# Study of Nine People in a Hallway: Some Simulation Challenges

Claudio Pedica
CADIA, Reykjavik University
claudio@ru.is

Hannes Högni Vilhjálmsson
CADIA, Reykjavik University
hannes@ru.is

## ABSTRACT

Many virtual environments rely on a social population that needs to behave in a plausible manner. Crowd simulations typically concern themselves with simulating collision free movement, while getting into and managing social contact with other humans is a less explored subject. Complex situations such as forming conversation groups and recognizing each other presence need to be tackled. This study presents a typical real-life scenario at a university, which in spite of its mundane nature uncovers a wide range of simulation challenges. We apply our reactive rule-based system that models territoriality, proxemics and social navigation to this scenario to demonstrate a possible approach to addressing some of these challenges and contribute to a vocabulary for social simulation.

## CCS CONCEPTS

• **Computing methodologies → Intelligent agents**; **Procedural animation**; *Agent / discrete models*;

## KEYWORDS

Virtual Characters, Social Simulation, Group Behavior, Nonverbal Cues, Animation

**ACM Reference Format:**
Claudio Pedica and Hannes Högni Vilhjálmsson. 2018. Study of Nine People in a Hallway: Some Simulation Challenges. In *IVA '18: International Conference on Intelligent Virtual Agents (IVA '18), November 5–8, 2018, Sydney, NSW, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3267851.3267900

## 1 INTRODUCTION

Many virtual environments rely on a social population that needs to behave in a plausible manner. This has given rise to a rich field of pedestrian and crowd simulation, such as those surveyed in [1]. These works, and some of the more recent improvements such as [11], are mainly concerned with generating realistic collision-free motion of larger groups, which in itself is a major challenge in a dynamic environment. However, social environments need to portray plausible smaller scale interaction, just as much as they need to portray plausible avoidance. This is very clear when observing behavior in places where people know each other, such as on a university campus.

This study presents a typical real-life scenario at a university, which in spite of its ordinary nature uncovers a wide range of simulation challenges. We have chosen to use our own social AI system, a reactive rule-based system that models territoriality, proxemics and social navigation [8] [9], to demonstrate a possible approach to addressing these challenges (see Figure 1). The aim is not to evaluate against a metric of believability but rather to invert the perspective and assume that a proficient simulation designer will always find a way to achieve a believable result. We then show the value of the underlying social principles that help untangle the complexity of a real-life scenario and facilitate the work of the designer. This paper is an extended version of a previous poster presentation [10], going into fully detailed analysis and discussion.

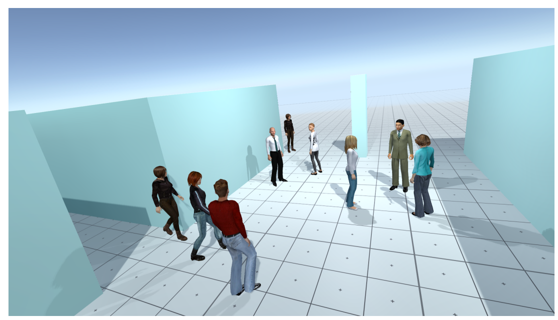


**Figure 1: Screen from a simulation of a real social situation that occurred in a university hallway. Creating this simulation highlighted some challenges in modeling a rich social situation.**

## 2 SCENARIO

To explore the complexity of a naturally occurring social scenario, we chose the hallway of a university, near the end of a class or exam period. The accompanying video[1] contains a recording of the scenario and a few important moments have been summarized in Figure 2. The situation reveals a great variety of social dynamics and nonverbal behavior. The event was located on a large hallway that opened up on two main doors that were kept open. The main doors functioned as entrances or exits depending from which way people were coming and going. On one side of the hallway there was a classroom closed behind a black door. A summary of all the entities involved in the scenario can be seen in Figure 3.

Standing across the hallway, there were two groups of people conversing separately: the first group, that we will call group A (GA), which initially included two members (GAM1 and GAM2), and the second group, that we will call group B (GB), which included three members (GBM1, GBM2 and GBM3). The two groups were

[1] https://tinyurl.com/ninepeople


*IVA '18, November 5–8, 2018, Sydney, NSW, Australia*

ACM ISBN 978-1-4503-6013-5/18/11…$15.00
https://doi.org/10.1145/3267851.3267900

not related to each other and will never come in contact during the scenario. After some time, the black door of the classroom will open and three more people will arrive in the hallway. We call these people arrivals to group A (GAA1, GAA2, GAA3). They will identify the people in GA as acquaintances and decide to join them. As they join the group there will be a new person entering the scene, arriving into the hallway from one of the main doors. We will call her a passerby (PASSERBY). She will swiftly head down the hallway and, in doing so, find her way through the two groups without disturbing the conversations or hit anyone along the way.

The real life scenario that we described above is of particular interest because it displays a number of basic human social skills used coherently and in perfect choreography. Here with social skills we refer to those human behaviors apt to sustain an effective co-presence and face-to-face communication. Altogether they express a body language which is instrumental for social interaction. Some behaviors are better described as relationships between individuals and they contribute to a sense of context [5]. With reference to the real life scenario we want to address specific behaviors such as glancing, mutual attention, group formation, rearrangement, territoriality, proxemics and proper navigation of the social landscape.

To a certain extent we can imitate the behaviors of the real scenario mechanically by means of a procedural model that generates the illusion of social awareness from a set of reactive rules [8] [9]. Originally we embarked on this exploration to asses the accuracy of our own procedural model. At first we considered doing a frame-by-frame comparison of a simulation against the video footage of the real life scenario. However we realized how that would miss the point of procedural generation. If we set the goal to reach the highest level of frame-by-frame similarity one could be tempted to heavily script the simulation to strictly respect the video timeline, transforming our work into an overly sophisticated exercise of linear animation. Moreover even assuming an infallible procedural model with perfect accuracy, we would need to set up the simulation with very precise initial conditions to hope to replicate one-to-one the timing of the events as seen in the video. It would be a daunting job which is outside the scope and practical objectives of the present work.

Instead of considering frame-by-frame accuracy with the video footage, we decided to look at the relative timing of certain behavioral events. From this perspective, instead of saying "GAA1 looked at GAM1 at second 00:09" we consider that "GAA1 looks at GAM1 right after coming out of the door as part of recognizing him". This refocuses the objective from reporting absolute time accuracy (with respect to the initial time of the video) to a reconstruction of the communicative intent. This in turn has lead us to view the analysis and synthesis of the requirements of the chosen scenario as something of general interest to the social simulation community.

## 3 INITIAL SIMULATION CONDITIONS

Using our existing social AI system [8] [9], in the Unity 3D Game Engine, we began from an empty scene that resembled the hallway with the two entrances and the classroom door. Then we created the groups GA and GB, created their members and assigned them to their groups; two characters for GA, and three characters for GB.

The system assigns the status of dyad to GA because it only has two members. Dyads allow for extra control over their members. By default, dyads are face-to-face formations where both members are orientated towards each other. For these types of groups we can specify an extra relative orientation for each member. This parameter allows forcing the default orientation of a member along a certain direction providing a simple way to set up the initial arrangement as close as possible to the real life scenario. On the other side we created GB which is a regular group of three members that will arrange autonomously following the laws of the *F-formation* (see Table 1 for definition of terms). Then we added the three characters inside the classroom. They will come out of the door in a specific order but at different times. Two of them will stay longer in the classroom and come out only few seconds after the first one. To have them wait in the classroom we set an activation delay that will keep them inactive for a certain time from the start of the simulation. We used the same solution for the passerby, that we will call PASSERBY, to simulate a late arrival in the scene.

## 4 SIMULATION

In the real scenario we see the three arrivals GAA1, GAA2 and GAA3 to come out of the door and head into the hallway. Once outside the classroom they will notice their friends GAM1 and GAM2. There will be mutual recognition and the three arrivals will decide to join their friends while the members in GA will welcome them into the group. We were able to model quite simply this rather complex course of events. In essence, we gave the three actors two main goals: to walk out of the door toward a destination point and to approach and join group GA. The two goals are competing but "reaching the destination" has a higher priority and gets executed before "joining the group". This rather mechanical sequence of behaviors will generally fail to conceive any sense of character social awareness. However in our case we found that the addition of gaze behavior brought the scene to life. With good timing and accurate meaning of its communicative function, that is with a clear distinction between glancing and attention shift, the characters looked more lifelike while performing a trivial mechanical execution of two sequential tasks.

Every character has a parameter that sets an "entity of special interest" that will act as a target for the attention shift. An "entity of special interest" could be a familiar person, a point of interests, a public speaker or anything peculiar. When an "entity of special interest" becomes visible and within a certain range then the character shifts attention and looks at it. The attention shift starts when the entity is within the character's *social space* and stops when the target is close enough, that is at the border of the *personal space* (see Table 1). Going back to our scene, two of the characters in the classroom have one of the member of GA as their entity of special interest, or in this case, their best friend. They will look at him as soon as they exit into the hallway and will keep looking at him until they get close enough to join him in conversation. This emergent behavior is generated by a simple composition of "moving to destination" and afterwards "joining the group" and, while this is happening, "shift your attention" to your best friend when he becomes visible. These are the first tokens of the vocabulary

**(a) Start**

**(b) Mutual Interest**

**(c) Approach**

**(d) Passing Glance**

**(e) Forming Group**

**(f) Stabilizing Group**

**(g) Passerby Approaches**

**(h) Passerby Avoids Groups**

**(i) End**

**Figure 2: The significant events of the real-world scenario**

**Table 1: Useful Concepts from Sociology**

| Concept | Source | Description |
|---|---|---|
| Civil Inattention | [2] | Appreciation of presence without expression of special curiosity or intent |
| F-Formation | [5] | System of positioning and orienting that maintains optimal conversation arrangement |
| O-Space | [5] | The space at the center of an F-formation that maintains the interaction and participants need to defend |
| Personal Space | [3] | Space reserved for interacting with close friends and family |
| Social space | [3] | Space for general social interaction with others |
| Location | [12] | An increment of space which is claimed by a person for a certain amount of time. |

that create the illusion of social awareness in the first beat of our imitation of the original real life scenario.

If we were a director on a movie set we would inform real actors what to do and where to go similar to what we did for the intelligent characters. But there is more to the first part of the simulation. So far we focused on the three characters walking out of the classroom but what happens to the members of GA while this is happening? In the real situation the two members of the group look at GAA1 (the first one to walk out of the door) to signal that they know him and, more importantly, that they are inviting him to interact (The relationship between these people is just a hypothesis. In fact we don't know the real relationships between them but, based on the video, it seems plausible to say that GAM1 knows GAA1. Regarding GAM2 we could assume that either she knows GAA1 or she simply turned her head because so did her fellow member in the group).

We resorted again to a simple approach to model this subtle communicative exchange between individuals by using the "entity of special interest". Both GAM1 and GAM2 have as "entity of special interest" GAA1 which is the first of the three characters to come out of the door. The members in GA will then shift their attention to

**Table 2: Challenging events of real-world hallway scenario**

| Time | Event | Description | Challenge | Simulation Approach |
|---|---|---|---|---|
| 00:09 | Mutual Interest | GAA1 shifts attention to GAM1 who does the same thing back. The held mutual gaze signals recognition and availability to interact. | The attention shifts has to happen while the character is moving and has to be a prominent gaze behavior easily identifiable as an attention mechanism. It has to start and end at the right time. | Dormant gaze rule for attention shift. Rule is activated only when entity of special interest is visible and within social range. It becomes target of the attention shift. Mutuality is guaranteed as GAA1 and GAM1 are mutual targets. Gaze rule has the highest priority and subsumes other rules for glancing. |
| 00:11 | Approach | GAA1 walks towards GA. Both GAM1 and GAM2 understand GAA1's intention to join the group. | The approaching phase has to last long enough to get the character closer to the group and signal its intention to join the interaction. Navigation has to accept the behavioral constraints of the F-formation once the character is close enough to the group. The other members have to signal awareness of the character's intention to join by stepping towards it once the character is within a certain range. | Navigation to a location within the group's participants space. Location is a point on imaginary formation before group has found a stable arrangement. Navigation ends as soon as the character enters the participants space to leave the control to the F-formation social forces. There is a location available for every member. The character is considered a group member already during this phase. |
| 00:12 | Civil Inattention | GAA2 and GAA3 come out of the classroom and enter the visual range of GAA1. GAA2 glances briefly at GAA1. | Civil inattention is a general reaction that signals awareness of other people. It is active all the time but has to submit to other behaviors that signal intention such as attention shifts. | Dormant gaze rule for quick glances. Rule is activated when an entity enters the visual range. Closest entity becomes the target of the glance. Rule keeps a memory of the last target to avoid mechanical repetitive glancing to the same entity. This rule has a weaker priority than attention shifts. |
| 00:16 | F-Formation | The members of GA are within range and they start arranging into a face-to-face formation. Within range means inside the participants space. | The self-organization of face-to-face interactions is a complex emergent phenomenon. At its core there is a behavioral relationship between all the members of a group. The relationship exists only when the members are within a certain range. The relationship ensures the enclosing of the o-space, which is essential to the interaction, and the emergence of the typical circular formation. | A motivational rule activated when the actor is a member of a group and is inside the participants space. The motivational rule generates social forces to constraint the actor to the behavioral rules of a functioning F-formation. The motivational forces are turned off as soon as the actor looses the membership or walks too far from the group's region. |
| 00:18 | Rearrangement on Obstruction | GAA2 steps aside to allow GAA1 to have an easy and equal access to the o-space. This is the last maneuver after which every member finds a stable position and orientation. | The F-formation is reassessed anytime its fundamental rules are broken. This often means forming the F-formation again. A special case arises when one member is obstructing another. A character is then too close to the o-space, blocking someone's line of sight. The rearrangement has to be consistent with the fundamental principle of F-formation that every member has to have immediate and equal access to the o-space. | A special case of the rule above activated when the actor is in front of another member. This rule subsumes regular F-formation social forces to let the actor move aside and backward to guarantee free and equal access to the o-space for every member of the group. |
| 00:19 | Group Avoidance | PASSERBY arrives in the hallway and walks through GA and GB without disturbing the conversations. | The navigation has to be obstacle free but also socially aware. The avoidance has to account for the character personal and social space and steer around groups without walking through them. | An extension to the velocity obstacle approach described in [4] can add a term for F-formations. |

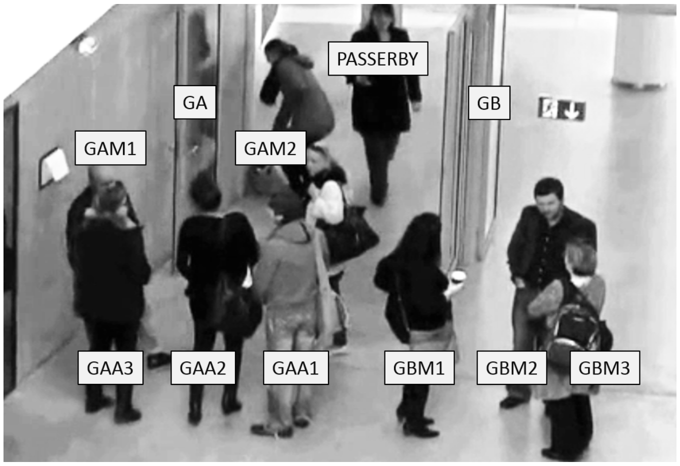



**Figure 3: The entities involved in the scenario**

him as soon as he is visible and within range, performing the actual bodily-visual feedback expected for the occurrence of recognition with an invitation to interact. At this point of the social simulation GAA1 will move toward GA with the intention to join while two more characters will be heading out of the door into the hallway. GAA1 will quickly glance at GAA2 as she overtakes him. The quick glance of GAA2 is a reaction of a general background gaze rule that we have implemented and not the result of a script that has been finely timed. This rule makes every character glance to the closest entity within their visual range. This is a way to simulate the common phenomenon of *Civil Inattention*. What at first seems to be a simplistic generalization for a gaze mechanism provides surprisingly good results in a highly dynamic environment where the state of the simulation changes so frequently. The characters involved are moving and their field of view scans the environment every time the glance rules activates. All this motion provides ample opportunities for picking different glance targets. For those situations in which the simulation's state is rather stable and the scene looks quite static, the glance rules implements a simple internal memory to keep track of who was the last target. By checking its internal memory, the glance rule prevents the actor from glancing at the same target over and over, a circumstance that looks very artificial and that can undermine character believability.

We left the three newcomer characters at the point when they just came out of the door and moved into the hallway. We assigned these characters to GA as part of the initial condition of the simulation. This will inform the characters that they belong to a group and should adapt their behavior to that context. When the characters stand outside the group's region they will try to get closer by entering the group's participant space. In doing so, the actor will navigate the environment (e.g. using path-finding and path-following) in an attempt to reach an imaginary *location* within the group's participant space. The imaginary locations exists only when actors have been assigned to the group but the members haven't established a stable F-formation yet.

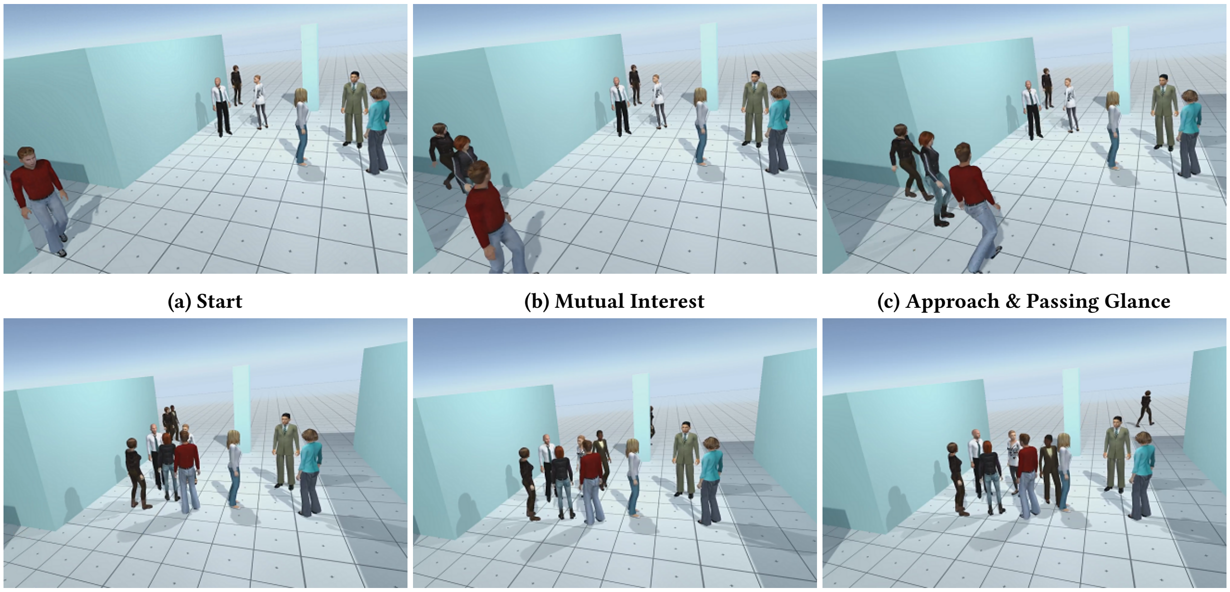
The imaginary locations shape an idealized F-formation or, more generally, a formation mesh that gives a structure to the group while the real emergent structure hasn't been formed yet. Every location is assigned to a group member and this setting, which is editable, is part of the initial condition of the simulation. The actor that has been assigned to a group too far away will walk towards its location until it hits the border of the group's participant space. At that point the rules of F-formation will take over and will constrain the character's motion to adjust its position and orientation in certain ways [7]. When the formation attains a stable state (e.g. nobody will re-position or re-orient) the imaginary locations will consolidate over the actual members position and the real structure of the member's formation will substitute the imaginary one. When the members have found a structure, they will preserve it after any rearrangement. A rearrangement may occur whenever a member moves away from it's current location, for example as a reactive response to an invasion of personal space or because the member is temporarily leaving the group or because the requirements for a valid F-formation aren't met anymore. In a valid F-formation every member has immediate and equal access to the *o-space*, the central area of the formation that has vital role for the members interaction [5].

When GAA1 and GAA2 walk towards the group and arrive to join the interaction, GAA2 will stand one or two steps in front of GAA1 preventing him from accessing the o-space. In the real scenario GAA2 realizes that it is obstructing GAA1 and decides to move aside and backwards; the formation is rearranging. In the simulation we implemented a variant of the F-formation rule that takes over the usual generation of social forces and handles the special case of when a member stands in front of another. Allowing equal access to every member takes priority over keeping the right distance and orientation. Once the F-formation rules are in place the characters know how to arrange and rearrange, generating complex motion dynamics that evokes the appearance of context awareness and intention.

While GA is finally forming and getting into a stable formation, we see that the PASSERBY is making her appearance from one of the main entrances and heading down the hallway. In the real scenario we see her navigating the environment with clear awareness of the people around her and the meaning of their grouping. She avoids any collision with the participants and takes care of not walking through any of the two groups, not invading their interaction space. Her counterpart in the simulation navigates the environment by means of path-finding and path-following. On top of them we run a RVO algorithm for obstacle avoidance. Our RVO algorithm of choice was the one proposed by [4] with social awareness modifications, including an avoidance term for social groups [6].

## 5 RESULTS

The results of simulating the social situation, as described above, are shown in the accompanying video. Significant social events that occurred in the simulation and corresponded to events in the real scenario are shown as screen shots in Figure 4. A comparison can be made between the two sets of images, real (Figure 2) and simulated (Figure 4), but perhaps more importantly we have summarized these events as a set of interesting challenges in Table 2, juxtaposed with how one can handle them.



(a) Start (b) Mutual Interest (c) Approach & Passing Glance

(d) Forming Group (e) Stabilizing & Passerby Approaches (f) Passerby Avoids Group

**Figure 4: The significant events of the simulation**

## 6 CONCLUSION

As with many other works in the intelligent virtual agents community, our approach to the simulation of social awareness and behavior is rooted in the social sciences. Solid knowledge of psychology and sociology contributes in many ways to the process of understanding and then imitating group interaction and social intelligence in general. Important benefits of this include the reinforcement of observational skills and a guidance for the design process. The work begins outside the realm of machines by observing a real life phenomenon. Observational skills mature by exposing oneself to social study. They provide a firmer ground when formulating the complex interplay of people, thanks to which even a simple reactive logic for a procedural model of behavior can provide satisfactory results.

The design of reactive rules needs spatial preconditions and constraints which are much easier to define by referring to the literature. For example, the proxemics theory of [3] clearly provides a ground for qualitative spatial reasoning on a division of space into sectors that identify a set of relations such as is-right-of, is-inside-of, is-left-front-of, etc. Once we have encoded the proxemics qualitative relations in the system it is then rather natural to write logic expressions that serve as preconditions for the activation and deactivation of reactive behaviors. The machinery that we realized is simple but when it works with the right distances it favors better timing for reactions, and timing is essential to attain a plausible appearance of social intelligence.

We encourage other researchers in the intelligent virtual agents community to share more annotated examples of specific simulation challenges, dissecting new social situations and calling attention to new theoretical concepts than can help us to model and design the complex behavioral fabric that we aim to recreate in computational technologies.

## ACKNOWLEDGMENTS

This work was supported by The Icelandic Research Fund under Grants Number 174444-051 and 130705-053.

## REFERENCES

[1] D. C. Duives, W. Daamen, and S. P. Hoogendoorn. 2013. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies* 37 (2013), 193–209.

[2] Erving Goffman. 1967. *Interaction Ritual: Essays in Face-to-Face Behavior.* Aldine Publishing Company, Chicago, IL.

[3] Edward T. Hall. 1966. *The Hidden Dimension.* Doubleday, New York, NY.

[4] Ioannis Karamouzas and Mark H Overmars. 2010. Simulating Human Collision Avoidance Using a Velocity-Based Approach. *VRIPHYS* 10 (2010), 125–134.

[5] Adam Kendon. 1990. *Conducting Interaction: Patterns of behavior in focused encounters.* Cambridge University Press, Cambridge.

[6] Karl Valdimar Kristinsso. 2015. *Social Navigation in Unity 3D.* Master's thesis. Reykjavik University, Menntavegi 1, 101 Reykjavík.

[7] Claudio Pedica and Hannes Högni Vilhjálmsson. 2008. In *Proceedings of the 8th International Conference on Intelligent Virtual Agents.* Springer, 104–116.

[8] Claudio Pedica and Hannes Högni Vilhjálmsson. 2010. Spontaneous Avatar Behavior for Human Territoriality. *Appl. Artif. Intell.* 24, 6 (July 2010), 575–593. https://doi.org/10.1080/08839514.2010.492165

[9] Claudio Pedica and Hannes Högni Vilhjálmsson. 2012. Lifelike Interactive Characters with Behavior Trees for Social Territorial Intelligence. In *ACM SIGGRAPH 2012 Posters (SIGGRAPH '12).* ACM, New York, NY, USA, Article 32, 1 pages. https://doi.org/10.1145/2342896.2342938

[10] Claudio Pedica, Hannes Högni Vilhjálmsson, and Karl Kristinsson. 2015. Study of Nine People in a Hallway. In *ACM SIGGRAPH Conference Motion in Games 2015 Posters (Motion in Games '15).* ACM, New York, NY, USA.

[11] Jose Ramirez, Devin Lange, Panayiotis Charalambous, Claudia Esteves, and Julien Pettre. 2014. Optimization-based Computation of Locomotion Trajectories for Crowd Patches. In *Proceedings of the Seventh International Conference on Motion in Games (MIG '14).* ACM, New York, NY, USA, 7–16. https://doi.org/10.1145/2668064.2668094

[12] Albert E. Scheflen. 1976. *Human Territories: how we behave in space and time.* Prentice-Hall, New York, NY, USA. 1160–1168 pages.