior in focused encounters, Cambridge University Press, Cambridge.

[2] PEDICA, C., AND H. VILHJÁLMSSON, H. 2010. Spontaneous avatar behavior for human territoriality. Applied Artificial Intelligence 24, 6, 575–593.

[3] SCHEFLEN, A. E. (1976), Human Territories: how we behave in space and time, Prentice-Hall, New York, NY, USA.

This work was funded by a grant from the Technology and Science Board of Iceland (Rannis Grant Number 093020009).

Contacts:
Claudio@iiim.is
Hannes@ru.is