Rannsóknasjóður

2017

Hannes Högni Vilhjálmsson
(School of Computer Science, Reykjavik University)

Character Territoriality: Social Spatial Reasoning for Digital Actors

The Icelandic Research Fund 2017
Project Grant — New proposal
Detailed project description

rannís

GUIDELINES

Please note that *all proposals and appendices* should be in English.

The highlighted text (gray) in this document is provided for guidance purposes only, and requested information should be added by applicant. Please insert appropriate material (i.e. text, pictures and/or tables) and add the principal investigator's name to the header and the project's name to the footer.

The detailed project description should provide the following information, and be divided into the following sections (the order and titles should not be changed):

a. *Specific aims of the project, research questions/hypotheses, feasibility, originality and impact*
b. *Present state of knowledge in the field*
c. *Research plan (time and work plan, present status of project, methodology and milestones) and deliverables*
d. *Management and co-operation (domestic/foreign)*
e. *Proposed publication of results and data storage (including open access policy)*
f. *Contribution of doctoral and master's degree students to the project*

The project description should not exceed **17 pages** including front page and guidelines (1.5 line spacing, 12 point Times/Times New Roman, or similar, 2.5 cm side margins). The bibliography is submitted separately. Convert the file to pdf before submitting. Please note:

→ ***Incomplete proposals will be rejected***
→ ***Proposals which exceed length limitations will be rejected without review***
→ ***Corrections or amendments after the previously announced deadline will be rejected***
→ ***The clarity and the overall quality of the presentation are taken into consideration when proposals are reviewed***
→ ***Applications that are not submitted using the most recent application form available will be rejected***

For further information please refer to the ***Icelandic Research Fund Handbook 2017.***

Please do not delete, overwrite, or amend this Guideline page in your submission

A. Specific aims of the project, research questions/hypotheses, feasibility, originality and impact

A.1. Introduction

We propose to push the boundaries in modeling and synthesis of social character behavior by enriching the state-of-the-art with socio-territorial reasoning and spatial knowledge. Our approach aims to make digital characters more believable as well as deepen our understanding of human social behavior. Digital characters need to show plausible behavior in accordance to the ongoing narrative and context in order for human participants to believe in what they are seeing and for them to be able to naturally interact with the simulated world. Our approach is to package two well established behavior components for digital character automation within a new social behavior component. The two existing components are: (1) Navigation System and (2) Animation System. The navigation system has more or less become standard while high quality animation systems are gaining more and more attention and are getting closer to a stable standard in industry and academia. The new component that we propose will take navigation and animation to a new level by coordinating actions in a way to appear socially plausible in a range of social situations. We envision a mechanical process that mimics the way people reason about space in social interaction. It includes qualitative representation of space as described in psychology and sociology by theories of *human territories*, *F-formation* and *proxemics*.

A.2. Background and Goals

This proposal builds on the success of our "Humanoid Agents in Social Game Environments" (HASGE) project, which received a Grant of Excellence from the Icelandic Research Fund (2008-2010). The goal of that project was to "develop methods in AI and real-time character animation" to make interactive character behavior more believable and engaging than previously possible, focusing on social behavior. HASGE took a holistic approach, defining and addressing control at many different levels of abstraction, ranging from low level animation to high level intent planning. One of the most successful outcomes of the project was a simulation platform based on social steering forces, where convincing conversation groups emerged from a set of basic reactive rules rooted in social theory [13]. The platform, originally termed *CADIA Populus* (see Figure 1) and later *Impulsion*, enabled a wide range of social simulation experiments within CADIA and at collaborating sites. Most recently, a paper from TELECOM ParisTech that incorporates our platform, won a Best Paper award at the International Conference on Intelligent Agents [14] and has lead to a quality journal paper [1]. Another very successful outcome was work on establishing international standards for the major interfaces involved in the generation of communicative behavior in humanoid agents. Working closely with the SAIBA consortium, the project was instrumental in rolling out the first official release

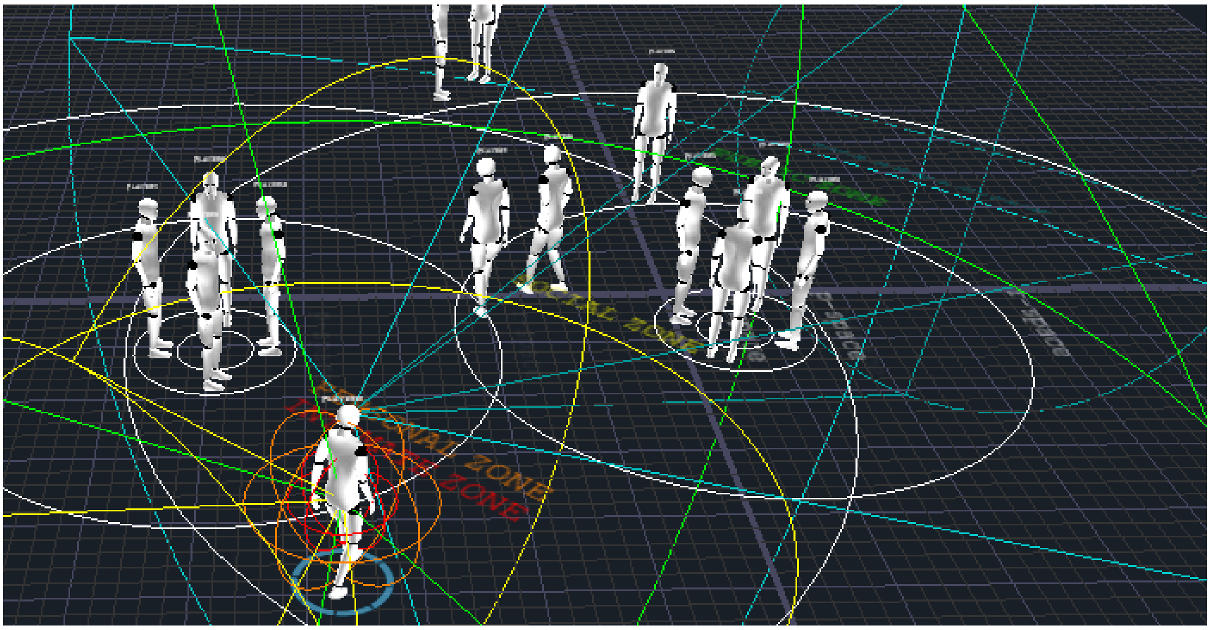

Figure 1: One of the most important results of the "Humanoid Agents in Social Environments" project was the CADIA Populus platform, which we now propose to advance by addressing four fundamental challenges related to human territorial behavior.

of the Behavior Markup Language standard [1].

Using the knowledge gained during the HASGE project and from projects that have employed the resulting technology, such as the "Icelandic Language and Culture Training in Virtual Reykjavík" (Icelandic Research Fund 2013-2015)[27], we have identified four opportunities to significantly advance the state of the art in social simulation:

1. Territorial Motion Interface
2. Arbitration of Conflicting Socio-Territorial Movement
3. Synchronization and Temporal Reasoning
4. Representing and Reasoning about Context

The focus of all challenges is on human territorial behavior, which is one of the most important concepts to emerge from the HASGE research. Before describing each challenge, a few words about the general approach and impact.

A.3. Spatial Representation and Reasoning

Space, like time, is one of the most fundamental categories of human cognition. It structures all our activities and relationships with the external world [24]. It also structures many of our reasoning capabilities and provides a frame of reference for interpreting non-verbal communication. Spatial reasoning, in our every day interaction with the physical world, in most cases is driven by

[1] `http://www.mindmakers.org/projects/bml-1-0`

qualitative abstractions rather than complete a priori quantitative knowledge [3]. A large number of knowledge sources contribute to establishing a mental representation of space. Reasoning about it often involves reasoning about change in spatial configuration, thus reasoning about space-time. In social interaction the behavior of spacing and the behavior of communication are interrelated, thus creating a connection between acting on space and communicating one's intention. The fabric of space relations contributes to a mental structure that, to a certain extent, provides basic context information within which communicative acts gain meaning. Spatial representation and reasoning has been applied to the field of virtual character and robotics mainly to solve problems of navigation. This proposal extends spatial reasoning toward the imitation of social interaction abilities.

A.4. Socio-Territorial Spatial Reasoning

We call our approach Character Territoriality and it includes high quality motion controllers, path-finding, space-time reasoning, context knowledge and reactive planning. We believe that this approach can increase virtual character believability in a number of diverse scenarios. Our approach can be applied to animation, visualization, games and the art and it's especially geared towards immersive storytelling where ambient life and character interaction are pivotal all along the narrative arc. The core of our approach is reactive spatial reasoning, which is a process made of several procedures that work on perceptual information and decide on an immediate responsive action. In our approach we call those procedures reactive behaviors as they implement the logic that then will control the animations. Each reactive behavior implements a constrained-based spatial reasoning; spatial relations between spatial entities (e.g. other characters) which can be expressed by constraints. In qualitative spatial reasoning, knowledge about entities or about the relationships between entities is often given in the form of constraints [15]. For example, we can formalize a rudimentary attention mechanism by creating a behavioral constraint to look at the closest visible entity. Constraint satisfaction is achieved by performing a task: a quick glance to the designated target. The result is a character that will look at passers-by and bystanders.

The machinery of spatial reasoning includes a qualitative representation of space intended to model the cognitive representation of space as described in psychology by the theories of human territories, F-formation and proxemics. They broadly qualify our approach as social spatial reasoning and more specifically as socio-territorial spatial reasoning. Both territories and proxemics are distance systems that include orientation. They are composed of an ordered sequence of distance relations and a division of space into sectors that identify a set of relations such as is-right-of, is-behind-of, is left-front-of, etc. This construct of spatial knowledge allows reactive behaviors to reason on the current state and request an action for actuation by generating a control stream that the architecture will then route to a motion controller interface in the animation sub-system.

A.5. Impact

We expect significant impact on several different fronts:

- **The Intelligent Virtual Agents Community** New constraint satisfaction perspective for solving social behavior planning and furthering the international standardization of territorial behavior control through the Behavior Markup Language (BML) and Functional Markup Language (FML) and the SAIBA consortium.
- **Games, Visualization, Immersive Experiences and Movie Industry** New tools for generating more believable interactive characters at lower cost using automation will be made possible. One can think of this as packaging powerful theories and algorithms behind intuitive user interfaces.
- **Computational Social Science** A new discipline that aims at deepening our understanding of human behavior by leveraging on big data and computational models. Spatial reasoning for territoriality fits into the big picture by inverting the simulation to infer people relations from tracking data.
- **Ambient Intelligence and Human Machine Interaction** Ambient Intelligence, Calm Computing, Disappearing Computing, and Ubiquitous Computing are new forms of machine interaction. New interaction models can emerge from our socio-territorial reasoning that take the users social context into account when creating new user interfaces for ambient intelligence.

What follows is a description of each research challenge which represents an opportunity to advance the state-of-the-art. These challenges also represent a natural work break-down, which is reflected in the project work plan.

A.6. CHALLENGE 1: Territorial Motion Interface

A human territory is not a physical thing. It's made of the space around people and between them. Indeed a territory is visually outlined by the pattern of human behavior and movement. Observing a human territory is an exercise in visualizing the negative space during people interaction. People can express meaning by *designing* the space around and between themselves. The movement they perform in doing so communicates their intent. In previous work we showed how it's possible to synthesize this type of communicative action by modeling human territoriality and the emergence of social self-organization. Digital characters that show behavioral patterns similar to those we see in human territoriality appear more aware of the surroundings, space and context [12].

Considering the novelty of the territorial perspective in the field of social signal processing, one the of hardest challenges we faced early on in previous work was the lack of high quality animation to render territorial behavior. Even now there is no standard practice to render such communicative actions neither in industry nor in academia. One of the important steps to take in this direction is to define the core motion set for animating territorial behavior. What type of motion controllers do we need to synthesize the actions of spatial assertion typical of human territoriality? Gaze, body positioning and orientation are certainly important but we might need more, since Scheflen identified four regions of the human body that are relevant to territorial behavior plus pointing and touching [19]. Once we identify the core motion set for territoriality and the technical requirements for animating those movements, we will formalize an interface between behavioral logic and animation control. The formal interface will serve as a cornerstone for future research work. We envision a behavioral logic able to reason on space knowledge and control the actuation of movement to design the spatial relations in character group interactions. This form of behavioral control can improve the dynamic believability of animation in such a way to make the behavior appear mindful and consequential in a vast range of different interactive scenarios. If the character were a puppet, the research question is akin to asking how many strings and where should they be attached to grant enough control for the imitation of human territorial abilities?

A.6.1 State of the Art

To tackle this challenge appropriately it is essential for us to have access to a high quality animation system. In the past few years the community has achieved remarkable results in character animation [20]. Many solutions are quickly becoming available to the industry and getting integrated in their work-flow. It is rather easy now to get access to an animation system for high quality motion and quickly build a prototype. Our choice is an animation system that combines blend trees with inverse kinematic and that allows for animation re-targeting. The Unity3D engine is an affordable option that offers these features and therefore we have chosen that engine as our experimentation platform.

Blend trees have been quickly accepted by the industry because they provide a good compromise between quality of motion, performance and productivity. Blend trees use a database of traditional mocap animation clips and a number of blending nodes organized compositionally.

This machinery alone can animate a variety of movements that are expected in social situations, ranging from walking and turning to nodding and gesturing. The community has already explored how to control those types of movements but there is an aspect of locomotion that has been consistently left out so far: holonomic locomotion. In recent work, Hughes et al. [8] showed how to apply holonomic locomotion to improve obstacle avoidance in crowd simulation. For the correct synthesis of territorial behavior we need holonomic locomotion, especially to realize positioning behavior in social interaction. While blend trees can be an affordable option for this type of bipedal locomotion, they aren't particularly adapt for gaze animation.

Body movement necessary to imitate human gaze are better synthesized using an inverse kinematic approach. A simple IK solver can take a target point as input and compute the bending angle of the spine bones to achieve high quality real-time gaze motion [21]. Both of these animation systems, blend trees and IK solver, are affordable components and relatively easy to get access to in Unity3D. They are ideal for fast prototyping and experimentation within the interface to core territorial motion.

A.6.2 Contribution

From our past work we know that adjusting gaze, position and orientation provides a basic level of control for rendering territorial behavior but further investigation will shed light on whether or not a division of the body into multiple regions can have a substantial impact on character believability. For the behavioral logic, the motion controller interface will have a purpose which is twofold: motion actuation and proprioception. On the one hand, the motion controller exposes a set of parameters that the behavioral logic can manipulate to actuate the character's movements. On the other hand the motion controller exposes a set of attributes that the behavioral logic can read to reason about the state of the character's body. The accurate description of the motion controller interface is going to be important in the overall economy of our approach to social simulation. The animation system that we plan to use supports animation re-targeting and there is a decoupling between the abstract representation of motion and it's actual manifestation as moving parts of a digital character. This is made possible by an emerging industry practice for the creation of humanoid characters in games and animation. As long as the motion controller interface is well defined and established, we can expect to also re-target the behavioral logic along with the necessary abstract movements. We are opening up to the possibility of creating a library of reusable behaviors, organized around the concept of human territory, that will represent an extra contribution in the work-flow of digital character design for immersive storytelling. In sum, the contributions are:

- We are going to identify the core motion set for the synthesis of territorial behavior and propose an interface between behavioral logic and animation
- We are going to contribute to the work-flow for the creation of humanoid characters with social abilities.

A.7. CHALLENGE 2: Arbitration of Conflicting Socio-Territorial Movement

In Character Territoriality, the behavioral architecture is composed of several reactive behaviors that work on perceptual information and decide on an immediate responsive action.

Conceptually we can think of a reactive behavior as a process that decides to activate a body region and perform an action for a certain amount of time. We design the behavioral architecture to guarantee two properties: quick reaction to contingencies and fluent choreography of several motions. To achieve the second property we need parallel execution of the reactive behaviors in order to simultaneously control more than one body region. Multimodality is an important component of what makes virtual characters to look socially plausible. However the management of multiple parallel behaviors that might simultaneously perform conflicting movement is quite challenging and we haven't solved the problem yet for social territorial behavior. It seems natural to design an arbitration mechanism to handle those conflicts. We intend to model this as a resource allocation problem and draw inspiration from scheduling solutions for operating systems. Within this model the body parts that the behaviors aspires to control are seen as resources to manage. The behaviors will request the control over a body region with the purpose of performing an action (e.g. controlling legs and torso for walking) and waits to get their turn if the body region is occupied. The arbitration mechanism will then handle allocation, waiting time, interruption, priority and rescheduling of requests.

Behaviors that generate gaze requests need a different arbitration mechanism than those generating requests for positioning. A typical scenario, historically quite challenging for us, is the one of a character who navigates through a dynamic environment while paying attention to several points of interest including other characters. This scenario is highly dynamic (e.g. the character is moving, other characters are moving and the point of interest might also change) and at a given time the architecture has to handle, let's say, three gaze requests to three different targets at once. The requests will compete for the control over the head-look apparatus. Which target should be looked at first and for how long? What is the best way to schedule the gaze requests to keep motion fluency with a high level of believability and responsiveness? We know that discarding or ignoring gaze requests doesn't produce a naturally looking attention mechanism. Possibly even worse if we arrange the requests sequentially and try to satisfy them all within the allocation times that they have requested. Often the end result tends to look mechanical and unnaturally repetitive. On the other hand, posture orientation works quite differently. In a scenario with two requests for body orientations it is perfectly legal to choose the orientation in the middle. Thus the arbitration strategy for posture orientation seems closer to those applied in command fusion architectures [16]. The management for positioning behaviors is quite similar. A group of reactive behaviors might constrain the locomotion apparatus simultaneously but none of them should exclude the others. For example, imagine a scenario where a character is moving closer to a group (first constraint) while keeping an equal distance to every member (second constraint) but wanting to stand in closer proximity of a certain member (third constraint). Once again, we could apply command fusion arbitration to combine the three control streams but this approach isn't always satisfactory as much of the quality for the end animation is determined by a careful parameter tuning. Either we need to improve the combination strategy or to reorganize the positioning and orienting behaviors in such a way that the architecture can scale robustly while producing high quality results.

A.7.1 State of the Art

Parallel behavior request arbitration has been applied in DiscoRT [10], a BML realizer for virtual and robotic conversational agents. A precious source of inspiration for us because they worked on resource allocation in the domain of communicative agents. However they focused on use cases where arbitration of locomotion requests was not needed. However, the work of Huang and Kallmann on motion planning and placement for virtual characters [6] shows how to synthesize locomotion for precise of departure and arrival position and orientation. Their idea of deforming a motion trajectory by using the desired orientation as a parameter is interesting and significant. Also their work on motion parametrization [7] is quite important in the economy of our goals and might be instrumental in designing the arbitration strategy for positioning and orientation required by Character Territoriality. We will investigate further and extend their work towards group interaction.

A.7.2 Contribution

We envision a new component for socio-territorial spatial reasoning that arbitrates between several concurrent reactive behaviors and resolves conflicts. We will design the arbitration component around the model of resource allocation where behaviors are requesters and motion capabilities are resources. The core of the arbitration component is going to be the arbitration strategies for gaze, positioning and orientation requests. In sum, the contributions are:

- Architectural extension to handle resource allocation between behavioral logic and motion control.
- Arbitration strategies tailored to schedule different classes of actuation requests from the spatial reasoning behavior logic.

A.8. CHALLENGE 3: Synchronization and Temporal Reasoning

So far we described our behavioral system as performing reactive spatial reasoning on abstract territorial structures. However we know well that timing is an important, if not crucial, factor for rendering plausible social behavior. The synchronization of the character's action is important to give the illusion of intentionality. Therefore we plan to extend spatial reasoning with temporal reasoning in such a way to handle the coordination of non-verbal behaviors to achieve even finer action timing than plain spatial reasoning does. It should be noted that space intelligence alone already introduces some level of synchronization. For example, a behavior that triggers a certain communicative action (e.g. attention shift or stepping sideways) when an event occurs within a given proximity, allows the action to happen at a specific instant in time. When more than one action is activated under similar conditions, the character's movement appears synchronized especially in highly dynamic environments. This way of controlling action timing through spatial constraints can achieve a rudimentary level of coordination but it's not ideal because it is hard to control. It doesn't make temporal constraints explicit in a clear way and makes it much harder for a designer to synchronize actions with finer precision, especially if there are many actions. It's a problem similar to the behavioral control of gesturing which the community has already been working on extensively and that we plan to take inspiration from.

Character territoriality controls the actions that communicate the spatial relations of characters and groups. The communicative actions that we aim to coordinate within Character Territoriality are: gaze, holonomic walking, posture orientation, torso orientation, salutation gestures and possibly pointing and touch.

A.8.1 State of the Art

Many BML realizers implement the temporal reasoning necessary for action coordination and, more importantly, temporal constraints are formally defined as part of the language. The most recent achievement in this area is the AsapRealizer2.0 [23] which extends BML to allow fluent behavior realization. They achieved a remarkable set of architectural features such as incremental plan construction, graceful interruption, top-down, bottom-up, and environmental adaptation of ongoing behaviors. Their work provides a solid ground for our plan of extending socio-territorial behavior with temporal reasoning. Also the classic work on SmartBody is relevant. They applied several promising techniques to solve the synchronization problem such as meta-controllers for time shift warp, blending and scheduling of underlying motion generation [22].

A.8.2 Contribution

The definition of BML has evolved over the years but much of its formalism is the results of the knowledge gathered from conversational settings with only two members: either a man and a machine or two machines "talking" together. We developed an expertise in group face-to-face interaction and we can contribute to the community by extending BML towards group dynamics. Mixing action synchronization and temporal reasoning in our territorial approach will contribute in several ways:

- We improve the results of our approach, which is mainly space-based, by including temporal reasoning.
- We stay compliant to the BML formalism thus preparing a common ground to discuss character territoriality with the rest of the community
- We lay down the basis for extending the SAIBA framework with spatial reasoning on human territories and proxemics.

It's not clear yet whether the spatial reasoning for territories should directly extend BML or pertain to the FML layer. It's a profound question that opens up a debate on whether spatial and temporal reasoning are clearly separated in the SAIBA framework or, conversely, there is a common space-time reasoning across all the layers from communicative function down to behavioral planning. Indeed our extensive work on groups sets us apart from the rest of the community and gives us the opportunity to contribute substantially to an area that hasn't been thoroughly developed yet.

A.9. CHALLENGE 4: Representing and Reasoning about Context

In social interaction, the communicative function of an action is interpretable only within the circumstances that form the setting of an event. Generally we refer to the setting as context. Defining context is hard because it is a broad concept that includes several representations at once. It has been identified as one of the future problems in Social Signal Processing [25][26] and one of the open issues for FML [2]. Context as such can include information about the established relationships among a group of people, the intended use of a location, the arrangement and type of furniture, the time of day, the cultural background including language and mannerism and, at a larger scale, even the historical moment in which an event is taking place. Those are examples of what might constitute contextual information and the different level of abstraction at which it presents itself. For virtual characters with social skills we know that contextual information is important for behavior generation. A territory is a spatial structure that provides a qualitative description of space; zones and areas with intended social purpose. Indeed a territory is a construct of spatial context as we know from the theories of Proxemics [5], Human Territories [19], and the essential work of Kendon on conducing interaction [9]. The region of space in which an action is performed contributes to the interpretation of its communicative function and the recognition of the role of the performer. In some situations the behavioral logic can reason upon it to deduce the role of the members in a social interaction by simply referring to their qualified position and orientation. Position and orientation within the territorial structure contributes to the definition of someone's role in the interaction floor. For example, the F-formation is the typical spatial context of a conversation. It's made of a nucleus, that includes the participants, and of a region, that encloses the nucleus, which usually is an area reserved for bystanders or new arrivals. Thus a character that stands somewhere in the region, oriented towards the group and focused on one of the participants, is likely to have the role of new arrival that wants to join the conversation. We can use spatial context not only in role recognition but also in behavior generation. Let's say a character has the goal of joining a conversation. Its intent will generate a plan that will break down into phases. Here the territorial spatial structure can inform every phase with context information that will have an effect on the behavior realization (e.g. approach the group by moving somewhere within the conversation's region) as well as provide a mechanism for behavior coordination (e.g. as the new arrival approaches look at her until she has reached the participants zone). When the spatial structures form spatial knowledge that is coherently shared among all characters within the frame of reference provided by Character Territoriality, the awareness of the territorial structure can contribute to the movement coordination which is so important for the simulation of plausible social interaction. Thus the territorial spatial structure can provide a degree of contextual information and serve as a foundation for defining a language for context reasoning and context awareness in a computational model of social intelligence.

A.9.1 State of the Art

Two important issues in FML are the definition and separation of contextual information and the management of multiple interaction floors and the roles of participants [2]. Some level of context description in FML has already been proposed in [18]. Our goal is to bring a territorial-spatial perspective into the context definition of FML. In doing so we plan to consider one of the most audacious observation of Albert Scheflen and expand on it. He described the territorial fields as somewhat recursively defined, with higher level territories able to embed lower level territories. We believe it is worthwhile to investigate further and explore the possibility to define context as a recursive and compositional structure of other contexts: a spatial context tree. With such representation we could model contextual information at different levels of abstraction. More specifically we could model virtual characters that belong to more articulated social settings, of the kind that has multiple interaction floors within a larger interaction floor. For example, a participant engaged in a private conversation inside a larger social group which is surrounded by a vast audience of viewers. At any given time the participant can have multiple roles, each one belonging to a specific level in the context tree. We envision a unified definition of context that starts at the higher level with the description of the social situation at hand, designed around Goffman's concept of frame [4]. An approach to model and architect this concept of framing for behavior generation is provided by Rovatsos et al. [17].

The social situation will then embed a description of static environmental boundaries and the hierarchy of territories, from the larger to the smaller instances all the way down to the participants. For every participant we will have a personal context describing static traits (e.g. gender, age, personality, etc), cultural preferences and personal relationships with other members.

A.9.2 Contribution

This unified definition of context has the potential to profoundly impact the community with both technical and theoretical contributions. It will provide context knowledge at every layer of reasoning of the SAIBA framework with the potential to affect the character's behavior in many ways. The great advantage of using territorial structures as context information is that they describe relations that are independent of specific distances. Thus standing close to someone's personal space will translate into different interpersonal distances if the social event happens in a park rather than an elevator. The reasoning logic will stay unaltered and adhere to the structure of space which is independent from quantitative representations. Moreover, territories exists in space-time. They are created and dismissed, and last for a certain period of time leading us to the idea of dynamic context definition. To the best of our knowledge, ours will be the first work to propose a dynamic definition of context and the unification of the ideas of Goffman, Kendon, Scheflen and Hall under a unique computational model. We already foresee the many challenges that such a complex representation of context might pose on the behavior realization, especially on arbitration and temporal synchronization. Our behavioral logic is made of a collection of spatial constraints that we plan to extend to the temporal dimension. How is the spatial context tree going to affect the constraint resolution process? How is a more expressive definition of context going to affect the whole agent architecture from behavior planning, to behavior realization and finally animation? We are going to tackle these challenges and explore what we see as an opportunity to bring a coherent and consistent paradigm of spatial territorial reasoning within the SAIBA behavior generation framework. In sum the contributions are:

- Introduce a territorial perspective at the intent level of the SAIBA framework.
- Introduce territorial-based context to FML for roles in single and multiple interaction floors.
- A recursive, compositional, dynamic definition of context.
- A unified definition of context comprising situation framing, territories and personal traits.

B. Present state of knowledge in the field

See the "Background" section and the "State of the Art" sections under individual research challenges above.

C. Research plan (time and work plan, present status of project, methodology and milestones) and deliverables

C.1. Methodology and Work Plan

While the research questions are organized around 4 distinct challenge areas, they form an architectural and conceptual whole that will advance at a steady pace throughout the project through frequent prototypes and analyses. Producing visible results early and fast serves multiple goals: (1) Internal feedback for iterating on solutions, (2) sharpen focus on tool usability and clear communication of ideas, (3) building a portfolio of experiments, some of which will succeed to inspire and educate the community. We will start with

Prototypes, each addressing a specific social scenario that relates to a particular challenge, will be produced every 6 months (see Figure 2). Prototyping involves both building models from theories that explain the situation studied and also collecting and analyzing reference data from the real world as a reality check. All prototypes will be analyzed and critiqued with our collaborators, but also made available to the general community for feedback, open discussion, re-use and modification.

Two kinds of additional formal evaluations will be performed on three of the prototypes: (1) Systematic comparison between the real world data and the simulated results, utilizing data visualization to highlight important deviations; (2) Visual perception study (in part based on [12]) where a subject first observes and then interacts with the digital actors through a full room scale Virtual Reality setup, which embeds the subject completely within the simulated scenario. Survey instruments, such those addressing social presence, as well as behavioral parameters, such as proximity and gaze, will be used in subject studies.

Our starting point is the *Impulsion* simulation platform that grew out of the HASGE project [11] [2]. The iterative approach means that the results of analyzing a prototype can potentially feed into any challenge area, e.g. a scenario produced in year 3 may call for improvements or additions at the level of animation support. Therefore, all challenge areas remain active throughout the duration of the project, but they are introduced one at a time to manage the overall complexity.

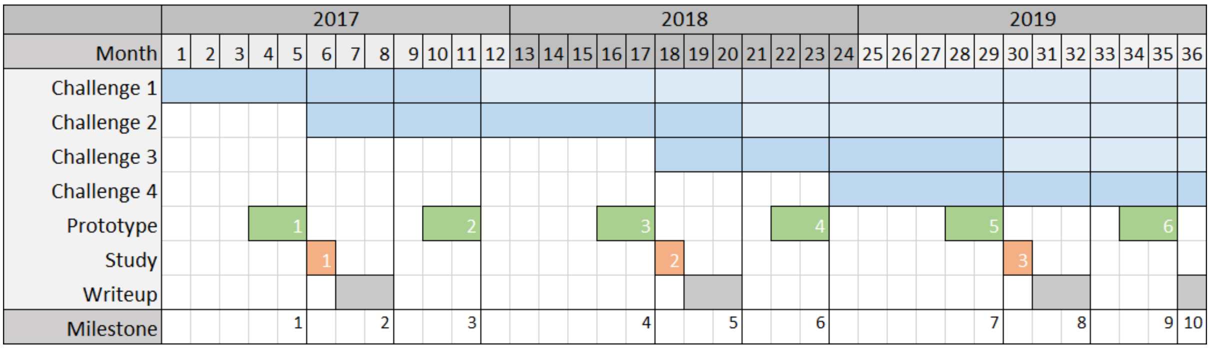

Figure 2: Three year iterative plan for integrating solutions from challenge 1 through 4 to deliver frequent working prototypes for analyzing and sharing with the research community.

C.2. Scenarios and Prototypes

A number of prototypes simulating challenging scenarios will result from the work. For now they are just suggestions since we expect to change or re-purpose them as we develop the approach and involve our collaborators. We are going to work on scenarios such as:

1. Approaching an individual by considering environmental features and proxemics.
2. Plausible, adaptable and controllable reaction to invasion of personal space.
3. The phases of entering and leaving a conversation with correct movement coordination.
4. Fluent, consistent and reactive gaze behavior in small crowds.
5. Adding variation by controlling the temporal features of group behavior.

We will make the prototypes by quickly wiring up available components and keeping a fail-fast approach. We aim at low fidelity prototypes that we can easily modify to test our hypotheses, as if they were interactive sketches. By keeping low production costs more resources will be available for experimental work and knowledge production with potential for discovery.

[2] `http://secom.ru.is/projects/impulsion/`

C.3. Milestones and Deliverables

Completing a prototype corresponds to an important milestone and deliverable, packing up the work so far as a comprehensible and focused demonstration for sharing (see Table 1). Study results, written up as publishable papers, are also an important type of milestone and deliverable. Final write-up will include a concrete proposal for BML and FML extensions.

Milestone	Type	Year	Description
1	Prototype	2017	Scenario 1 related to challenge 1
2	Paper	2017	Results from Study 1
3	Prototype	2017	Scenario 2 related to challenge 2
4	Prototype	2018	Scenario 3 related to challenge 1 and 2
5	Paper	2018	Results from Study 2
6	Prototype	2018	Scenario 4 related to challenge 3
7	Prototype	2019	Scenario 5 related to challenge 4
8	Paper	2019	Results from Study 3
9	Prototype	2019	Scenario 6 related to all challenges
10	Paper	2019	Overall Results

Table 1: Important project milestones and deliverables

D. Management and co-operation (domestic/foreign)

D.1. Management

The Principal Investigator will conduct local weekly project meetings as well as individual meetings with doctoral students and MS and BS students working on theses. Regular Skype meetings will be conducted with our foreign collaborators. Two workshops with our collaborators are scheduled during the first two years. The first will introduce and integrate the teams, sharing initial knowledge and setting goals. The second workshop will focus on lessons learned so far, goal revision and on joint publishing and dissemination strategies. Both workshops will discuss opportunities for joint EU and US funding for related work.

D.2. Collaboration (see attached Letters of Intent)

Dr. Catherine Pelachaud, Director of Research CNRS at LTCI, TELECOM ParisTech

Dr. Pelachaud's research group at TELECOM ParisTech has already shown successful integration of our first-generation character territoriality technology *Impulsion*. A paper that built on this integration won the best-paper award at IVA [14] and resulted in an ACM journal publication [1]. Dr. Pelachaud and her colleagues will continue to integrate features into their virtual agent platform, help with evaluation and provide invaluable feedback on how well our solutions generalize across other systems and scenarios.

Dr. Takehiko Nagakura, Department of Architecture, Massachusetts Institute of Technology

Dr. Nagakura heads the Computation Group of MIT's Department of Architecture, which focuses on the development of innovative computational tools, design processes and theories, and on applying these in creative, socially meaningful responses to challenging design problems. Dr. Nagakura and his PhD student Paloma Gonzales have developed methods for collecting, analyzing and visualizing human-space interaction, which they will contribute, while jointly working on scenarios and simulations with us.

Dr. Kjartan Pierre Emilsson, Chief Executive Officer, Sólfar

Dr. Emilsson co-founded Sólfar in October 2014 with two other game industry veterans. Their mission is to create and publish pure VR games that thrill players, and push the boundaries of virtual reality entertainment. Currently developing multiple projects across all major VR platforms, including PlayStation VR, HTC Vive and Oculus Rift. As representatives of the storytelling industry, Sólfar will help design scenarios with high relevance and potential industry impact. We will seek opportunities for joint tech demos and publicity.

E. Proposed publication of results and data storage (including open access policy)

An integral part of this project is active communication and open sharing with the social signals and simulation community, both in academia and in industry. This is to receive useful early feedback, but also to inform the field about the possibilities that territorial management affords. This will revolve around the prototypes, which will be shared through a web-site promoted in social media. Papers will be published at quality conferences with open access options such as IVA, AAMAS and AIIDE and submitted to high impact journals such as ACM's Trans. on Graphics and Trans. on Computer-Human Interaction. We will also aim for a paper in a quality non-technical journal such as Environment and Behavior or American Behavioral Scientist. Talks and demos at industry oriented venues will also take place, especially with our industry collaborator Sólfar. Collected natural human behavioral data in raw form will not be shared, to preserve privacy of subjects, and will be stored on file servers with strict access control.

F. Contribution of doctoral and master's degree students to the project

The bulk of the research will be carried out by 2 doctoral student and part-time Research Affiliate, Claudio Pedica. Claudio is the original author of the social simulation platform and an interdisciplinary expert with extensive experience in social theory, technology and interaction design. Doctoral students will be recruited through our collaborators and academic network and will each be expected to focus on 2 of the 4 challenges. The MS students will contribute to creating and testing prototypes, mainly through paid summer work. A limited number of MS and BS students will also get a chance to participate throughout the school year in exchange for course credits, some of which will result in MS or BS theses.