Computational Proxemics: Simulation-based analysis of the spatial patterns of conversational groups

Kavin Preethi Narasimhan

Submitted for the degree of Doctor of Philosophy

Queen Mary University of London

2016

Declaration

I, Kavin Preethi Narasimhan, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below. I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material. I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Kavin Preethi Narasimhan

September 2016

Details of collaboration and publications:

- K. Narasimhan and G. White. "Agent Clusters: The Usual versus the Unusual", in Y. Demazeau et al. (eds), Advances in Practical Applications of Heterogeneous Multi-Agent Systems – The PAAMS Collection, Lecture Notes in Artificial Intelligence 8473 (Springer 2014), pp. 244–255.
- K. Narasimhan and G. White. "Look, Who's Talking: Simulations of Agent Clusters", in Y. Demazeau et al. (eds), Advances in Practical Applications of Heterogeneous Multi-Agent Systems – The PAAMS Collection, Lecture Notes in Artificial Intelligence 8473 (Springer 2014), pp. 375–378.
- K. Narasimhan and G. White. "An Agent-Based Analysis of F-Formations",

in Y. Demazeau et al. (eds), Practical Applications of Agents and Multi-Agent Systems, Lecture Notes in Artificial Intelligence 7879 (Springer 2013), pp. 239–250.

- K. Narasimhan. "SPAGE: An Action Generation Engine to Support Spatial Patterns of Interaction in Multi-agent Simulations.", in S. Omatu et al. (eds), Distributed Computing and Artificial Intelligence Volume 151 of the series Advances in Intelligent and Soft Computing (Springer 2012), pp. 273–280.

- K. Narasimhan. "Towards modelling spatial cognition for intelligent agents", Proceedings of the 29th Annual European Conference on Cognitive Ergonomics (ACM 2011), pp. 253–254.

Computational Proxemics: Simulation-based analysis of the spatial patterns of conversational groups

Kavin Preethi Narasimhan

Abstract

In real-world conversational groups, interactants adjust their body position and orientation relative to one another in order to see and hear clearly. We use an agent-based modelling approach to compare alternative models for simulating the spatial patterns of conversational groups. The models are based on simple rules that control the movement, positioning, and orientation behaviour of individual agents, which in turn leads to the emergence of agent clusters. We identify which model alternative produces agent clusters with characteristics typical of real-world conversational groups.

The centroid-based approach, where agents readjust their position and orientation with respect to the group centroid point, is a commonly used method to simulate conversational groups, but has not been empirically validated. This thesis replicates, evaluates, and validates the centroid-based model in a systematic way. Another model, where agents perform positional-orientational readjustments to see as many neighbours as possible within a 180° field of view, called the field-of-view approach is proposed, implemented, evaluated, and validated.

Analysis of the spatial patterns of conversational groups has hitherto mostly relied on visual verification. We, novelly, use a combination of qualitative and quantitative methods to analyse the spatial patterns of conversational groups. Evaluations show that the field-of-view model and centroid-based model produce agent clusters with significantly different social, spatial, and temporal characteristics.

Validation is performed using a dataset which captures the spatial behaviour of 21 participants for the entire duration of a party. This validation shows that the characteristics of agent clusters resulting from the field-of-view model most closely reflects the characteristics of real-world conversational groups. We also show that a local neighbourhood influence works better than an extended neighbourhood influence to simulate conversational groups. The influence of objects in the environment on the spatial patterns of agent clusters are also discussed.

Submitted for the degree of Doctor of Philosophy

Queen Mary University of London

2016

Contents

List of Figures

List of Tables

Acknowledgements

I am grateful to several people for helping me through this journey. I have and will always cherish these fantastic people in my life – I am because you are.

Firstly, I thank my principal supervisor, Graham White, for his valuable guidance and phenomenal support through all these years. He has been a mentor and friend to me, and has helped me overcome several professional and personal struggles. I am grateful to him for helping me reach the finish line. I thank my second supervisor, Paul Curzon, who has been very supportive and offered insightful feedback on my work over all these years.

I sincerely thank my examiners, Marco Gillies and Alan Penn, for their constructive and insightful comments on my thesis during my viva, and for pointing the way to future directions in the field.

I'm grateful to my Cognitive Science research group family. I thank Pat Healey for his valuable guidance and mentoring throughout my PhD. I thank Matt Purver for his advice and support on several occasions. Christine Howes has been my go-to person since my first day as a PhD student. I am truly grateful for her encouragement and support and her timely and valuable feedback on my thesis. I thank Pollie Barden for being the best writing buddy and a stellar friend. Arash Eshghi, Stuart Battersby, Colombine Gardair, Mary Lavelle and Chrystie Myketiak for being inspirational seniors and helping me along the way. Julian Hough, Nicola Plant, Sam Duffy, Saul Albert, Paolo Masci, Kleominos Katevas, Jeni Maleshkova, Toby Harris, Claude Heath and the GHack team for being the best PhD buddies and officemates and for helping me all along.

I thank my extended EECS family who have helped me in various aspects of my research. Claire Revell, Colin Powell and Sue White for helping me run the filming study. Vlado Kitanovski and Riccardo Mazzon for helping me understand the basics of computer vision processing. The EECS IT team for supporting me throughout my PhD and for being great colleagues. Fabrizio Smeraldi for his insightful advice and encouragement on several occasions. Jennifer Murray for being so kind and understanding and Melissa Yeo for keeping track of my progress.

The Centre for Research in Social Simulation at Surrey has been a second home to me over the last year and a half. I am truly grateful to Nigel Gilbert for his extraordinary support. He was kind enough to give me time to complete my thesis despite being needed elsewhere. His pioneering work on agent-based modelling has inspired my thesis. I thank Jie Jiang for proofreading my thesis. I have thoroughly enjoyed the lunchtime and corridor conversations on agent-based modelling, and just about any topic under the sky, with all my wonderful colleagues at CRESS.

I thank my friends for being my support system. I'm particularly grateful to Kamesh and Girija for being my family away from family. I thank all my friends from school, college and work for motivating me from all over the world, particularly Liz, Earnest, Gowri, Aparna, Prasi, Reha, Egi, Anu, Sneha, Vaishu and Deepa.

Last, and most importantly, I thank my family for their love and support, always. Aswin and Vidya, my brother and sister-in-law, for being my pillars of support. My in-laws, Rajan, Asha, Sandeep and Kavitha, for being kind and understanding, and loving me unconditionally. My grandmother, Prema, and my nephew, Krish – the sunshine of my life – for being who they are.

This thesis would not have been possible without the help and support of three people. My husband, Kiran, for being my best friend and staying strong through difficult times. My parents, Narasimhan and Ananthalakshmi, for loving me and caring for me so limitlessly, and for providing me the warmth and safety of home even after being away from home for so long.

This work was supported by the Queen Mary University of London College Doctoral Training Account award and the Rabin Ezra Scholarship Award.

I may have inadvertently missed some names here, and may have not done enough justice to those mentioned here, however, I am, and my thesis is, because of several wonderful people to whom I will forever be indebted.

Part I

Introduction and Literature Review

Chapter 1

Introduction

This thesis uses an agent-based modelling approach to investigate the role of movement, positioning, and orientation behaviour in the functioning of conversational groups. There is a huge body of work on the spatial organisation of face-to-face encounters, notably Goffman (1961, 1963, 1971), Scheflen (1964, 1975), Birdwhistell (1952, 1970), Hall (1966) and Kendon (1973, 1979, 1990, 2010), among others. Several computational models have also been proposed to simulate the spatial organisation of face-to-face encounters. Most of these models use what we refer to as the centroid-based approach, to simulate conversational groups, but the results of the approach have not been empirically validated in a systematic way. In this thesis, we have replicated the centroid-based approach, as well as implementing an alternative model to simulate the spatial arrangements of conversational groups. We evaluate and validate the results of both models, using a combination of qualitative and quantitative methods, to identify which approach works better for simulating conversational groups.

This chapter introduces the focus of our inquiry – the spatial orientational behaviour manoeuvres involved in everyday face-to-face interactions. We begin by providing two examples that demonstrate the routine yet significant characteristics of spatial behaviour in social situations. We then turn to delineating the scope of our work in this thesis and introducing some relevant terminology. In section 1.3, we explain our research methodology, and then in section 1.4, we introduce the research questions addressed in this thesis. Finally, we present the chapter structure and a brief description of each chapter, as a road map to the reader.

1.1 Setting up the Scene

Imagine the following scenario. Ann and Bob are standing next to each other and talking in a crowded railway platform. Claire walks by and wants to get past them to wait for her train, but she recognises the duo are engaged in conversation, and tries to avoid walking in between. However, the station is pretty crowded, so she is forced to walk through. As she does so, she dips her head, and apologises (utters "sorry!"). She walks past the duo, and moves to a more comfortable location, where she is no longer considered to be an *intruder* in the conversation. Claire also faces away from Ann and Bob, as if to clearly denote – through her body position, orientation and posture – that she is not involved in Ann and Bob's conversation. Claire's behavioural manoeuvres indicate her "non-membership" in the conversation, not just to Ann and Bob, but to everyone in the scene at the time. Therefore, given the current spatial configuration between Ann, Bob and Claire, it is more likely that a random person would walk in between the duo and Claire, rather than interrupting Ann and Bob. In other words, the spatial-orientational arrangement of the trio serves as a cue, indicating to people the presence of a conversational group and a non-participant (in the conversation) individual standing nearby, and by virtue of that, imposes a need to adopt spatial manoeuvres appropriate to the situation.

The previous example was only concerned with one conversational group, but what if there are many co-located conversational groups? Consider a different social scenario, e.g. a cocktail party, wherein there are many small conversational groups distributed across the party hall. Jim, an invitee, enters the party venue and would like to join his friends Anna and Marie, talking to one another across the party hall. Jim begins to walk towards the duo, intending to join them in their ongoing conversation. However, en route Jim runs into John, an acquaintance, and stops to interact with him instead. In the meantime, Anna and Marie notice Jim, so they break off their interaction temporarily, and start walking towards Jim (talking as they walk). When they arrive, Jim acknowledges their arrival, and they join the ongoing conversation between John and Jim. So a four member conversational group is now formed between Anna, Marie, John and Jim. Furthermore, the way in which these interactants stand with respect to one another (i.e. their position, orientation and posture) clearly manifests a social-spatial relationship, which others in the party hall recognise and respect. That is to say, other people in the party hall wouldn't normally stand near, walk close to, or walk through the conversational group. On the other hand, if a person intends to join the group, they would do so using suitable mechanisms, e.g. arrive near the group, and wait for an invitation to join just like Anna and Marie did.

These scenarios are common occurrences which all of us are aware of, and acknowledge, by virtue of growing up in any society that has its own social norms. There is also a huge literature on the topic, focusing on the characteristics, functions, and social relevance of spatial behaviour in conversational groups. Some of this research used observational strategies to capture, as closely as possible, the dynamics of processes involved in naturally occurring conversational groups. Others have used subjects and/or confederates to stage conversations and analyse the resulting behaviour. In the last couple of decades, findings from these studies have motivated the computational modelling community, to develop models capable of simulating conversational groups. However, a downside of these models is that the results of the simulation (or animation) have not been systematically validated to confirm if the characteristics of simulated groups reflect the characteristics typical of human conversational groups.

The lack of validation poses an evident risk that models are not capable of producing realistic outcomes, or they are actually doing it, but lack the evidence to confirm it. Even worse is the possibility that models are overrating the mechanisms needed to simulate the spatial organisation of conversational groups, e.g. until the simple and elegant boid model was proposed (Reynolds, 1987), flocking behaviour was simulated using complex scripting procedures.

In this thesis, we address the current gap by evaluating and validating the agent-based models we have developed for simulating the spatial organisation of conversational groups. Our work: (1) accounts for the credibility and effectiveness of the models, and more importantly, (2) provides a systematic assessment of how rules concerning the movement, positioning, and orientation behaviour of individual agents leads to the formation of agent clusters resembling human conversational groups.

1.2 Delineating the Scope of Inquiry

Human conversations entail different forms of behaviour, some are explicit (e.g. speech and gestures), while others are comparatively subtle (e.g. body position and orientation). At a time when sociological inquiries focused on spoken utterances as the main aspect of conversational gatherings, several notable sociologists emphasised the need to study the non-verbal behaviours and paralinguistic features facilitating the conduct of interactions. Goffman noted that the use of language in conversations, as well as the behavioural mechanisms that provide the medium for dialogue exchange, should be studied as an exclusive field of sociology (Kendon, 1988). He suggested calling the field *Interaction Order* (Goffman, 1983), *public order* (Goffman, 1963) or *interaction ethology* (Goffman, 1971).

Birdwhistell (1970) proposed *Kinesics* as the study of movement related non-verbal behaviour associated with the conduct of face-to-face interactions. Kendon (1988) notes that the works of Goffman and Birdwhistell are complementary to one another. Researchers like Edward T. Hall and Adam Kendon further advanced the field – *proxemics* is the study of the use of space surrounding the body (Hall et al., 1968) and *F-formation system* denotes a system of spatial-orientational behaviour organisation in conversational groups (Kendon, 1990).

In this thesis, we examine the characteristics of the positional-orientational behavioural dynamics of interactants engaged in face-to-face conversations (conversational groups) using computer simulations, hence we have used the title *computa-*

tional proxemics. We focus on modelling the movement, positioning and orientation behaviour of individual agents, using different rule sets, to test the emergence of agent clusters akin to human conversational groups. The different rule sets each correspond to one model. We have a baseline called Model 1, an alternative to the most common approach to simulate the spatial patterns of conversational groups called Model 2a, and lastly, an adaptation of the most commonly used centroid-based approach called Model 2b. We consider a party context for our simulations and therefore call it the Party World models.

In Models 2a and 2b, we also vary the scope of neighbourhood, based on which agents readjust their position and orientation. If the scope is local, agents readjust their position and orientation only with respect to immediate neighbours (Model 2a LA and Model 2b LA), whereas if the scope is extended, agents perform readjustments with respect to every other agent in the group (Model 2a EA and Model 2b EA).

We evaluate the spatial, temporal, and social characteristics of agent clusters resulting from the different Party World models, based on the conceptual frameworks of Hall's (1966) proxemics and Kendon's (1990) F-formation systems. We then validate our models by comparing the characteristics of agent clusters resulting from our models with the characteristics of real-world conversational groups. We created the Drinks Reception (DR) party dataset exclusively for the purpose of this thesis to validate our models. The dataset contains 4720 images and values of the position and orientation of 21 people who participated in a drinks reception party.

1.3 Agent-based Modelling as a Research Method

Agent-based modelling provides the means to experiment with rules concerning the behaviour of individual agents, and thus to grasp the mechanisms of emergence at a group-level or macro-level. The method has continued to gain popularity over the last few years as a way of conducting systematic scientific research (cf. Gilbert and Troitzsch (2005); Railsback and Grimm (2011)).

We use an agent-based modelling approach to devise, implement, evaluate, and validate alternative models aimed at simulating the spatial patterns of naturally occurring conversational groups. There are three key entities in any agent-based model: agents, environment, and the rules for interactions among agents and between the agents and the environment (Gilbert, 2008; Macal and North, 2010). The agents in our models, defined as entities in a computer simulation, are intended to represent humans. The environment denotes the virtual space within which agents reside and function. We specify the rules for interactions as program constructs which determine how, when, and with which other entities, agents interact during the simulations. In our models, agents interact by relating their movement, positioning and orientation behaviour with respect to one another, as well as the environment. By adopting agent-based modelling as our research method, we take advantage of the following methodological possibilities:

Replicability: Wilensky and Rand (2007) note that replication of agent-based models is a useful exercise to verify and validate the models, as well as to develop a shared understanding among modellers. For example, Reynolds' (1987) boid model has been replicated extensively, to simulate the coordinated steering behaviour of agents. Huth and Wissel (1992) proposed a fish schooling model similar to Reynolds' (1987) boid model, and went a step further to validate the simulated schooling patterns against the characteristics of real fish schools. The most commonly used approach to simulate the spatial patterns of conversational groups is what we refer to in this thesis as the centroid-based approach. However, to date, the accuracy or performance of the centroid-based approach has not been validated. To address this limitation, in this thesis we have replicated the centroid-based approach, as well as evaluating and validating the outcomes of the model. We also provide the specifications of the other models that we devised and implemented for the purpose of this thesis. We have evaluated and validated the alternative models, too.

Specificity: Compared to textual descriptions, which are open to different interpretations, the clear, precise and complete descriptions of agent-based models allow less room for misinterpretations (Gilbert and Troitzsch, 2005). The specificity aspect is useful in clarifying the causal relationship between the spatial behavioural rules specified in each of our Party World models and the characteristics of agent clusters resulting from the models.

1.4 Research Questions

This thesis addresses the following research questions:

1. What are the differences between the Party World models in terms of the F-formations they generate?

2. What are the differences in the spatial patterns and durations of agent clusters resulting from the different Party World models?

3. Which Party World model produces agent clusters whose spatial patterns most closely resemble to those of human conversational groups?

The first question is concerned with the emergence of F-formations, i.e. to examine if the Party World models generate agent clusters that qualify as F-formations. Kendon (1990) notes that an F-formation system arises whenever interactants organise their spatial-orientational behaviour with respect to one another during face-to-face encounters. Therefore, we examine if the movement, positioning and orientation behaviour of agents in our models leads to the emergence of agent clusters that are akin to F-formations.

The second question is concerned with examining the differences in the spatial patterns and temporal characteristics of agent clusters resulting from the different models. We examine how the positional-orientational behavioural rules in the different models affect the spatial layout and duration of agent clusters. The last question is concerned with validating the Party World models, i.e. confirming if, and how, the social, spatial and temporal characteristics of agent clusters resulting from the models relate to the characteristics typical of real-world conversational groups.

1.5 Thesis Structure

This thesis has ten chapters divided into four parts. Table 1.1 provides an overview of the thesis structure:

Parts	Chapters
I Introduction and Literature Review	1. Introduction 2. The Role of Spacing and Orientation in Conversational Groups 3. Computational Models of Conversational Groups
II Method	4. Party World: Our Agent-based Models and Simulations
III Results: The Validation of Party World Models	5. Analysis of the Emergence of F-formations 6. Analysis of the Spatial Patterns and Temporal Characteristics of Agent Clusters 7. Sampling the Spatial Patterns of Real-world Conversational Groups: The Drinks Reception (DR) Dataset 8. Validating Party World Models using the DR dataset 9. Findings from a User Evaluation Study
IV Conclusions	10. Conclusions and Future Work
Appendix	A. Full tables for Statistical Analysis

Table 1.1: Overview of thesis structure

- **Part I Introduction and Literature Review:** Part I has three chapters. This chapter is intended to set the scene for the research undertaken and reported in this thesis. Chapter 2 is a review of previous research explaining the role of movement, body position and orientation in the functioning of human conversational groups. Chapter 3 is a review of existing models for simulating the spatial patterns of conversational groups. It is notable that the research reviewed in chapter 3 was mostly built on the conceptual basis reviewed in chapter 2, however, the models (and their outcomes) have not been systematically validated in most cases. The current gap in the literature is the basis for our Party World models and their evaluation and validation presented in the later parts of the thesis.

- **Part II Method:** Chapter 4 provides a detailed description of the Party World models. Kendon (1988) notes that cocktail parties are an example of social situations that involve multi-focused gatherings, i.e. many individual conversational groups co-located and functioning simultaneously in a well-defined space, with an overarching social context. The party scenario described in section 1.1 is an example. The Party World models presented in this thesis explore the social-spatial dynamics of such multi-focused gatherings in a party context.

- **Part III Results:** This part consists of five chapters presenting our findings based on the analysis of the characteristics of agent clusters resulting from the different Party World models. Firstly, chapter 5 focuses on testing the emergence of F-formations in the Party World models. We also perform a timeline analysis, to explain how the spatial patterns of agent clusters undergo changes over time depending upon a number of factors, and how the patterns of changes vary between the different Party World models. Chapter 6 focuses on detecting, evaluating and analysing the different types of dyadic and multi-party arrangements resulting from the different Party World models. We also examine how long agent clusters exist, and identify the differences between models in terms of the durations of agent clusters resulting from the models. Chapter 7 describes the process involved in filming, post-processing, and annotating the images captured at a drinks reception party, to obtain the feet position and head orientation of 21 participants, for the Drinks Reception (DR) dataset. In chapter 8, we compare the characteristics of agent clusters resulting from our models with the characteristics of conversational groups captured in the DR dataset. Lastly, in chapter 9, we present the findings from an evaluation study, conducted before implementing the Party World models, to present the reasons which led to developing the Party World models, and to suggest further research possibilities.

- **Part IV Conclusions:** The last part of the thesis consists of chapter 10, which presents a summary of our findings and conclusions. We show the limitations of the centroid-based approach to simulate the spatial patterns of conversational groups. We provide evidence that suggests that our alternative to the centroid-based approach (i.e. Model 2a) produces agent clusters whose characteristics are most similar to the conversational groups in the DR dataset. We also show that a local neighbourhood influence is better than an extended neighbourhood influence to produce realistic spatio-temporal arrangements of conversational groups. Chapter 10 also recognizes the limitations of our models, thereby suggesting the directions for future research. Lastly, we present our implementation of a semi-immersive virtual world installation, where users can participate in conversational groups with computer controlled humanoid characters, and our proposal to improve the installation.

1.5.1 Thesis Contributions

1. We have devised and implemented alternative models to simulate the spatial patterns of conversational groups using an agent-based modelling approach.

2. We have used, for the first time, a set of new techniques for evaluating the spatial patterns of conversational groups, the analysis of which has hitherto been restricted to visual verification or user feedback.

3. We have validated the Party World models using the DR dataset created exclusively for the purpose of this thesis. The dataset contains the position and orientation of 21 participants for the entire duration of a party. To the best of our knowledge there are no accessible datasets which provide the position and orientation of so many people for the entire duration of a party.

4. We have established the causal relationship between the different models and the social, spatial and temporal characteristics of agent clusters resulting from the respective models. We have shown the differences between models in terms of the characteristics of agent clusters they produce. We have validated the different models, showing which of them produces agent clusters that most closely resemble real-world conversational groups. To the best of our knowledge, existing models to simulate conversational groups have not been validated against real-world conversational groups.

5. We have highlighted the advantages and limitations of the different methods we have chosen to simulate, evaluate, and validate the spatial patterns of naturally occurring conversational groups.

Chapter 2

The Role of Spacing and Orientation in Conversational Groups

We consider a conversational group as a collection of two or more individuals engaged in face-to-face communication. The body position and the relative orientation of individuals in a conversational group are key behavioural aspects of face-to-face interaction. They serve two important purposes. Firstly, they allow individuals to see and hear one another clearly. Secondly, they provide a structural basis for a conversational encounter, to facilitate the uninterrupted exchange of verbal and non-verbal information between interactants. In this chapter, we review the literature focusing on the role of interactants' body position and orientation in the functioning of conversational groups.

We begin with a brief overview of the different subject areas in which the spatial organisation of conversational groups has been investigated. In the same section, we also describe the focus of our inquiry. In section 2.2 we clarify the terminology used in relation to the study of conversational groups. In section 2.3, we review the conceptualisation and significance of the space surrounding the human body, both for personal and collaborative use. We then review theories concerning the role of interpersonal distances and mutual orientation in the spatial organisation of conversational groups in sections 2.4 and 2.5 respectively. We review the *F*-

formation system, a unit of social spatial behaviour organisation, which emerges from the spatial behaviour coordination of interactants, and the *equilibrium theory*, which accounts for the persistence of spatial configurations of conversational groups in section 2.6. We review the different types of F-formations, and the role of the physical environment in influencing the spatial configuration of F-formations, in section 2.7. In section 2.8, we investigate how people deal with deviations in spatial behaviour in expected, as well as unexpected, social situations. This is followed by a brief review of the social dynamics of co-existing conversational groups in occasions such as parties in section 2.9. Finally, we conclude the chapter by discussing the key issues that motivate our research.

2.1 Studies on Spatial Behaviour and the Focus of Our Inquiry

The role of proprioceptive cues in face-to-face communication, and the organisation of spatial behaviour in conversational groups, have been studied under different banners, such as cognitive anthropology, conversation analysis, ethology, ethnomethodology, exchange theory, network analysis, sociolinguistics and symbolic analysis (McDermott and Roth, 1978). In addition to this, *kinesics*, *proxemics*, and *coenetics* are three related yet different disciplines, that focus exclusively on the role of non-verbal, body-based communication.

Birdwhistell (1952, 1968) introduced kinesics, the study of body behavioural communication, focusing on how humans use their bodies as instruments of communicative behaviour. Birdwhistell suggested that the ability of people to behave appropriately in a public setting using bodily manoeuvres, and conversely, to identify appropriate or inappropriate behaviour of that sort, is a consequence of the fact that body motion has an intertwined role with language in human communication and social interaction. Birdwhistell also noted that interaction of communication does not end when interactants lapse into silence, only to resume again at the onset of spoken utterances; rather, the exchange of communicative signals is a continuous process that also involves other forms of bodily communication, such as body movement, gaze and gesture. He therefore considered the two modes – body motion and language – as infracommunicational systems that are interdependently merged.

Goffman remarked that the human body, as a whole or its parts, can be used to *relay* or *exude* information to other individuals (Goffman, 1963, p. 14). Relay denotes the intentional passing of information, while exude denotes information made apparent by one's bodily presence. Patterson and Edinger (1987) explain that non-verbal behaviour, expressed by bodily actions, represent relayed communication when they are intended to achieve a specific goal. On the other hand, just the bodily presence of a person conveying information to others (e.g. regarding one's position, orientation and posture), is referred to as indicative behaviour or exuded communication. Goffman's works also led him to propose *interaction order*, an area of study focusing on the traffic rules of interaction, and not with why people interact or what they achieve when they interact (Goffman, 1983; Kendon, 1988).

Hall et al. (1968) proposed *proxemics*, a field of study that focuses on the perception and use of space by humans, for performing various daily activities. Hall differentiates proxemics from kinesics in two ways. Firstly, proxemics considers the role of physical settings (i.e. the environment) in governing the use of space, whereas kinesics is only indirectly concerned with it. Secondly, proxemics has a narrower scope than kinesics. Proxemics is principally concerned with the position and orientation of the body and the positional-orientational relationships between individuals. Interpersonal distance, eye contact, openness of arms and legs, touching and holding, forward-back lean of trunk, and the direction of orientation of a speaker towards his addressee are examples of proxemic behaviour (Mehrabian, 1968). Kinesics, on the other hand, covers a much wider array of bodily communication, ranging from whole body motion through to subtle behaviour such as eyebrow motion (Kendon, 1977, Ch. 6).

Wescott (1966) proposed coenetics, a biosocial study of communication – an umbrella discipline that includes kinesics, proxemics and linguistics. Wescott identifies communication as anything that ranges from the collision of subatomic particles to the graceful exchange of conversations. He organised the communicative patterns covered by such a broad classification along a linear scale, with patterns enabling a minimal level of communication at one end, and those enabling a maximal level of communication at the other end. Wescott notes that the use of language, a maximal level of communication, has evolved over billions of years from other more minimal forms of communication. Kendon (1979) and Ciolek (1983) endorse using the term coenetics to refer to the study of spatial-orientational organisation of face-to-face interactions.

Researchers have also tried to understand human social-spatial behaviour based on insights from the discipline of animal ethology (Becker, 1973; Vine, 1975). Drawing on the territorial behaviour of animals, Hediger (1961) defined *territory* as a section of space occupied by an individual or a social unit of individuals, with a specific internal structure and a definite spatial extent. Drawing on Hediger's (1961) work, Hall (1966) suggested that in the case of humans, an individual's territory acts as an invisible bubble of space that one maintains between oneself and others (and other things). Hediger suggested that territories enable individuals to protect themselves, and to perceive, i.e. to see and hear, one another clearly. Territories also have a temporal dimension: Cheyne and Efran (1972) and Scheflen (1975) note that a territory is an area that is bounded for a time in a discernible way and used by an individual or group of individuals.

Lindskold et al. (1976) has pointed out that, despite being comparable forms of behaviour, in animals territoriality involves a primary imperative to defend, whereas in humans, territoriality results from acknowledging and contributing to the privacy and social norms of courtesy. In a similar vein, Efran and Cheyne (1973) note that, while aspects of animal and human territorial behaviour may be related, they are not the same, and therefore, cannot share the same forms and mechanisms.

This thesis agrees with the principal idea behind these varying yet related branches of studies – that communication is a major component of the human use of space. The focus of our investigation is the role of spatial behaviour of individuals in establishing and sustaining the spatial layout of conversational groups. Within the scope of our inquiry, we do not deal with any emotional, physiological or psychological causes, situated in individuals, which might give rise to their spatial behaviour.

Rather, we consider communication in the sense Birdwhistell defines it: in the presence of one another, people can actually communicate using non-verbal behaviour (e.g. posture, position and orientation), even when there is a lack of explicit phonetic exchange (Kendon, 1977, ch. 6). We aim to simulate the dynamics of spatial behaviour in conversational groups as a form of communication in its own right. In Wescott's scale of communication, the scope of our research is neither at the minimal end nor at the maximal end, but at the level of interactants coordinating their spatial behaviour with respect to one another to provide the structural basis for conversational groups.

2.2 Clarification of Terminology and Definitions

Goffman (1963) uses the term *interactants* to refer to individuals participating in conversational groups. Kendon (1990) uses the terms *members* and *participants* interchangeably. We use the terms participants, members and interactants interchangeably throughout this thesis.

As noted before, we regard a conversational group as a collection of individuals engaged in face-to-face interaction. Within conversational groups, also referred to as conversational gatherings (Kendon, 1973), participants are influenced by one another's presence. Goffman (1963) considers mutual influence as a consequence of co-presence, i.e. the condition in which two or more individuals are close enough to perceive, and be perceived by, one another in whatever they are doing. Gill et al. (2000) note that co-presence is a precondition for communication; it allows people to hear, see and touch one another, and thus has a bearing upon how individuals coordinate with one another during face-to-face encounters.

Co-presence is essential for face-to-face communication, but co-presence does not mean people will communicate. As such, co-presence enables two different types of social gatherings: focused gatherings and unfocused gatherings (Goffman, 1963). Activity is always located, people always do something from somewhere (Kendon, 2010). Even when doing nothing, people are still present somewhere. The physical presence of an individual at a particular place at a particular time is manifested by several features, including their position (coordinate position on the ground plane) and their orientation (rotation of the body with respect to the median line extending outwards from the body). When two or more people come together to perform a joint activity, they tend to relate their position and orientation with respect to one another, such that they demonstrate their *withness* with one another during the occasion (Scheflen and Ashcraft, 1976). On such instances, where people actively demonstrate their withness with one another, they are considered to be members of a *focused gathering*. On the other hand, those instances where people merely acknowledge and momentarily react to the presence of others in their immediate vicinity, are referred to as *unfocused gatherings* (Goffman, 1963, p. 24).

In focused gatherings individuals engage in a joint activity, whereas in unfocused gatherings, although co-present in the same physical space, individuals do not pursue activities jointly (Patterson and Edinger, 1987). Conversational gatherings, card games, musical performances, fist fights, and the more formal interactions, such as interviews and lectures, are examples of the former (Argyle and Kendon, 1967). Collections of individuals at waiting areas or lounges in airports and railway stations and lifts are examples of the latter (Kendon, 1979). Both focused gatherings and unfocused gatherings can occur within the same *social occasion* that provides the overarching social context for the gatherings (Goffman, 1963, p. 18). For instance, if the social occasion is a party, it is common to find several clusters of conversational gatherings, whereas on a crowded platform, it is common to witness both focused and unfocused gatherings.

In conversational groups, the spatial behaviour of participants – such as their movement, posture, position and orientation – influence, and are influenced by, the conduct of interactions within the group (Kendon, 1990). F-formation system refers to the behaviour coordination mechanism by which participants relate their spatial behaviour with respect to one another in focused gatherings. The spatial behaviour coordination mechanisms adopted in F-formation systems are readily recognised by those directly engaged in them, as well as by those who look at them (Scheflen and Ashcraft, 1976). In the absence of F-formation systems, co-present individuals can demonstrate their non association or lack of withness with one another. Avoidance behaviour such as stepping aside, turning away, or looking away, are examples of behaviour demonstrating non-association.

Figure 2.1 shows the hierarchical relationship we have derived based on the definitions of social occasion, co-presence, focused and unfocused gatherings, F-formation systems and conversational groups. In the following sections, we will review how individuals relate their body position and orientation with respect to one another in co-present social encounters.

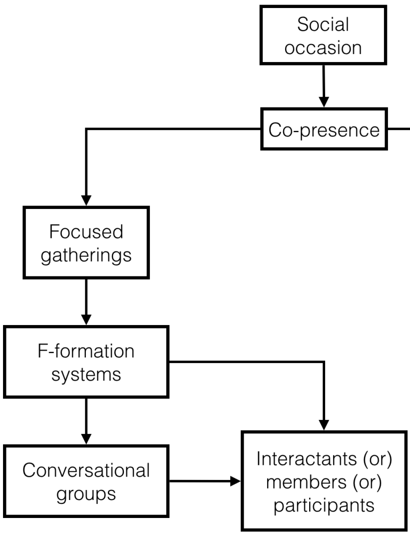


Figure 2.1: Classification of co-present social encounters

2.3 Characterisation of Space Surrounding the Body

In this section, we review the characterisation of the significance of the space occupied by the human body, and the space surrounding the body. As noted in section 2.1, territory is the space surrounding an individual's body, with a definite spatial extent and an internal structure, used for the purposes of self-protection, and to see and hear one another clearly (Hediger, 1961; Hall, 1966). Territories also have a temporal extent in that they are spaces used by an individual, or a group of individuals, for a bounded time period (Cheyne and Efran, 1972; Scheflen, 1975). Goffman identified eight types of territories to which individuals claim spatial access and situational access(Goffman, 1971, pp. 29-41). Scheflen also identified five types of territories by considering an individual's body as the frame of reference (Scheflen, 1975). Following is a comparative review of the territories identified by Goffman, Scheflen and others:

1. *Personal space* is the inviolable space around an individual's body. Hayduk (1978) and Patterson and Edinger (1987) define personal space as the area surrounding the body into which others cannot step in without causing discomfort. Personal space is also referred to as *individual space* (Vine, 1975) or the *body buffer zone* (Ciolek, 1983). Considering an individual's body as the point of reference, the extent of personal space varies depending on the density of the crowd in the environment, and the character of the occasion in which the individual finds himself (Goffman, 1971).

   Vine (1975) notes that while personal space signifies the minimum tolerable distance, *social space* signifies the maximum separation distance an individual will tolerate from others. The distinction is similar to Hediger's (1961) classification of personal space and social space in animals.

2. *Body sheath* indicates the skin that covers the body and the clothes worn by an individual. Body sheath is the minimal configuration of an individual's territorial claim to personal space. Scheflen's (1975) definition of *cubit* indicates the space occupied by the human body elbow to elbow when standing or seated on a chair. Scheflen (1975) defines k-space as the minimal space allocation required to contain the whole body of an individual.

3. *Possessional territory* is the extent of space occupied by the personal belongings of an individual. Scheflen (1975) uses the term *spot* to denote the area of space claimed by putting down one's personal possessions, such as a purse, bag or camera.

4. *Use space* is the spatial extent demarcated by the instrumental needs of an individual. Scheflen (1975) uses the term *o-zone cubit* to denote the area of space just outside an individual's k-space, which could be reached by extending the arms.

Scheflen also notes that a *location* is a time bound integration of spots and k-spaces used by an individual.

5. *Stall* is the well bounded space to which individuals can claim territorial rights on a temporary basis. E.g., boxes at the opera.

6. *Information preserve* denotes the set of facts about an individual, to which he expects to control access, while in the presence of others.

7. *Turns* are queues or row formations where the spatial order, in which an individual is positioned with respect to other individuals, will determine the order in which the individual's request to goods or services is dealt. Scheflen (1975) uses the term *modules* to refer to a row or queue of locations (see above the definition of a location), or multiple rows and columns of locations, as in the case of crowds of spectators.

8. *Conversational preserve* is the right of a group of conversing individuals to preserve the spatial boundary of their group from intrusions and overhearing.

Lyman and Scott (1967) also identifies four types of territories based on the situational factors influencing the behaviour of individuals. *Public territories* are those to which an individual has freedom of access but not freedom of action, meaning that a certain level of normative behaviour is expected of individuals in public territories. *Home territories* are those where regular participants have a relative freedom of behaviour, a sense of intimacy, and a control over the area. An *interactional territory* denotes the area of space that serves as a social membrane encompassing social gatherings. Lastly, *body territory* denotes the space occupied by the human body.

It may be noted that Goffman's (1971) conversational preserve and Lyman and Scott's (1967) interactional territory denote the same thing, i.e. a spatial territory exclusive to a group of interacting individuals. Likewise, Scheflen's (1975) k-space and Lyman and Scott's (1967) body territory denote the same thing, i.e. the region of space occupied by an individual's body.

Becker (1973) and Patterson and Edinger (1987) uses the term *jurisdiction* instead of territory. Becker argues that territory conveys a sort of permanent ownership, whereas jurisdiction is suggestive of a temporary defence of space, for a specific instrumental purpose. Patterson and Edinger (1987) agree with the notion of jurisdiction. Furthermore, Goffman (1971) notes that territories could be demarcated using explicit spatial markers that announce territorial claim, e.g. a library visitor leaving books at the table. Goffman uses the term *central makers* to denote the spatial markers or objects used to demarcate territories. Based on this, Becker (1973) argues that when spatial markers are not used to demarcate territories, it rather makes sense to refer to them as jurisdictions.

Hayduk (1978) differentiates personal space from territory based on the concept of mobility. He suggests that both personal space and social space are mobile spatial units, because when an individual moves, the regions of space surrounding his body also move. On the other hand, Hayduk considers territory as a relatively fixed space, which is applicable only when people are stationary. Knowles (1973) notes that personal spaces do not have any explicit boundaries, or in other words, they have invisible boundaries. Therefore, as Becker (1973) also acknowledges, personal space, territory, and jurisdiction are three loosely related conceptualisations referring to the space surrounding the body.

2.4 Interpersonal Distance

Conversational groups consist of two or more individuals interacting with one another. Patterson and Edinger (1987) note that "personal space" is not a useful term in the context of studying the role of space in conversational groups. As social interactions always deal with two or more participants, they suggest using the terms "interaction distance" or "interpersonal distance" instead. In this section, we review theories about interpersonal distances.

Interpersonal distance is the actual extent of spatial separation between the foot positions of individuals. Deutsch (1977) provides an operational definition of interpersonal distance as the length between the centres of the toe-to-toe distance between two individuals. Maintaining an appropriate interpersonal distance is a pre-requisite for interactants to be able to face one another more or less directly (Deutsch, 1977), and to see and hear one another clearly while interacting (Hall, 1966). Interpersonal distances also help in keeping the "talk lines" open between interactants (Goffman, 1963, p. 161).

The interpersonal distance suitable for face-to-face conversations falls in between the personal space and social space regions. Vine (1975) refers to this distance, which will vary also based on the orientation of the body, as the *equilibrium distance*. An exact measure of the equilibrium distance is hard to establish, but there have been some suggestions. Goffman (1963) indicates that participants would find it difficult to interact when they are more than a few feet apart. At the same time, interpersonal distances less then a foot and a half might be too close. Goffman (1963) notes that, at closer ranges, interactants may try to compensate for the spatial closeness by looking away.

McDowell (1972) suggests somewhere near 97 centimetres (approximately 3 feet) as the average interaction distance, and around 48 centimetres (approximately 1.5 feet) as the violated, or encroached, personal space condition. Sommer (1962) recommends interpersonal distances ranging between 5 feet and 5.5 feet for face-to-face conversations. Efran and Cheyne (1973) found that outsiders are reluctant to walk in between interactants separated by a distance of about 48 inches (i.e. 4 feet). The most detailed account of interpersonal distances, however, is that of Hall (1964) and Hall (1966).

2.4.1 Classification of interpersonal distances

Hall proposed four different spatial regions centred around an individual's body: intimate, personal, social and public distance zones (Hall, 1966). We can think of the four distance zones as concentric regions of space extending outwards from an individual's body:

1. The *intimate distance* zone extends from the body sheath up to 18 inches outwards.

2. The *personal distance* zone starts at the farthest extent of the intimate zone and extends up to 4 feet.

3. The *social distance* zone ranges between 4 feet and 12 feet.

4. The *public distance* zone is the region beyond the 12 feet mark.

Each of these distance zones are further divided into a close phase and a far phase. In the far phase of the intimate distance (i.e. between 6 inches and 18 inches from the body), hands can reach out and grasp the extremities of each other's bodies. In the close phase of the intimate distance (i.e. full body contact up to 6 inches from the body), it is easy to touch each other. Intimate distance is considered to be too close for face-to-face conversations because of the greatly escalated sensory inputs, especially olfaction, thermal, and sound inputs. As the name suggests, intimate distances are almost exclusively used for intimate activities, and in some cultures, are considered inappropriate in public places (Hall, 1964).

Personal distance is the equivalent of personal space (Patterson and Edinger, 1987). The close phase of the personal distance, extending between 1.5 feet and 2.5 feet, is the minimum distance people expect to keep from one another. But the far phase of the personal distance, between 2.5 feet and 4 feet, is considered ideal for discussing subjects of personal interest and involvement. Likewise, the close phase of the social distance, between 4 feet and 7 feet (+ or - minus 6 inches at each end), is commonly preferred for attending casual social gatherings. Hall (1964) found that at this distance interactants tend to have a 60° visual access to one another, which allows seeing one another's head, shoulder and trunk.

The far phase of the social distance, between 7 feet and 12 feet, offers the advantage of permitting easy shifting back and forth between one's activity, and whoever else is in the room (Hall, 1964), but at this distance people struggle to converse comfortably (Hall, 1966). Lastly, the public distance signifies the space outside the circle of an individual's involvement. In the close phase of the public distance, between 12 feet and 25 feet from the body, participants cannot reach one another easily. This distance is mainly used for defensive or evasive purposes. The far phase of the public distance extends beyond 25 feet from an individual. Hall (1966) notes that 30 feet is the distance set around important public figures.

Figure 2.2 visualises the four proxemic zones. Despite Kendon's (1973) concerns of there being no systematic data to support Hall's (1966) findings, the study served as a precursor to several later investigations on spatial behaviour in face-to-face encounters, such as Watson and Graves (1966), Cheyne and Efran (1972), Efran and Cheyne (1973) and Deutsch (1977), among others.

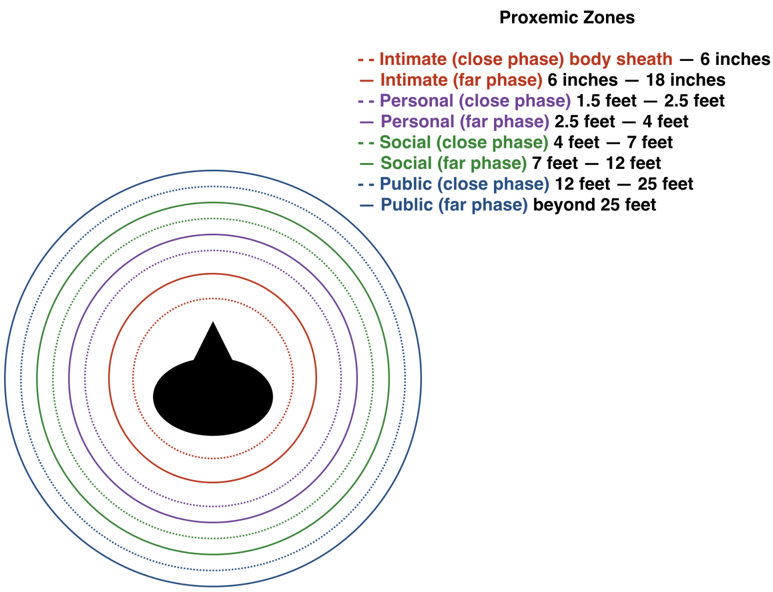


Figure 2.2: Hall's (1966) classification of spatial zones

2.4.2 Interpersonal distance as a cue to identify conversational groups

Hall's studies were primarily concerned with how interactants regulated the distance between their bodies, but not how others perceived these distances. Scheflen (1964) noted that even a mere glance at the relative positioning and orientation of individuals indicates to a great extent the nature of an encounter, and the kind of ongoing conversations. The experiments conducted by Cheyne and Efran (1972) and Efran and Cheyne (1973) sought to investigate how interpersonal distances in a dyadic conversational group were perceived by others. The aim of the experiments was to assess whether or not passers-by interrupted the dyads (i.e. by walking through the group) under different scenarios, where the gender composition of groups, the activity pursued within groups, and the spatial parameters of the environment were varied.

Cheyne and Efran's (1972,1973) experiments showed that fewer people intruded the group when the interpersonal distance between interactants was about 40 to 48 inches, i.e. the equivalent of Hall's personal distance. On the other hand, greater interpersonal distances (between 50 inches and 56 inches) led to more intrusions. The experiments also showed that interpersonal distances serve as cues to identify the nature of conversations between interactants, e.g. closer interpersonal distances were perceived as intimate conversations, which in turn influenced intrusion behaviour. The gender composition of groups also influenced how others perceived interpersonal distances. Mixed-gender dyads and female dyads were more effective in prohibiting intrusions than male dyads.

2.4.3 Factors influencing interpersonal distances

The factors influencing the interpersonal distances preferred for face-to-face encounters include: familiarity between individuals and their previous interactions (Patterson and Edinger, 1987); interactants' gender, age and the situations in which encounters occur (Becker, 1973; Scheflen, 1975); the cultural background of participants (Hall, 1964, 1966; Hall et al., 1968; Scheflen, 1975); and the dominance and social status of participants (Vine, 1975; Patterson and Edinger, 1987).

Hall found that the topic of conversation (Hall, 1964), and the feelings interactants have towards one another, also influences interpersonal spacing (Hall, 1966). He found that situational personalities, such as when interactants are angry or emphatic about the point being made, motivate interactants to move closer to the other person. Sussman and Rosenfeld (1982) and Patterson and Edinger (1987) also support the notion of situational personalities affecting interpersonal distances.

Lastly, and importantly, the practicalities associated with being able to use different sensory signals at different distances influences the choice of interpersonal spacing. For instance, touch is only possible when interactants are reasonably close, but voice can be used at comparatively longer distances (Goffman, 1963; Patterson and Edinger, 1987, p. 162).

2.4.4 Interpersonal distances preferred for face-to-face conversations

Combining Hall's findings with Efran and Cheyne's findings, we are able to deduce that the region between the personal distance and the close phase of the social distance is preferred by interactants, as well as being respected by others. This resonates with Vine's (1975) findings about the equilibrium distance being in between the personal and social spaces. A universally acceptable magnitude for the equilibrium distance is hard to establish due to the variety of factors capable of influencing the choice of interpersonal distances in different scenarios. The most evident finding is that interpersonal spacing is a form of behavioural relation, which allows interactants in conversational groups to be able to perceive one another, which in turn facilitates the exchange of communicative signals. It is also evident that standing too close to one another is not preferred for face-to-face conversations, except for intimate relationships. At the same time, standing too far imposes practical difficulties. Therefore, somewhere between standing too close and too far is an ideal interpersonal distance that interactants prefer for face-to-face conversations.

2.5 The Role of Body Orientation

In face-to-face encounters, after adopting suitable interpersonal distances, interactants will reorient their head and body with respect to one another. Kendon (1973) calls an individual standing alone a *solo*, but the moment he is joined by another person, he becomes a *with*. Scheflen and Ashcraft (1976) note that the process where interactants adjust the position and orientation of their bodies with respect to one another, which often precedes any form of verbal communication, is an essential step to affirm withness in conversations. Scheflen and Ashcraft (1976) also note that orientations held by an individual are *locked-in* in stationary tasks and interactions, and often held that way until a particular sequence of behaviour is completed. The process is referred to as *orientational hold*. In this section, we review the purpose and significance of interactants' body orientation in conversational groups. We will review the role of whole body orientations, as well as those of individual body parts, such as the head, body, and feet.

Scheflen and Ashcraft (1976) use the term *position* to denote the orienting of one or more body regions towards someone or something. They use the term *stance* to indicate the way in which a body region is deployed and framed by the hands in marking off an orientation. Deutsch (1977) uses the term *posture* to denote the body lean or stance adopted by an individual. Scheflen and Ashcraft (1976) also note that, both position and stance can involve one or more regions of the body, and can occupy and claim a *segment of space*, which extends beyond the body.

The spatial extent of segments are delineated by the way in which body regions are held. For a person standing with his feet on the ground, different regions of his body, such as the head, torso and feet, can be oriented such that each body region commands a well-defined segment of space. When all body parts are oriented in the same direction, they all claim the same spatial segment. Kendon uses the terms *transactional segment* (Kendon, 1990) and *use space* (Kendon, 2010) to refer to the segment of space marked by an individual's body position and orientation at a particular time.

Kendon (1990) notes that the position and orientation of the body, together with the spread of limbs, serves as a frame to delineate an individual's transactional segment. He further notes that transactional segments are delineated by the position and orientation of the lower body – to be more precise, the position and orientation of feet placement. He also explains that, although it is possible for each body region to be oriented differently, in face-to-face interactions, the head-body orientation are usually not misaligned by more than 45°. This is because, beyond the 45° limit, people have to strain their head and neck to face someone or something.

The purpose of an individual's transactional segment is that it serves as an immediately accessible spatial zone, where the individual can carry out their intended activity – this is also the reason for referring to it as a use space (Kendon, 2010). Figure 2.3 shows Kendon's (2010) illustration of the transactional segments claimed by an individual, while in pursuit of two different activities. To the left is an individual's transactional segment when watching TV and to the right is his transactional segment when writing at the table. As evident from the figures, an individual drawing while seated at a table commands a much smaller spatial segment, than when sitting at the couch and watching television.

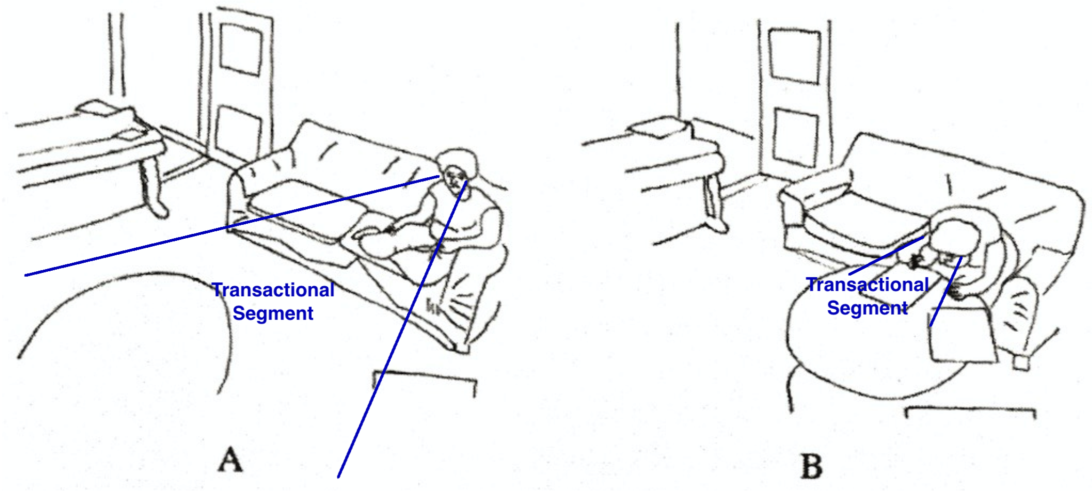


Image reproduced with kind permission of the publishers.
Figure 2.3: Kendon's (2010) representation of the spatial extent of transactional segments differing based on an individual's activity

Transactional segments are sustained for the entire duration of an activity. Just like safeguarding one's personal space, an individual endeavours to protect his transactional segment against intrusions (Kendon, 1990, 2010). Transactional segments may be poorly defined spaces, but when an individual's stance focuses and fixes an orientation, people avoid walking through it or gazing into it. In such cases, Scheflen and Ashcraft (1976) note that an individual's transactional segment becomes a territory.

Deutsch (1977) provides the most detailed formulation of an individual's transactional segment. He combined the findings of Hall (1966), Efran and Cheyne (1973) and Kendon (1978) to devise a trapezoidal shape transactional segment for individuals. Figure 2.4 shows Deutsch's (1977) formulation of an individual's transactional segment. The shorter of the two parallel sides of the trapezoid is called the *base* – imagine a line drawn by connecting the two perpendicular lines dropped from each shoulder of an adult human. The base is 2 feet long to accommodate the shoulder-to-shoulder width of a person. The side opposite to the base is the *crown* measuring 6 feet long. Deutsch (1977) conceptualised the length of the crown to extend two feet past either side of the base by taking into consideration Efran and Cheyne's finding that 2 feet is the closest distance passers-by come next to a conversing dyad. The width between the base and the crown is 2 feet long, i.e. half of the maximum extent of Hall's (1966) personal distance. Therefore, when two people arrive at a vis-à-vis arrangement such that their crowns are just merged, they would be spaced at 4 feet from one another (see figure 2.5a). It also possible for interactants to stand closer than 4 feet, in which case the transactional segments are more overlapped (see figure 2.5b). Kendon (1978) suggests that the maximum deviation between the head-body orientation of an individual is 45°, consequently, Deutsch (1977) formulated the non-parallel sides connecting the crown and the base to be lines drawn at 45° angles to the straight ahead. These sides, referred to as the *legs* of the transactional segment, are each 2 feet and 10 inches long.

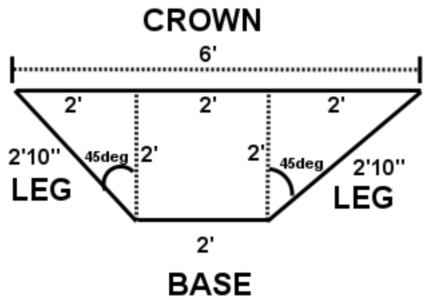


Figure 2.4: Deutsch's (1977)formulation of an individual's transactional segment

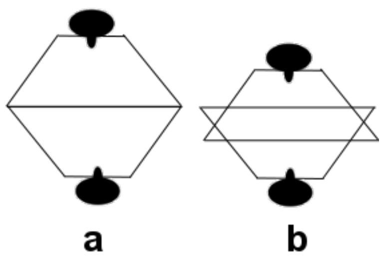


Figure 2.5: Transactional segments (a) just overlapping (b) greatly overlapping

Earlier it was noted that individuals strive to keep their transactional segments free of intrusions, but now we have described examples of overlapping transactional segments. This is because overlaps of transactional segments are avoided for individual activities, but for collaborative endeavours such as face-to-face conversations, interactants strive to orient their bodies towards one another, such that their transactional segments overlap (Kendon, 1990, 2010). When interactants adopt appropriate interpersonal distances, and point their torsos and faces towards each other, a *channel of space* gets established between them. This channel of space serves as the medium to transmit and receive communicative signals between interactants (Hall, 1966; Scheflen and Ashcraft, 1976). Overlapping transactional segments help in establishing the channel of space needed to facilitate the exchange of communication between interactants.

Deutsch (1977) provides an operational definition of the relative orientation between interactants as the number of degrees the frontal body plane of one individual needs to be turned to be parallel to the frontal body plane of another individual. Ciolek (1983) simply uses the term *angle* to refer to the relative orientation of interactants. Figure 2.6 shows the median and the frontal body planes of an individual. In face-to-face conversations, the relative orientation of interactants can extend from face-to-face to back-to-back. Of these, the range of angles between face-to-face and side-by-side, i.e. starting from about 0° between interactants to about 180°, are preferred for conversations (Argyle and Kendon, 1967).

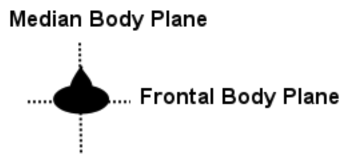


Figure 2.6: Frontal and median body planes of an individual

Figure 2.5a, where interactants are in a vis-à-vis (i.e. face-to-face) arrangement, is an example of 0° relative orientation between participants. If individuals are not facing one another vis-à-vis, then the imaginary line formed by connecting the non-overlapping edges of the crowns of their respective transactional segments, is referred to as a *free bonding site* (see figure 2.7). Deutsch (1977) notes that a free bonding site may be used to entertain new members willing to join a group.

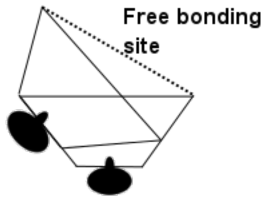


Figure 2.7: Overlapping transactional segments with a free bonding site

So far, the review has focused on the body orientation of individuals and the relative orientation of dyads, but in reality, conversational groups can involve more than two members. In multi-party conversational groups, there is usually a speaker, i.e. the one who has the conversational turn, a primary addressee, i.e. the intended recipient of a communicative signal originating from the speaker, and one or more listeners, who are also members of the conversational group. Interactants adapt their head and body orientation based on the conversational roles they play during face-to-face encounters. Head orientation (and eye gaze) is one of the principle ways in which a speaker denotes the primary addressee (Kendon, 1973). But at the same time, a speaker can sustain attention towards other participants (i.e. listeners), simply by splitting the orientation of other parts of their body (Scheflen, 1975), e.g. head and upper body can face one person, while lower body faces another.

Addressee(s) and listeners can also demonstrate their conversational roles using orientational cues, e.g. they tend to orient their head and neck towards the speaker, and hold it there until she finishes. If a new speaker begins talking, listeners then orient their head and neck towards the new speaker (Scheflen and Ashcraft, 1976). Furthermore, in multi-party interactions, each speaker-listener relationship sustains a dedicated channel of space. Kendon (1970, 1973) found that the characteristics of these channels varied depending upon the conversational roles of participants. He suggested that an *axis of interaction* gets established between a speaker and the primary addressee, to enable the uninterrupted exchange of communicative signals. Other listeners in the conversational group are referred to as *non-axial members*, because their orientation may be directed at the primary axis of interaction, but does not fall on the axis itself. Kendon (1970, 1973) also noted that non-axial members were able to engage in sub-conversational units in bigger conversational groups.

2.5.1 Factors influencing body orientation

A variety of factors influence participants' body orientation during co-present situations. The nature of an encounter, i.e. whether it is focused or unfocused, is a major influence. Kendon (1973) notes that, in focused encounters, head and body are aligned to orient towards and face one another more or less directly. Conversely, when two or more people are forced to stand in close proximity, Scheflen and Ashcraft (1976) suggest that they could demonstrate their non-association, or lack of withness, by orienting their bodies (or regions of it) in ways that discourage, break off, or avoid any involvement.

In general, co-present individuals could organise themselves in arrangements conducive of focused encounters, i.e. *sociopetal* arrangements. Or, they could be in *sociofugal* arrangements, which allows functioning alone, even when in the presence of others. Hall (1963) notes that vis-à-vis arrangements are sociopetal, while back-to-back arrangements are sociofugal. While sociofugal arrangements tend to pull people apart, and orient away from one another, sociopetal arrangements tend to pull them in and face one another.

Participants' status and dominance also influences their body orientation. Patterson and Edinger (1987) point out that equal status participants face one another more directly than those of unequal status. Mehrabian (1968) notes that for a speaker, there is more shoulder orientation towards a high-status addressee, than there is towards a low status addressee.

The stages in interaction also influence body orientation. Argyle and Dean (1965) and Kendon (1990) suggest that at the start of new conversational turns, participants tend to momentarily orient their head and/or whole body away from the rest of the group. However, over time, speakers gradually align their bodies to face others.

Personal relationships between people can also influence their body orientation. Mehrabian (1968) found that the relative orientation between interactants decreases as closeness increases. Scheflen (1975) notes that eye-to-eye gaze and vis-à-vis orientation at close interpersonal distances means an intimate relationship between interactants.

Scheflen's (1975) observation also indicates another important aspect: interpersonal distance and orientation work together in serving as indicators of participants' affiliation with one another. In this regard, any lapses in one form of behaviour will be compensated for by the other. For instance, Mehrabian (1968) found that when standing too close, people tend to turn away to decrease any feelings of uneasiness. Likewise, when people are forced to face one another more directly, unless they share an intimate relationship, they step away from standing too close to one another. Vine (1975) notes that, generally in face-to-face encounters, interpersonal distance varies with the orientation held by participants, and likewise, orientation varies with interpersonal distance. In fact, there are systematic means by which the two forms of spatial behaviour are intertwined. In the following section, we will review the means by which interpersonal distances and the relative orientation of interactants provide a structural basis for focused encounters.

2.6 Systems of Spatial Behaviour Organisation

In focused encounters, participants enter into and maintain a *working consensus* (Goffman, 1963; Argyle and Kendon, 1967; Kendon, 1973), which provides the rules for conducting interactions within the group. A key component of the working consensus is adherence to a system of spatial behaviour organisation, which allows participants to see and hear one another clearly, and gives them uninterrupted access to the activity occurring within the group. As Kendon (1973) notes, while such systems of behaviour organisation are not concerned with the actual exchange of words between interactants, they enable establishing and maintaining the conditions necessary for the exchange of words. The spatial-orientational behaviour relationship also allows interactants to exclude non-members from their conversational encounter.

Goffman (1963) uses the term *public order* to indicate the order that governs the conduct of persons when they are in the immediate presence of others. In focused encounters, such as in conversational groups, the positional-orientational relationship formed by interactants demonstrates their withness. Conversely, in unfocused encounters, the lack of a distinct positional-orientational relationship between individuals is what confirms the lack of withness. In this section, we review the spatial-orientational order by which interactants demonstrate their withness, and how compromising the order would suggest a lack of withness.

McDermott and Roth (1978) use the term *interaction machinery* to refer to the ways in which people organise their bodies for joint activities in focused encounters. We review two specific instances of interaction machinery – the F-formation system and the equilibrium system – whereby participants in conversational groups, and other focused encounters, organise their position and orientation with respect to one another.

2.6.1 F-formations

A *facing-formation* or *F-formation* is formed when two or more people adopt appropriate interpersonal distances and face one another, such that their individual transactional segments overlap, leading to the emergence of an overlapped space called the *o-space* (Deutsch, 1977; Kendon, 1990). The o-space is also called the *orientation space* (Deutsch, 1977; McDermott and Roth, 1978), *engagement space* (Gill et al., 2000), or a *shared transaction segment* (Kendon, 2010). Establishing an o-space is indicative of sustaining a joint focus of attention. It shows withness and the involvement of participants in a jointly pursued activity within a focused gathering (Goffman, 1963; Kendon, 1973; Scheflen and Ashcraft, 1976).

The overlapping transactional segments seen in figures 2.5a, 2.5b and 2.7 are examples of F-formations. Figure 2.8 illustrates, more clearly, an o-space emerging from the overlap of the individual transactional segments of three interactants. F-formations are readily recognisable social-spatial arrangements that can be found everywhere (McDermott and Roth, 1978), e.g. in the party hall, on the road, in the atrium of a mall, etc.

Just as individuals strive to protect their use space, the members of an F-

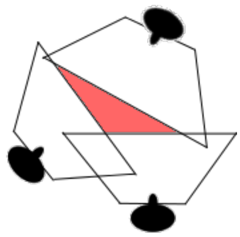


Figure 2.8: A triadic F-formation and the enclosed o-space (coloured spatial zone)

formation strive to protect their o-space from any kind of intrusion. They achieve this by facing one another, while having their backs turned towards others who are not members of their F-formation, e.g. figure 2.9 shows outsiders in the vicinity of a triadic F-formation. Goffman (1963) uses the terms *knots*, *clusters*, *face engagements* and *conversational circles* to refer to the likes of F-formations. Scheflen and Ashcraft (1976) use the term *connection units*. Both Goffman and Kendon note that face-to-face interactions could be the sole activity within F-formations, or that interactions can just be a part of other collaborative activities, e.g. playing cards, fist fights, group studies, etc.

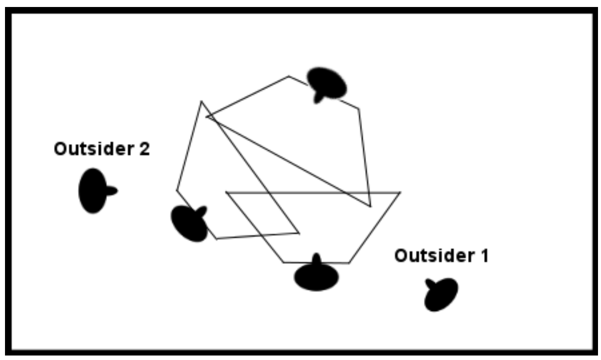


Figure 2.9: Outsiders around a triadic F-formation

Basically, any spatial pattern that people arrange themselves in during face-to-face encounters is referred to as a formation (Kendon, 1990), e.g. the circle of a conversational group, the line formation of a queue, the semi-circular arrangement of audience in a performance, etc. Among these, F-formations are those arrangements where the spatial-orientational relationship of participants leads to the emergence of an o-space. The spatial patterns established by the coming together of interactants in conversational groups are normally F-formations (Kendon, 1990).

Participants in an F-formation seek to establish and maintain the stability of their F-formation at all times (Kendon, 2010). Even if one participant in an F-formation alters their position or orientation, leading to the disruption of the o-space, others in the F-formation will readjust their position and orientation to compensate for the disruption (Goffman, 1961, 1963). For instance, Kendon (1990) provides an example of the spatial-orientational dynamics in a triadic F-formation, where the laughter of one interactant causes him to temporarily withdraw from contributing to the o-space. This in turn causes another participant in the F-formation to move in closer and restore the overlap of transactional segments, i.e. restore the stability of the o-space.

2.6.1.1 The F-formation system

The system of spatial-orientational behaviour organisation, which causes individuals to arrange themselves in F-formations, and as such, to sustain their arrangement for the entire duration of a collaborative activity, is referred to as the *F-formation system* (Kendon, 1990). During a face-to-face encounter, an F-formation system influences participants to adopt appropriate interpersonal distances and body orientations, such that they all have equal, direct, and exclusive access to the o-space formed by the overlap of their individual transactional segments (Deutsch, 1977; Kendon, 1990). Having access to the o-space allows the participants in an F-formation to be able to perceive one another more or less fully. Goffman (1963) uses the term *eye-to-eye ecological huddle*, to denote the formation of a spatial arrangement, whereby participants have maximum opportunity to perceive one another in an unobstructed manner.

An F-formation system is the means by which participants in a focused gathering preserve the integrity of their encounter. Kendon (2010) notes that an F-formation system serves to distinguish random assemblages of people, with no recognisable spatial patterns, from focused gatherings where a clear spatial pattern is sustained.

F-formation systems enable participants to contribute and cooperate to sustain the o-space (Goffman, 1963; Kendon, 1990), and to correct for any disturbances to the o-space (Deutsch, 1977) – an aspect referred to as the *maintenance of ar-*

rangements. As noted before, when a participant changes his position, orientation, or stance, the spatial layout of an F-formation changes. To keep the o-space intact in the face of changes, other participants in the F-formation compensate for the disruption, by readjusting their position, orientation and/or stance. Kendon (1990) indicates that participants' positional-orientational readjustments to sustain o-spaces may be coordinated about an equilibrial point, and that, whenever there is a shift in the equilibrial point, participants perform corrective actions to restore the o-space equilibrium. The spatial behavioural coordination involved in sustaining the equilibrium of o-spaces illustrates the notion of a *working consensus* described by Goffman (1963).

Condon and Ogston (1967) also remark how interactants in conversational gatherings tend to share patterns of changes in bodily movement in precise harmony, as if performing a dance. Gill et al. (2000) suggests that interactants do this to sustain the stability of the engagement space. Kendon (1970, 1973) re-emphasised Condon and Ogston's (1967) notion regarding *interactional synchrony*, a behaviour coordination mechanism, wherein the boundaries of a spatial movement wave from a listener coincides with the movement wave of a speaker. The earlier example in which the laughter followed by an outward movement of one interactant instigated an inward movement from another interactant, is an example of interactional synchrony (see section 2.6.1). Kendon (1970) pointed out that the coordinated spatial manoeuvres of interactants are anchored through their joint focus of attention manifested in the form of o-spaces.

While restoring the stability of o-spaces, it is possible that only a part of an F-formation system reacts to the changes. Lyman and Scott (1967) notes that in the event of perturbations, F-formation systems can re-bounce or restore their stability in a quick and prompt manner, due to the temporary dependence between parts. That is, whenever the system is confronted by changes, the entire system does not have to respond to it, unless the changes are huge enough for the whole group to react. Where changes affect only one part of the system, only members at that end of the system need to alter their position and orientation to compensate for the deviations. Deutsch (1977) notes that there are three main advantages to only a section of members reacting to perturbations: (1) it reduces the number of participants that must enter into the full duration of field change; (2) creates a condition in which other participants, not actively engaged in the readjustment, to focus on continuing the ongoing conversational exchange, or collaborative activity, without any pauses; and (3) it anchors the present stability of the system, so there is an anchor point for the restoration phase.

2.6.1.2 Conditions for F-formation arrangements

In F-formations, when people position their body parts (especially their lower bodies) at an angle less than 90° from one another, they are said to be mutually oriented towards one another. If, in an F-formation, adjacent pairs are oriented towards each other, then the spatial configuration of such an F-Formation is considered to be a *closed arrangement*. On the other hand, when the angle subtended between the orientations of some participants is more than 90°, an F-formation is considered to be an *open arrangement*. The latter category is also referred to as co-orientation arrangements that have u-spaces instead of o-spaces(Scheflen and Ashcraft, 1976).

Kendon (1973) notes that in focused gatherings, where a common focus of attention is maintained, participants tend to stand close to and face each another, such that the angle through which their heads would have to rotate from their current orientation while speaking is less than 90°. Deutsch (1977, 1979) also stipulate that, for a conversational group's spatial pattern to qualify as an F-formation, each individual's transactional segment must overlap with at least one other person's transactional segment in the group. Furthermore, the base of the transactional segment of any one interactant must not be more than 90° out of phase with the base of the transactional segment of at least one other interactant in the group.

2.6.1.3 Spatial domains of F-formations

There are three regions in an F-formation: the o-space, the p-space, and the r-space (Kendon, 1990). The o-space, as discussed before, is the core region, at the centre of the spatial arrangement of an F-formation. The existence of F-formations rely on the existence of the o-space, so participants strive to preserve the integrity of the o-space at all times. The *p-space* is where participants stand along with their personal belongings. Ciolek and Kendon (1980) indicates that the p-space roughly accommodates participants' body depth, i.e. about 45cm to 65cm. Entering into the p-space of an F-formation system ensures access and rights to the o-space (Kendon, 1988). Beyond the p-space is an area referred to as the *r-space*, which acts as the buffer zone between an F-formation system and the rest of the world (Kendon, 1988, 1990).

The r-space establishes the influence of an F-formation beyond the extent of the p-space (Kendon, 1990). It does this by governing the behaviour of potential new members, passers-by and associates (Kendon, 1988). *Potential new members* are those who wish to gain entry into an existing F-formation system. One of the ways to gain entry into an F-formation is to arrive at the r-space, and wait there, as a means of expressing willingness to join the F-formation (Deutsch, 1977). *Passers-by* are those who walk past F-formations en route to their destination. In an experiment, Efran and Cheyne (1973) found passers-by waiting outside a dyadic F-formation, awaiting their turn to squeeze through a narrow strip of space around the F-formation, instead of walking through it. In this sense, an r-space is seen as the region in an F-formation which passers-by use in a limited fashion, for purposes such as navigation, but not for stationary activities (Ciolek and Kendon, 1980). Lastly, those who are not directly engaged in an F-formation, but are positioned around it, and intermittently participate in it, are identified as *associates* of the F-formation (Kendon, 1990). Associates tend to stand in the r-space.

Ciolek and Kendon (1980) proposed two further regions around the r-space of an F-formation. The first one, referred to as the *b-space*, was proposed purely for want of a region that is distinctly different from the three regions that are central to F-formations, i.e. a truly public space in that sense. A further region, referred to as the *c-space*, was proposed as a transitional zone between the b-space and the r-space. The intention was for the c-space to serve as a boundary zone, which helps participants in the F-formations to pay selective attention to outsiders. Only those that cross the boundary of c-spaces are likely to be potential new members, passers-by, or associates. Hence, participants only need to pay attention to those that progress past the c-space, while others beyond this region can be ignored as outsiders in a true sense.

Figure 2.10 represents the different spatial zones within and immediately around F-formations, starting from the most private o-space, and extending outwards to the least private, or the most public, b-space. In addition to the five zones, Ciolek (1983) added an *a-space*, to signify a region that is outside the normal unaided sensory processing of an F-formation system.

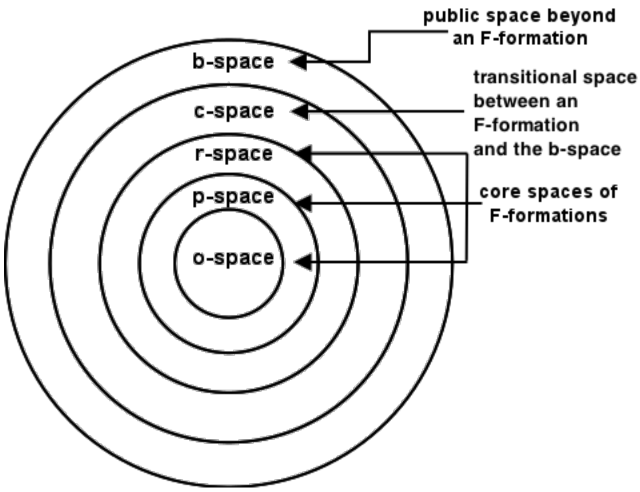


Figure 2.10: The spatial zones within and immediately around F-formations

2.6.1.4 F-formations are temporally regulated systems

F-formations are temporally regulated systems. At the end of a face-to-face encounter, members leave, and the F-formation comes to an end. Passers-by then have the freedom to walk through the spatial region that was previously claimed by the F-formation. McDermott and Roth (1978) note that facing-formations have clear beginnings and ends. An F-formation begins when the first of its members (two or more people) establish an o-space (Scheflen and Ashcraft, 1976; Deutsch, 1977). It then exists for as long as the o-space exists, resisting any changes in participation, such as new members joining or existing members departing.

Lyman and Scott (1967) also support the notion that conversational encounters are temporally regulated systems – they note that participants establish an interaction territory just for the duration of an encounter. Goffman (1963) remarks that the temporal duration of face engagements depends on the primary activity carried out within the groups. He suggests that, with a clearly defined task, the duration of face engagements can even last for many hours, while on the other hand, in the absence of a focused activity, face engagements may be as short as exchanging greetings.

Furthermore, F-formation systems are units of behaviour organisation at the level of interactions, and not at the level of individuals (Kendon, 1990). Consequently, an F-formation ceases to exit only when its last member renounces their position and the orientation from which they contributed to the F-formation system (Scheflen and Ashcraft, 1976; Deutsch, 1977; Kendon, 1990). Usually all members contributing to an o-space leave together, thereby drawing the F-formation to a close (Deutsch, 1977, 1979). However it also possible that a single individual remains, although, in such cases, the last person would not continue to stay for very long.

2.6.1.5 Leaving behaviour in F-formation systems

It is possible for members to leave while F-formations are still in progress. F-formation systems are highly stable spatial-temporal units, where even slight changes in the position and orientation of one participant, instigates compensatory moves from other participants. But when a person is planning to leave the F-formation, their spatial manoeuvres need no longer be compensated for. To that end, participants announce their intention of withdrawing from an F-formation using elaborate farewell mechanisms or departing procedures (Goffman, 1963).

Deutsch (1977) indicates that, once the spatial configuration of an F-formation is established by its participants, those that leave the arrangement go through a "rite of departure". The following is one form of leave taking ritual that participants follow when leaving F-formations (Kendon, 1990). Participants move back a few steps, then move further forward than their normal position to engage in salutations or departing greetings, and then walk away from the o-space. After walking away a few steps, they might change the direction of their movement to go to the desired exit point. Kendon (1979) identified manoeuvres such as withdrawing the body from the p-space, looking away, and momentary preoccupations with foci that were never a part of the central focus of attention, as indicators of leaving behaviour.

Deutsch (1977, 1979) identifies a three stage process for departure from F-formation systems. The *out-phase* or *preparation phase* is where a participant notifies their intention to leave the F-formation system. From this point onwards, other members in the F-formation stop compensating for the individual's spatial-orientational manoeuvres. The *in-phase* or *confirmation phase* is where the member who is about to leave the group steps in closer than their current position in the p-space, for any parting salutations. Lastly, the *away phase* or *finalization phase*, is where the member has left the group permanently.

The same three phases apply even if more than one individual, or all members of an F-formation, decide to leave at the same time. The preparation phase involves participants notifying one another regarding their decision to leave and the confirmation phase involves participants moving closer than before to exchange parting salutations. Lastly, the finalization phase involves participants coming out of the spatial-orientational arrangement they are presently in, and beginning to walk away. Thus, an F-formation system comes to an end. Deutsch (1979) notes that, while the exact steps followed may not be the same across cultures to end F-formations, each of them will have their own form of parting rituals for formally ending F-formations of face-to-face encounters. Goffman (1971) remarks that individuals who leave an F-formation using an elaborate farewell mechanism prior to the official closure of the F-formation will hesitate to return to it any time soon, to avoid embarrassment.

2.6.1.6 Significance of F-formation systems

F-formation systems signify the spatio-temporal relationship between interactants. Deutsch (1979) proposes that the space-time relationship of interactants in F-formations are of the following nature. Participants adapt their position, orientation, and posture with respect to one another, such that, given a constant set of environmental parameters, they maintain the same spatial-orientational arrangement, by compensating for the lapses caused by each other's movements. Participants also temporally coordinate their position, orientation and posture with respect to one another, such that the movement waves of one participant coincides with the movement waves of another participant (cf. section 2.6.1 for the laughter example).

F-formation systems are considered central to the functioning of face-to-face encounters. McDermott and Roth (1978) note that, irrespective of who the participants are, what they do, or what the consequences of their behaviour are, participants have to negotiate the issues associated with occupying the same space at the same time. McDermott and Roth suggest that F-formation systems are interactional machinery that people use to resolve issues pertaining to co-presence. Ciolek and Kendon (1980) and Kendon (1977, 1990) further add that F-formation systems provide:

1. A unit for analysing the organisation of behaviour in face-to-face encounters.
2. A unit that enables equal participation of all interactants.
3. A favourable network for the unobstructed exchange of communication signals, because F-formations provide a clearly delineated boundary, which helps participants shun the outside world, and focus solely on their joint activity.
4. A means of distinguishing between interactants and outsiders.
5. A primary interaction system that allows interactants to establish secondary systems of interaction, such as speech, gaze, gesture, etc.
6. A socially significant spatio-temporal arrangement within a physical environment. Studies focusing on the social use of physical spaces could assess the potential of the environment to support the functioning of F-formation systems.

2.6.1.7 Spatial behaviour of associates and outsiders

McDermott and Roth (1978) recognise that a conversational gathering establishes two distinct categories of people: those that are directly involved in the gathering and those that are beyond it. Knowles (1973) and Kendon (1990) have described that the spatial boundary of a conversational group serves as a marker, to distiniguish between people directly involved in the functioning of a group, and others who are external to the group. Goffman (1961) indicates that individuals exhibit different spatial behavioural traits depending upon their membership in, or proximity to, a focused gathering. In the immediate presence of non-members such as newcomers, associates and by-standers, face engagements become accessible, i.e. susceptible to intervention (Goffman, 1963). Consequently, both participants, as well as others in the immediate surroundings, who may be part of the social gathering within which the encounter occurs, cooperate to maintain the spatial and social integrity of a focused encounter (Kendon, 1988).

Kendon (1990) identifies a participant, or a member, in an F-formation as someone whose transactional segment falls over the o-space, but without intersecting the body of another individual. Participants in an F-formation system actively perform the spatial-orientational readjustments required to maintain the integrity of o-spaces. Outsiders to an F-formation also adapt appropriate forms of spatial behaviour to demonstrate their non-association in the focused encounter (Scheflen and Ashcraft, 1976). They do it as a means of expressing basic social manners and courtesy (Lindskold et al., 1976).

Different categories of outsiders demonstrate different forms of spatial-orientational behaviour around F-formation systems. As noted before, potential members are those interested in joining an ongoing F-formation (see section 2.6.1.3). A potential member can gain entry into an F-formation system using the following procedure. The individual arrives at the r-space and waits there. Kendon (1990) refers to this as waiting at an *outer position*. He also notes that the exchange of utterances or gestures from an outer position serves as access requesting, or access granting, mechanisms.

To allow a new member to join a group, the crown of the transactional segments of one or more existing members becomes non-overlapped, thereby creating a free-bonding site. E.g. imagine transitioning from the arrangement in figure 2.5b to the arrangement in figure 2.7. This flags an indication for the new member to join the group at the free-bonding site. When participants change their position and orientation to create a free bonding site, they are keen to maintain at least a partial overlap of their transactional segments. They do this in order to maintain the integrity of the conversational occasion even as the group undergoes membership changes (Deutsch, 1977).

Goffman (1963) also describes the situation where existing members of an F-formation system invite a new member to join the group. As the arrival of the new member is pre-acknowledged, existing members begin to readjust their position and orientation even as the new member starts walking towards the group. By the time the new member arrives, there is already a free-bonding site from where the new member can join the F-formation. The new member could then directly walk into the F-formation without waiting in the r-space. Once in, the new member also gets equal, direct, and exclusive access to the o-space. On the other hand, if the group does not wish to include a new member, they can simply ignore the arrival of the new member, and continue their ongoing activity in a closed manner (Lyman and Scott, 1967).

If an existing member leaves a F-formation only briefly, their place is usually not taken or given up, so they would be able to walk back in without the need for any explicit rejoining rituals. However, if the F-formation has undergone changes in their absence (either in shape, or in members, or both), then they will have to wait at the outer position to be formally accepted into the group again. If for any reason, one temporarily leaves the conversational group, they do so after saying “excuse me!” or using a hand gesture, as means of notifying other participants about their temporary exit, and to ensure that co-participation will soon resume (Goffman, 1971).

Passers-by are those who navigate around (or through) existing F-formations (see section 2.6.1.3). The experimental study by Efran and Cheyne (1973) found that passers-by who arrived outside a dyadic F-formation awaited their turn to squeeze through the r-space, instead of walking through the F-formation. Likewise, Knowles (1973) found from an experimental study that the size of a group, and the status of its members, affects the permeability of the spatial boundary of the group. A bigger group with more members is relatively impermeable when compared to smaller groups. Similarly, it was also found that passers-by of higher status are more likely to permeate interactional territories. Conversely, the higher the status of members in conversational groups, the less likely it is for others to permeate the spatial boundaries of their group. Lindskold et al. (1976) also found that permeability of a conversational group depends on the nature of conversations and the interactional relationships between participants.

Associates are people intermittently associated with the proceedings of an F-formation (Kendon, 1990). Consequently, they tend to demonstrate selective attention and contribution to the maintenance of o-spaces. Goffman (1963) identifies bystanders as unratified members of an F-formation. Bystanders do not gaze or manoeuvre into o-spaces unwarrantably. The only exceptions to unwarranted entry are non-persons such as kids, servants, pets, and courtiers, who are assumed not to influence the functioning of focused gatherings (Goffman, 1963, p. 40).

2.6.2 The Equilibrium Theory

In this section we review the equilibrium theory, which, like the F-formation system, presents a means of spatial behaviour organisation in focused encounters. The *equilibrium* state is defined as the stable state maintained by a group of interactants as a result of adjusting their proximity, orientation, and gaze behaviour with respect to one another in specific ways (Argyle and Dean, 1965; Argyle and Kendon, 1967; Ciolek, 1983). Similar to the notion of an o-space in an F-formation system, it is noted that members strive to establish and sustain equilibrium in conversational groups, as a means of demonstrating their working consensus. Consequently, anything that tends to change the pattern of equilibrium in a conversational group causes interactants to adopt compensatory moves aimed at restoring the equilibrium. Patterson and Edinger (1987) suggest that sustaining equilibrium is an adaptive mechanism that accounts for compensations to inappropriate arrangements undertaken by interactants. They also note that such adaptive mechanisms are aimed at achieving homoeostasis within conversational groups.

Argyle and Dean (1965) and Argyle and Kendon (1967) identified that, between a pair of interacting individuals, there is an invisible but conversationally perceived equilibrium point, which is a function of variables such as eye contact, physical proximity, and the intimacy of the topic of conversation. It is noted that interactants endeavour to maintain each of these behavioural aspects at optimal levels throughout their interactions. For instance, interactants prefer to stand neither too far nor too close, but at the optimal distance necessary for the smooth exchange of utterances. As described before, Hall (1966) identified four distance zones, which cater to different types of interactions. Among these, an equilibrium position that works for a conversational pair is exclusive to that particular pair, i.e. depends on what distance settings work for a particular conversational pair on a particular social occasion. If positioned too far from the ideal distance, interactants will move closer, whereas if too close, they will step away from one another. This readjustment process will continue until the equilibrium state is re-established.

If interactants are unable to readjust their interpersonal distances for any reason (e.g. it may be a crowded situation), then other variables contributing to the equilibrium will take over. For example, Argyle and Dean (1965) suggest that people can adopt suitable eye-contact mechanisms to compensate for any lapses in interpersonal distances. When standing too close people tend to look away to compensate for the more than desired physical proximity. Argyle and Dean (1965) also clarify that such a compensation does not mean that eye-contact mechanisms act as a substitute to interpersonal distances, rather, in the event of any one variable changing drastically, other variables will over compensate in the reverse direction to re-establish the equilibrium position.

Argyle and Dean (1965) suggested that the readjustment process, aimed at maintaining equilibrium in conversational groups, makes use of approach and avoidance forces. It is also noted that at any moment, a summation of these forces results in a state of equilibrium, which acts as a theoretical value of the desired physical closeness between interactants (von Cranach, 1971). Such an equilibrium is susceptible to initial conditions. For instance, Argyle and Dean (1965) found that, if interactants stood in close proximity at the start of conversations, then their interaction had much less eye-contact even when spacing increases during the course of interaction.

They suggest two possible reasons for this: either there is a separate equilibrium for each contributing variable such that they cannot fully compensate for one another, or different variables have different levels of influences on the state of equilibrium.

As explanations of the spatial-orientational organisation of face-to-face encounters, the F-formation system and the equilibrium theory have some similarities, and some differences. The F-formation system is principally concerned with sustaining the stability of the o-spaces, while the equilibrium theory is concerned with sustaining an equilibrium within conversational groups. On the other hand, the former does not conceptualise stable o-space manifestations as a tension between opposing forces, whereas the latter explicitly does so. The possibilities of combining the two theories to explain the spatial configuration of conversational groups is yet to be explored.

2.7 Different types of Spatial Arrangements of Conversational Groups

Kendon (1973) explains that the stable spatial-orientational arrangement of interactants' bodies serves to delineate the spatial boundary of a conversational encounter. In this section, we will review the different types of spatial arrangements of conversational groups, delineated by the arrangement of interactants' bodies in space.

Kendon (1990) argues that dyadic interactions assume one of the following spatial arrangements: *vis-à-vis*, *L-shaped* or *side-by-side* arrangements. Vis-à-vis arrangements means people stand facing one another almost directly. L-shaped arrangements means interactant's bodies appear to be aligned along the two arms of the letter L. Side-by-side arrangements are formed when people stand facing in the same direction.

Ciolek and Kendon (1980) propose three further categories of dyadic arrangements: the *N-shaped* arrangement, the *C-shaped* arrangement and the *V-shaped* arrangement. Ciolek and Kendon originally proposed six different types of dyadic F-formations, where the I-shaped arrangement and H-shaped arrangement denote the side-by-side and vis-à-vis arrangements, respectively. The L-shaped arrangement is the same as the one proposed by Kendon (1990). In addition to these, the N-shaped arrangement signifies interactants standing facing one another, such that their body planes are parallel and displaced by approximately half a body width. The V-shaped arrangement signifies interactants standing facing one another such that the angle subtended by the intersection of their frontal body planes outside the formation is approximately 45°. The C-shaped arrangement signifies interactants standing next to one another such that the angle subtended by the intersection of their frontal body planes is about 135°. Figure 2.11 shows the six types of arrangements of dyadic face-to-face encounters. The N, H and V-shaped arrangements are considered as closed arrangements, and the L, C and I-shaped arrangements are considered as open dyadic arrangements.

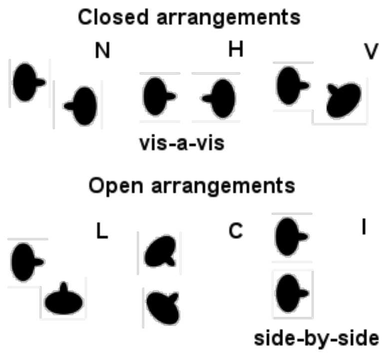


Figure 2.11: Ciolek and Kendon's (1980) six types of arrangements of dyadic F-formations

Watson and Graves (1966) proposed a differently organised, but related, schematism of the spatial arrangement of dyadic interactions. Figure 2.12 shows the dyadic spatial arrangements, originally proposed by Hall (1963), but later reorganised in a linear axis by Watson and Graves (1966). Of these arrangements, Hall (1963) identified arrangements 0, 1, 2, 4 and 8 as being commonly used in face-to-face interactions. Arrangement 0 tends to be used for direct communication, where participants intend to reach one another with maximum intensity. Arrangement 0, which is a face-to-face arrangement, is also called as a *sociopetal arrangement*. Arrangement 2 is a more casual or less involved conversational engagement. Arrangement 4 is suitable when interactants are watching or discussing about an external focus of attention. Hall (1963) also argued that arrangement 6 is used as a mechanism for disengaging oneself from an ongoing conversational encounter. Lastly, a back-to-back arrangement, also called as a *sociofugal arrangement*, signifies that participants choose to stay conditionally involved, while they remain uninvolved at other times (Hall, 1966).

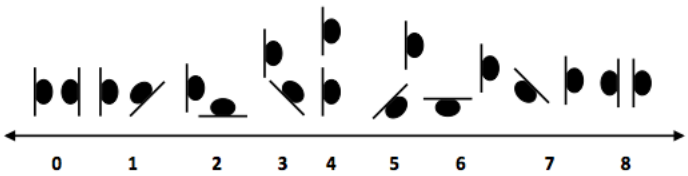


Figure 2.12: Watson and Graves's (1966) classification of dyadic arrangements along the sociofugal-sociopetal axis

The spatial arrangements of multi-party conversational groups have not been explored in as much detail as those of dyadic groups. Only a few general observations have been made. Scheflen (1975) proposes that multi-party conversational groups may be circular, triangular, or square shaped. Earlier, Scheflen (1964) reported that, given the freedom to move around the furniture in a room, freestanding conversational groups will have interactants arrange themselves in the form of a circle. Kendon (1990) also agreed that multi-party conversational groups comprising of three or more people tend to assume circular arrangements. At the same time, he also stressed that a circular arrangement is not the only possibility for multi-party conversational groups.

Kendon (1973) proposed that groups with equal participation rights for all interactants tend to assume circular shapes, whereas groups with non-equal participation rights assume parallelogram, triangle, semi-circular, or oval shapes. He also noted that groups with non equal participation rights are likely to have a *head position*, where the member with maximum participation rights will be located.

Scheflen (1964) suggested that dyads could assume parallel body orientations, while triads may be a parallel orientation dyad facing a third vis-à-vis, and quads may be two parallel orientation dyads facing one another vis-à-vis.

2.7.1 Influence of the environment on the spatial arrangement of conversational groups

The characteristics of the physical environment (i.e. the physical space) where the groups function affects the spatial patterns of conversational groups. Kendon (1973) uses the term *site* to refer to the well-defined physical space where conversational groups occur. The positional-orientational arrangements of interactants, which determines the spatial layout of conversational groups, are affected by: (1) whether sites are open, semi-open, or closed spaces, (2) the presence of furniture and other artefacts in sites, and (3) the presence of other individuals in sites, e.g. passers-by, associates and bystanders (Hall, 1966; Kendon, 1973; Vine, 1975; Ciolek and Kendon, 1980).

Marshall et al. (2011) found that interactants tend to stand closer to one another in open spaces than in closed physical spaces. Ciolek and Kendon (1980) found that, even with less pedestrian traffic, interactants preferred closed spatial arrangements in relatively open spaces.They also found that dyads preferred L-shaped arrangements for F-formations in the middle of a sidewalk, while different types of open arrangements were preferred when standing close to walls. In general, the degree of openness of the spatial arrangement of conversational groups are inversely proportional to the openness of the physical setting, where encounters occur (Ciolek and Kendon, 1980).

Goffman (1963) noted that in crowded open sites, where conversational groups are more susceptible to intrusions, interactants prefer open arrangements, or semi-open patterns, which allows them to monitor the happenings in their environment. In order to avoid intrusions, Kendon (1973) suggests that conversational groups tend to withdraw themselves into a room, or to any other private space. In the event of focused gatherings happening in open spaces, others avoid walking through the group. These avoidance habits are acquired as part of growing up in any society that has its own social norms.

Knowles (1973) notes that, while crowds attract passers-by, interactional units such as conversational groups tend to repel passers-by. Conversely, it is also expected that conversational groups do not function in spaces ordinarily used for other purposes, such as pedestrian traffic. This is one reason why corridor conversations are usually short (Kendon, 1973), while on the other hand, conversational groups functioning within well-defined spaces are subject to fewer intrusions, and sustained for longer durations (Ciolek and Kendon, 1980). Marshall et al. (2011) also found that the duration of conversations is shortened in confined physical spaces. Cheyne and Efran (1972), Efran and Cheyne (1973) and Ciolek and Kendon (1980) also found that interactants tend to stand close, facing one another at an angle of about 90°, while interacting in narrow spaces, such as corridors and sidewalks.

Several factors influence the spatial patterns of conversational groups even within closed spaces. Sommer (1962) found that the size of the room influences the spatial closeness of the arrangement of chairs in the room. He found an inverse relationship between the two, i.e. the bigger the size of the room, the closer the arrangement of chairs. Marshall et al. (2011) found that the shape and height of furniture and artefacts at a tourist information centre influenced the size (i.e. number of interactants) and spatial configurations of conversational groups.

Vine (1975) suggested that the boundaries of conversational groups are more explicitly defined when using semi-fixed features, like chair arrangements. Kendon (1990) has also made the same point – for seated conversations, the layout of furniture determines the spatial arrangement of interactants. Even in seated arrangements, interactants have the freedom to adjust their orientation, as appropriate to the context of their interaction. For example, if the arrangement of furniture allows interactants to sit within reasonable extents of interpersonal distances, they prefer to sit across from one another at a certain angle (Steinzor, 1950; Sommer, 1962). But when forced to sit at distances beyond the acceptable range, Sommer (1962) found that interactants preferred sitting side-by-side, rather than face-to-face, to compensate for the longer interpersonal distance.

Sommer (1965) also found that dyads preferred corner seating for casual conversations, because it allowed them to be close to one another, while avoiding direct eye contact all the time. He also found that the openness in the arrangement allowed interactants to pay attention to other events occurring in the environment. Furthermore, he found that co-operating pairs preferred sitting next to one another in side-by-side arrangements, because it made sharing things easier, while competing pairs preferred sitting in vis-à-vis arrangements, because it enabled complete visual access to one another's ongoing activities.

Just as the environment influences the spatial arrangement of F-formations, the existence of F-formations also influences the environment. Kendon (2010) suggests that there is a two-way relationship between F-formation systems and the physical space within which F-formations occur. For example, a corridor is a physical space, which people use to commute between two ends. But when an F-formation system occurs in a corridor, it immediately affects the flow of pedestrians through the corridor. Passers-by no longer walk in the usual trajectory that they would take in the absence of F-formations. Instead, they pause just outside the F-formation, then either squeeze around the group, or slow down, duck their heads and walk through the group.

2.7.2 Other factors influencing the spatial arrangement of conversational groups

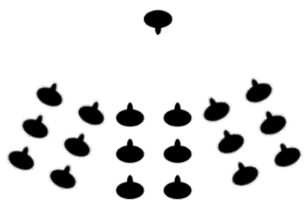
The spatial arrangement of F-formations are also influenced by the activities pursued within gatherings. Knowles (1973) suggests that, when the activity pursued is inclusive in nature, the spatial boundary of a gathering is more permeable, than when the activity performed is exclusive in nature. He provides the example of a group of members jointly looking at an object or artefact as an example of an inclusive activity, while engaging in face-to-face conversations is an exclusive activity.

Scheflen and Ashcraft (1976) also make an observation that there are two ways in which participants can sustain an orientation space: (1) by merely orienting towards the same focus of orientation to co-act and behave in unison, or (2) by coming together to form links and orient towards one another to demonstrate mutual involvement. The latter leads to closed configurations, whereas the former leads to comparatively open spatial configurations.

Kendon (2010) found that semi-circular open arrangements are formed when there is an external focus of attention. He gave the name *horseshoe-shaped* arrangement for such a configuration (see figure 2.13a). Kendon (2010) also used the term *pyramid formation*, to signify the likes of students facing a lecturer standing at the front of the classroom, as if they were all focusing at the apex of a pyramid (see figure 2.13b). Kendon (1973) also found that the level of involvement between interactants influences the spatial arrangement of conversational groups. If there is utmost mutual involvement, interactants prefer intimate or personal distances and a vis-à-vis arrangement, whereas at the social-consultative distance, a vis-à-vis arrangement does not reflect as much of a mutual involvement between interactants.

The spatial arrangements of F-formations are also affected by the stages of a conversation. At the initial stage of a conversation, Kendon (1990) suggests that interactants adopt a vis-à-vis arrangement, for the purpose of exchanging greetings. And then, as conversations progress, interactants would shift to other arrangements, such as the L-shaped, V-shaped, or other multi-party arrangements.

(a) Horseshoe-shaped F-formation



(b) Pyramid-shaped F-formation

Figure 2.13: F-formation shapes proposed by Kendon (2010)

2.7.3 Adaptation of the spatial arrangement of conversational groups

The spatial arrangement of conversational groups can change throughout the course of an encounter. Deutsch (1977) suggests that participants use a *probe-proposal-acceptance* strategy to organize and reorganize the spatial arrangement of their bodies during face-to-face encounters. The *probe phase* signifies the initiation of an arrangement by any one participant in the group. The *proposal phase* involves the initiation of changes to the spatial arrangement of an existing group. The *accep-*

tance phase seeks to gain the approval of members for a modified spatial arrangement of the group, resulting as a consequence of the proposal phase. In summary, the same F-formation can assume different spatial patterns at different times, e.g. an F-formation can start as a vis-à-vis arrangement while dyads exchange greetings, then turn into an L-shaped arrangement, and then become horseshoe-shaped as more members join the group (Kendon, 1990).

2.8 Deviations and Reactions to Deviations in Spatial Behaviour

Deviant situations are when usual expectations of spatial behaviour are not observed, or are temporarily compromised. In the event of deviant spatial behaviour, individuals, interactants, and where appropriate even outsiders, undertake corrective actions to restore the stability of the situation.

Argyle and Dean (1965), Felipe and Sommer (1966) and Hayduk (1978) suggest that unwarranted intrusions into personal space are responded with flight, withdrawal, or avoidance behaviour. McDowell (1972) identifies maintaining a frozen posture, avoiding touching, and shifting to a more comfortable location as common avoidance reactions. Other avoidance reactions include: drawing in of the hands and arms towards the body, avoidance of eye contact, turning both head and shoulder in a direction facing away from the intruder, and using physical objects such as books, purses and coats as barriers between oneself and the intruder (Felipe and Sommer, 1966; Scheflen, 1975; Patterson and Edinger, 1987).

Lyman and Scott (1967) propose three types of disruptions to personal territories – violation, invasion and contamination. They also propose three types of reactions to such disruptions – turf defense, insulation and linguistic collusion. *Violation* refers to the unwarranted use of a territory, *invasion* refers to the likes of trespassing, and *contamination* refers to situations where the spatial integrity of interactional territories have been compromised. The compensatory moves, *turf defense* indicates responses initiated when an intruder cannot be tolerated at all, *insulation* indicates the placement of barriers to avoid entry into egocentric territories, and *linguistic collusion* indicates the mechanisms adopted to explicitly affirm the non-association or the "outsider" status of territorial invaders.

Goffman (1971) also identifies three elements involved in the territorial disruption-restoration process. First is a consideration of the roles of the offender, victim and offence. *Offender* is one that does not exercise control over his body or intent, *offence* is over stepping into the personal territory of another individual, who becomes a *victim* of the offence. The second element of Goffman's (1971) disruption-restoration process is the ritual work that is performed in the offensive situation, e.g. offering an apology, which is then accepted. Last is the deed, which is the actualisation of the offence, and the ritual work performed to ameliorate the worst implications of the offence.

Invasions into interactional territories invite almost similar defense and avoidance mechanisms. Ford et al. (1977) suggest that, when interpersonal spaces are violated, people are left with a cognitive impression, wherein the perceived interpersonal distance seems significantly smaller than the actual interpersonal distance. Consequently, in the event of territorial compromises, interactants could react dramatically to restore the ideal interpersonal distance. Efran and Cheyne (1973) found that walking in between interacting dyads is seen as a violation of social norms and customs.

Lindskold et al. (1976) found that the status of interactants, their age, gender, group size and activity are factors, which govern the amount of spatial intrusions that conversational groups become subject to. Cheyne and Efran (1972) found that, the closer the interactants stand with respect to one another, the less likely they are to be intruded on by outsiders. They also found that, in situations where it is evident that talk lines are open between interactants, or in other words, where it is evident that two or more people are actively conversing, it is less likely for the group to be invaded. Knowles (1973) found that intrusions decrease with an increase in the group size, i.e. bigger groups are confronted with fewer intrusions than smaller groups.

Invading interactional territories does not just affect interactants, it also affects intruders. As noted before, Efran and Cheyne (1973) found that, when passers-by arrived next to dyads interacting in a narrow corridor, they awaited their turn to squeeze through a narrow strip of space around the dyads. On occasions where passers-by are forced to walk through conversational groups, they express their discomfort using verbal or non-verbal reactions. Passers-by tend to avoid eye contact (Cheyne and Efran, 1972), duck their head, and utter phrases like "excuse me!" and "sorry!" (Goffman, 1963; Knowles, 1973).

Felipe and Sommer (1966) remarks that avoidance reactions, such as the ones described here, are deployed based on many factors, e.g. the relationship between oneself and an intruder, territorial constraints applicable to oneself, victim's attribution of motives to the intrusion, and the situation in which intrusions occur. In normal circumstances, intruders' close approaches are confronted with the sort of *barrier behaviour* or *avoidance barrier* described here, or a stare, a still closer approach, and an aggressive posture (Patterson and Edinger, 1987). At the same time, violations of personal spaces may be inevitable in crowded situations, and therefore, violators are not held responsible for intrusions as intensively as in less crowded situations (Patterson and Edinger, 1987). People still tend to look away, or orient their body parts away from one another, or stiffen their bodies (Hall, 1966). People also fix their eyes on infinity except for exchanging passing glances (Hall, 1964). In general, eye-contact decreases with increased spatial proximity, whereas the length of mutual glances increases with longer interpersonal distances (Argyle and Dean, 1965).

2.9 Co-located Conversational Groups in Multi-focused Gatherings

Apart from face-to-face encounters, Goffman (1963) identified two other social units: the social occasion and the social gathering. The *social occasion* is the social-psychological unit that provides the frame of reference for face-to-face encounters. The *social gathering* is the overarching social situation that binds the behaviour and actions of a set of persons, who remain accessible to one another during any one continuous period of time. Goffman (1963) also suggested that a set of rules called *situational properties* apply to social gatherings. These rules determine how individuals relate their behaviour with respect to one another, taking into consideration their co-presence in a social situation.

A party is an example of a social occasion, and a birthday party group, or a drinks reception party group, are examples of social gatherings. Social occasions such as parties are often a platform for multiple small focused gatherings (Goffman, 1983). In other words, a party is like a big interactional territory that includes several knots of conversational groups, also called F-formation systems (Lyman and Scott, 1967; McDermott and Roth, 1978). Therefore, parties and party-like social situations are also called multi-focused gatherings (Kendon, 1988).

An appropriate trigger can restructure the situation in a multi-focused gathering, and cause several independent knots to come together to form one big group (Goffman, 1963), e.g. a singer or a musician at a party could draw the attention of guests. There might even be a point where the party becomes a performance and guests play audience to the performer.

On the other hand, when a group has too many participants, then side involvements of smaller groups can start to occur. Nonetheless, because of the overarching focus of attention within the larger group, the nature and character of collaborations within smaller groups are normally underplayed. Goffman (1961, 1963) uses the terms *subordinated groups* and *byplays* to denote the side involvements or subgroups in a large group. He also makes the distinction that, in a big group, members may or may not commit to maintaining a joint focus of attention, whereas in a focused gathering, participants commit and contribute to sustaining a common focus of attention (Goffman, 1961).

2.10 Techniques for Studying Spatial Behaviour

In this section, we briefly review the different techniques that have been used to research social-spatial behaviour in humans, both in natural as well as laboratory settings. The review highlights some of the advantages and disadvantages of these methods, which sets the scene for the computational models of social spatial behaviour that started emerging in the late nineties. The studies reviewed in this chapter used the following methods: (1) unobtrusive observation, e.g. the study of F-formation systems based on the analysis of video recordings in Deutsch (1977, 1979) and Kendon (1990). (2) chair placement or selection, e.g. the study of seating arrangements preferred by visitors in Sommer (1962, 1965) and Becker (1973). (3) felt board or stimulus object tests, e.g. the study using human head-sized objects in Argyle and Dean (1965) and the study of permeability of the group boundary in Knowles (1973). (4) staged interactions, e.g. the study of passer-by behaviour around conversational groups in Cheyne and Efran (1972), Efran and Cheyne (1973) and the study of personal space violations in McDowell (1972). (5) paper and pencil tests, e.g. Hall et al. (1968) asked participants to arrange coins and pencils in configurations that reflected being too close, far apart, side by-side, or next to each other. When using the paper and pencil technique participants are expected to complete a study by drawing on their previous experience of the phenomenon being studied. Hayduk (1978) considers this requirement to be cognitively demanding for participants.

Gathering participant feedback on stimulus photographs and role-playing techniques have also been used to study spatial behaviour (Mehrabian, 1968; Becker, 1973; Patterson and Edinger, 1987). Mehrabian (1968) notes that the use of photographic stimuli can yield information about postures and arrangements, but not about the dynamics of spatial behaviour. Hayduk (1978) proposed the stop distance technique, which involves participants stopping at desired distances when approaching experimenters or confederates, or vice-versa. A downside of this approach is noted as being the considerable practice required for experimenters and confederates to ensure they do not introduce variations in walking speed or other stopping cues.

Other data collection methods such as questionnaires and survey studies are less preferred for capturing the dynamics of spatial behaviour. Hall (1963) and Hall et al. (1968) explain the unproductivity of such techniques. Hall notes that spatial behaviour in humans is akin to the tone of voice or stress and pitch exercised when speaking a language – subtle and intuitive – hence making it difficult for participants to consciously recall and describe the nuances of their spatial behaviour. In research on social spatial behaviour in humans, unobtrusive observation has been the preferred approach.

2.11 Summary

In this chapter, we identified the systematic means by which interactants organise their position and orientation with respect to one another, during face-to-face encounters. We also reviewed how the spatial orientational arrangements of interactants' bodies serves to delineate the spatial arrangements of conversational groups. We reviewed the different types of dyadic arrangements, such as the H, V, C, L, N and I-shaped arrangements. Although not investigated in as much detail as the dyadic patterns, we found that circular, triangular, square, parallelogram, oval and semi-circular arrangements are suggested for multi-party conversational groups. There is no evidence, however, regarding the actual strategy used by interactants to arrive at a particular spatial arrangement on a particular occasion.

Understandably, the most basic requirement is for interactants to be able to see and hear one another clearly. Assuming it is true that an overlap of transactional segments inherently makes it possible for interactants to perceive one another clearly, we do not know yet what causes interactants to assume specific spatial arrangements. This is especially an issue in multi-party interactions where it may not be possible that the overlap of transactional segment achieved is uniform for all interactants. For example, in a group of seven or eight people, unless interactants stand in a perfectly circular arrangement, it is not possible to achieve the same level of transactional segment overlap for all participants. But for this to happen, the space where conversational groups occur should be big enough, and free of furniture and other objects. As noted in section 2.7, Scheflen (1964) also raises this point. In reality, conversational groups function everywhere – on the road, in the party hall, at the beach etc., and not all places will have the space required to host perfectly circular arrangements.

Even if it were possible to host reasonably big circular arrangements, how do participants arrive at such an arrangement? The notable contributions reviewed in this chapter – proxemics by Hall (1966), F-formation systems by Kendon (1990), the structure of transactional segments by Deutsch (1977) – were all based on observations of dyadic or triadic conversations. What happens in groups of more than three participants? Do participants arrive at a spatial arrangement by adjusting their position and orientation to achieve desirable interpersonal distances and overlap of transactional segments? Or do they readjust their position and orientation in view of sustaining specific spatial arrangements?

Computational models afford the possibility to answer these questions. It is possible to translate theoretical and empirical constructs into computational models and run them as simulations to gain clarity about processes where information is incomplete, inadequate or prone to ambiguity (Gilbert and Troitzsch, 2005; Gilbert, 2008). Computer simulations also provide an experimental test bed for investigating phenomena that may otherwise be difficult, challenging or hard to test in real life. For instance, Scheflen (1964) suggests that the function of a particular behaviour can be determined by comparing what happens when the behaviour is present with what happens when it is absent in the same situation. Evidently, it is challenging to turn on and turn off spatial behaviour in naturally occurring conversational groups. However, it is possible to run computer simulations, with precise control over the spatial behavioural characteristics of agents, to synthetically generate conversational groups, and analyse the outcomes. In this thesis, we conceptualise and implement alternative computational models to control the movement, positioning and orientation behaviour of individual agents in a simulated environment, to understand the emergence and persistence of the spatial patterns of conversational groups. But before that, in the next chapter, we will review existing computational models that aim to simulate the spatial patterns of conversational groups. We will focus on highlighting the relationship between the rules used to control the spatial behaviour of individual agents and the resulting arrangements of conversational groups.

Chapter 3

Computational Models of Conversational Groups

In this chapter we review existing computational models that enable agents, i.e. virtual representations of humans, to form conversational groups. The models draw insights from theories of human interaction, such as Hall's (1966) proxemics and Kendon's (1990) F-formation systems, to simulate the spatial behaviour of agents. We review algorithms and the rules used in models to control the movement, positioning, and orientation behaviour of agents engaged in face-to-face encounters. We also discuss the similarities and differences between models.

We begin our review by clarifying the definition of different types of agents, i.e. actors, virtual humans, avatars and non-player characters, and the need for generating believable social spatial behaviour for agents. Just as research on human territorial behaviour derived its inspiration from research on animal territorial behaviour (cf. section 2.1), many of the existing models of conversational groups are adaptions of models of coordinated group movement in animals. In section 3.3, we review Reynolds's (1987, 1999) boid model, which simulates the coordinated group movement of flocks of birds, and is one of the most commonly adapted models for simulating the steering behaviour of agents. In section 3.4, we review the social force model (Helbing and Molnar, 1995; Helbing et al., 1994) used to simulate pedestrian movement. In section 3.5, we review existing approaches to simulate the spatial patterns of conversational groups, followed by a review of the evaluations of simulated conversational groups in section 3.6. We then review models of social spatial behaviour for robots, and models for automatic detection of conversational groups in sections 3.7 and 3.8, respectively. We conclude the chapter by presenting a discussion of the limitations of existing models to simulate the spatial patterns of conversational groups.

3.1 Different Types of Agents

Gilbert (2008) defines agents as computer programs that interact with one another within a computational environment. Reynolds (1987) defines an actor as a computer code that is capable of interacting with other actors within a virtual environment (also computer code), based on rules of interaction (again, computer code). Macy and Willer (2002) consider actors as autonomous entities within agent-based models. Therefore, in terms of implementation, there are not many differences between agents and actors, and the terms have been used interchangeably to represent various entities, such as humans, birds, and animals.

Virtual humans are agents that only represent humans. They are intended to replicate the appearance, behaviour, and/or cognitive characteristics of real humans. Swartout et al. (2006) note that virtual humans look like, act like and interact with humans. Avatars are virtual characters controlled by real humans. Avatars can also be part-autonomous and part-controlled by users. Reynolds (1999) and Gillies et al. (2008) use the term *non-player characters* to denote avatars that are autonomous and not controlled by humans. In the rest of this chapter, we use the term *agents* to denote all types of computational characters.

3.2 Why Simulate Spatial Behaviour for Agents?

The need to generate believable movement, positioning, and orientation behaviour for agents engaging in face-to-face encounters has been discussed extensively in the literature. Cassell and Vilhjálmsson (1999) proposed *BodyChat*, a system in which embodied conversational avatars could be controlled to navigate in a virtual environment, and engage in face-to-face conversations. The system required users to manually control the movement of their avatars. Once users moved their avatars within the field of view of other avatars, BodyChat automated the facial expressions and head movements of avatars, but users continued being responsible for navigation and speech content. Evaluation of the BodyChat system showed that users felt that if the spatial behaviour of avatars was automated, they would engage in longer conversations. Similarly, Jan and Traum (2005) proposed a model where characters interacted with one another using speech and non-verbal behaviour, such as nodding, gestures, posture shifts, and gaze. The model, however, did not simulate the movement, positioning, and orientation of characters. Evaluation of the model showed that it is essential to automate the movement and spatial positioning of characters for realistic and believable simulations of face-to-face conversations.

Pedica et al. (2010) also emphasized that simulating territorial behaviour for virtual characters increases the believability of characters. In Pedica et al.'s (2010) study, participants found that videos where characters demonstrated territorial behaviour were more believable than the ones where no such behaviour was demonstrated. Cafaro et al. (2012) found that it was a burden for users when they had to manually control the spatial behaviour of their avatars in online platforms, such as Social Life and World of Warcraft. Over the years, several other studies such as Jan and Traum (2007), Vilhjalmsson et al. (2007), Pedica and Vilhjálmsson (2008, 2010, 2012), Ennis and O'Sullivan (2012) etc., have all emphasised the need to automate spacing and orientation behaviour for agents, as well as robots (Yamaoka et al., 2010; Koay et al., 2014).

3.3 The Boid Model

One of the early approaches to simulate the navigation of agents was to manually script the path of each agent, from a source point to a destination point, within the virtual world (Reynolds, 1988). Manual scripting works when there are only few agents to animate, but when there are a large numbers (e.g. in the order of thousands), it may be arduous, or even impossible, to manually script the navigation of each character.

The boid model proposed by Reynolds (1987) was one of the early approaches to automate the navigation of agents. The boid model does not specify the path taken by agents to move from one point to another, instead, agents navigate based on rules controlling their velocity and acceleration. Reynolds (1988) refers to this as *self-directed action*.

Particle systems is a technique used to model the movement of a large number of entities (*aka* particles), to simulate different kinds of fuzzy phenomena, such as fire, smoke, cloud, etc. (Reeves, 1983). The boid model, which is an advanced form of a simple particle systems model, simulates flocking behaviour in hundreds of birds using a distributed behaviour modelling strategy. The strategy involves specifying simple behavioural rules for each individual agent in the system, and then the target phenomenon, such as bird flocking, emerges naturally from the local interactions between the agents and their environment. Reynolds (1987) refers to the emergent pattern of agent motion in the boid model as *impromptu flocking*. This is because flocking behaviour is not hard-coded in the model, but it emerges naturally from the behavioural rules specified for each agent in the model.

The agents in Reynolds's (1987) boid model are called boids to represent bird-like objects[1]. The navigation of boids in Reynolds's (1987) flocking model is a consequence of combining the outcomes of three heuristic spatial behavioural rules: (1) separation – avoid collision with nearby boids; (2) cohesion – stay close to nearby boids; and (3) alignment – move ahead in the same direction as that of nearby boids. The outcome of each of these rules is a force component that is capable of steering the motion of a particular boid at a particular speed in a particular direction. To ensure that outcomes do not cancel each other out, a net steering force is calculated, based on a prioritised and weighted combination of the three force components. The net steering force calculated for each boid, at every time step, drives the movement

[1] Although Reynolds (1987) presents his boid model in the context of bird flocks, he notes that the model outcome can be considered to represent other collective behaviour, such as schools of fish and herds of animals.

of the boid in the model world. More precisely, the net steering force translates into actions, that control and reassess the position, orientation, and the velocity of boids (Reynolds, 2000).

At the system level, the collective movement of boids in Reynolds's (1987) boid model resembles bird flocking. In other words, controlling the navigation of each individual boid in the model using the separation, cohesion, and alignment rules, leads to the emergence of a coordinated motion pattern that mimics the flocking of real birds. Reynolds's (1987) boid model has only been evaluated using face validity, i.e. outcomes of the simulation were found to be visually convincing as bird flocking behaviour. Nonetheless, it is one of the most popular models[2], for automating the steering behaviour of agents.

Reynolds (1987) argued that assuming global knowledge of the environment is both unrealistic (in comparison to real bird behaviour) and computationally expensive. Consequently, boids were programmed with a local spatial awareness, which only allows them to perceive and respond to two or three nearest neighbour boids. Each boid in the model has access to a global database that stores the position, orientation, and velocity of other boids (and other objects, if any) in the environment. Based on this information, boids work out their nearest neighbours using the distance formula. Then, at every time step, the net steering force required to drive the motion of a boid is calculated based on the relative position and orientation of utmost two or three other neighbouring boids. The ability to work based on such a local perception, and yet yield the most realistic flocking behaviour, is the highlight of the boid model.

Seek, flee, arrival, pursuit, offset pursuit, evasion, wander, path following, wall following, flow field following, containment, leader following, and unaligned collision, are additional steering behaviours that Reynolds (1999) added to his original boid model. Several models that we review in this chapter, and our Party World models described in the next chapter, have adapted various aspects of the boid model to automate the spatial behaviour of agents.

[2] The model proposed in Reynolds (1987) has been cited over 7400 times.

3.4 The Social Force Model

Helbing and Molnar (1995) proposed the *social force model* to simulate the movement of pedestrian agents within a virtual world as if they are subject to *social forces*. The forces do not directly act on the agents, but they constitute a net motivation, which causes an agent to move in certain ways. Helbing et al. (1994) describes this as a *motivation to act*. The net motivation is calculated from an equation containing three terms. The first term describes an agent's acceleration to a desired velocity. The second term denotes the distance an agent needs to maintain from other pedestrian agents and obstacles, such as walls and borders. The last term models the attractive influences that cause an agent to move towards a particular destination within the simulated environment. The three terms come together to create a net force that causes an agent to move in a particular direction within the virtual world. Running the model with multiple agents, each driven independently by the social force model, results in a simulation of pedestrian traffic. Again, as is the case with Reynolds's (1987) boid model, the social force model also has only three simple rules governing the motion of each individual agent in the system, and yet, at the system level, the collective motion of agents mimics pedestrian traffic.

The pedestrian motion resulting from the social force model has two distinct characteristics. Firstly, there is a possibility of multiple lanes resulting from the model simulation, however, within each lane, pedestrians either move from left to right or from right to left. Secondly, at narrow exits, the model causes an oscillatory flow in pedestrian motion. Imagine a one dimensional wall, with a very narrow exit right at the middle, through which agents try to pass through from either direction. For some time, agents moving in one direction would continue to pass through the exit, e.g. pedestrians moving from the left to the right. During this time, the social force keeps accumulating for pedestrians waiting to pass through from the other direction, i.e. from the right to the left. Eventually, when the accumulated social force exceeds a particular threshold, one of the agents from the other direction will manage to squeeze through the exit, followed by others waiting to move in the same direction. However, the flow could get reversed any time, when the accumulated social force exceeds the threshold for pedestrians waiting to move in the other direction. In this manner, the oscillatory flow in pedestrian motion continues for the entire duration of the simulation.

Helbing and Molnar (1995) suggested the possibility of adapting their social force model to simulate other collective phenomena including group dynamics. However, Pedica and Vilhjálmsson (2010) argue that simulating the spatial dynamics of large groups, such as crowds of people, flocks of birds, and schools of fish, is essentially different from simulating the spatial dynamics of smaller units such as conversational groups. Forming conversational groups is believed to be a more refined behaviour than coordinated crowd motion (Vilhjalmsson et al., 2007). While flocking and pedestrian simulations entail continuous movement in space, one of the most important requirements for mimicking the process of forming conversational groups is that agents should stop moving and engage in nuanced positional-orientational readjustments. Consequently, for models simulating conversational groups, the main idea adapted from the boid model and the social force model is to use a combination of force components, to control the movement and the coordinated positional-orientational readjustments of agents. In the following section, we review and deconstruct the rules underlying models, which aim to simulate the spatial patterns of conversational groups.

3.5 Models to Simulate the Spatial Patterns of Conversational Groups

Jan and Traum (2007) proposed a model to deal with the dynamic movement and positioning of characters engaged in face-to-face encounters. The model was intended to support multi-party interactions. The model calculates four force components based on which movement and positional readjustments are initiated for agents. The first component $\overrightarrow{F}_{speaker}$ is an attractive force that causes an agent to move towards the position of the current speaker agent. Next, $\overrightarrow{F}_{noise}$ is a repulsive force that causes characters to move away from sources of disturbances (e.g. noise). The third component $\overrightarrow{F}_{proximity}$ is a repulsive force that causes characters to move away from standing too close to one another. Lastly, $\overrightarrow{F}_{circle}$ causes characters to organise themselves in circular arrangements.

At each time step, a net motivation force, calculated as the weighted sum of the $\overrightarrow{F}_{speaker}$, $\overrightarrow{F}_{noise}$, $\overrightarrow{F}_{proximity}$ and $\overrightarrow{F}_{circle}$ forces, acts upon a character causing it to move or readjust its position. Jan and Traum's (2007) model did not explicitly deal with the orientation of characters. This could be the reason why characters don't appear to face one another, even when they are arranged in a circular spatial configuration – see the spatial-orientational arrangement of agents in figure 3.1. Jan et al. (2007) extended their model, to incorporate culture-specific interpersonal distance behaviour for characters, based on the metrics proposed by Hall (1966). Measures of the intimate, personal, and social distances attributed to different cultures were given as input to the model. In the resulting simulations, characters were found to adjust their interpersonal distances based on pre-defined relationships, and cultural affiliations.

Figure 3.1: Spatial arrangement of characters in Jan and Traum's (2007) model

The spatial pattern of the cluster in figure 3.1 draws our attention to another important point. The characters are arranged in a circular pattern, even though $\overrightarrow{F}_{circle}$ was only one of the forces used in the model, in fact the least weighted one. This could imply two things. Either the weighted addition of the attractive and repulsive forces inevitably, i.e. with or without the $\overrightarrow{F}_{circle}$, causes agents to arrange themselves in a circular pattern. Or, even if weighted the least, $\overrightarrow{F}_{circle}$ has a dominant effect in causing agents to arrange themselves in a circular arrangement. It is not clear which of these reasons leads to the emergence of the circular patterns, because the unique role of each individual force component used in Jan and Traum's (2007) model has not been systematically assessed. There is a lack of reasoning regarding what each rule does to influence the overall spatial arrangement of agent clusters.

The model proposed in Jan and Traum (2007) inspired several later models to simulate the spatial patterns of conversational groups. For instance, Talbot and Youngblood (2013a,b) proposed a force directed graph model to arrange virtual actors in a play on the stage. The model uses a rule called as *centering and encircling groups*, to create a dummy vertex that is connected to every other actor node, and is used as a pulling force to center characters around this point. It is suggested that the dummy node causes characters to arrange themselves in circular spatial configurations.

Pedica and Vilhjálmsson (2008) also proposed a model, inspired by Reynolds's (1999) steering model and Jan and Traum's (2007) dynamic movement and positioning model, to enable readjustments to the position and orientation of avatars engaged in face-to-face encounters. The model uses a weighted and prioritised combination of force vectors to drive the steering behaviour of avatars. The *keep personal distance* force enables avatars to avoid stepping into one another's virtual equivalent of personal distance zones. The *keep conversation equality* force motivates avatars to maintain a shared group space of constant size. It seems that the *keep conversation equality* rule causes each and every avatar in a group to stand at an equal distance from, and facing towards, the centroid point of the group. Here, centroid is calculated as the average of the (x,y) values of the positions of each individual avatar within the group. The model also has a *keep conversation cohesion* force, which prevents avatars being isolated, by motivating them to join ongoing groups.

The three forces used in Pedica and Vilhjálmsson's (2008) model are combined using different weights and priorities to constitute a net steering force that causes an avatar to readjust its position and orientation. Readjustments occur whenever

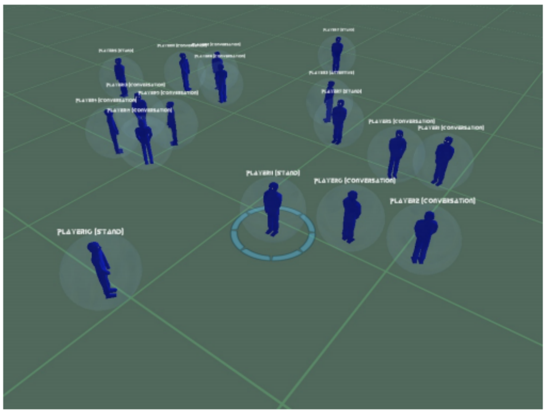


Image reproduced with kind permission of the publishers.
Figure 3.2: Spatial arrangements of characters in Pedica and Vilhjálmsson's (2008) model

conversational groups are disturbed, e.g. when an existing member leaves a group, or a new member joins a group. They describe avatars wandering around coming to a stop just outside a conversational group they wish to join. The newcomer is accepted into the group, if the avatar continues to stand facing towards the centre of the conversational group, for longer than a predetermined duration. In order to accommodate the newcomer, the other avatars in the group readjust their position and orientation based on the net motivation force, calculated as a weighted sum of the *keep personal distance*, *keep conversation equality* and *keep conversation cohesion* forces. Figure 3.2 is an outcome from Pedica and Vilhjálmsson's (2008) model, which show the centroid-driven, equally spaced arrangements of avatar clusters.

Pedica and Vilhjálmsson (2010) added two further components to their steering algorithm, the *keep conversational common attention* and the *keep conversational domain awareness* forces. Pedica and Vilhjálmsson (2008, 2010) emphasise that the steering algorithm entails only reactive spatial behaviours and no complex cognitive processing. The reactive spatial behaviours enable agents to perceive their environment (which has other agents in it), choose a reaction based on the state of the environment, and then carry out a physical motion (e.g. changing position and/or orientation) corresponding to the chosen reaction. The outcomes of these models were not formally evaluated, but it has been suggested that visual confirmations show that the spatial configurations of avatars look realistic.

Pedica and Vilhjálmsson (2012) proposed a behaviour tree approach, for choosing spatial behavioural reactions based on weighted priorities assigned to different territorial behaviour alternatives. The tree structure meant that higher priority behavioural nodes might be implemented at the cost of lower priority behavioural nodes. Say there are two conflicting priorities, one emphasising that a character should position and orient itself in such a way as to demonstrate continued participation in a conversational group, while another expects the agent to pay attention to an external event. In this case, the alternative with a higher priority will be chosen over the one with a lower priority. It has been asserted that such a priority based behaviour tree implementation enhances the earlier version of the steering model in Pedica and Vilhjálmsson (2010). Figure 3.3 is an image from a video accompanying Pedica and Vilhjálmsson (2012). The image shows the circular arrangements of agents. More recently, Pedica et al. (2015) note that they have extended their territorial behavioural model, to allow interactant agents to have equal access to the o-space. The model is also intended to get passer-by agents to avoid walking near simulated conversational groups, by modelling a repulsive force, which causes passers-by to move in a direction away from the group.

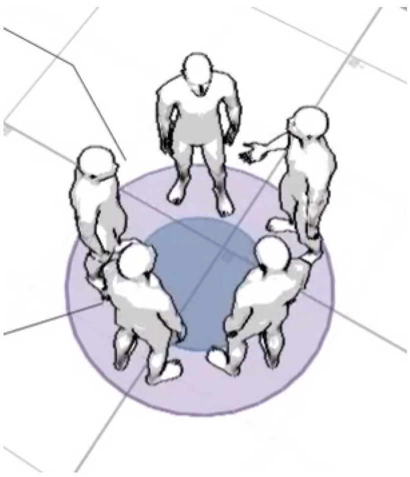


Image reproduced with kind permission of the authors.
Figure 3.3: Spatial arrangement of characters in Pedica and Vilhjálmsson's (2012) behaviour tree model

Lakshika et al. (2012) proposed a model to simulate the spatio-temporal dynamics of conversational groups, based on four simple rules similar to Pedica and

Vilhjálmsson (2008, 2010): "keep conversational distance", "keep center of the conversation", "keep visibility", and "keep distance to nearest neighbours". The difference, however, are the evolutionary aspects used in the model, whereby the four simple rules, and the parameters involved, evolve during the course of the simulation using a genetic algorithm approach. The goal was to derive the optimal rules and parameter conditions for simulating the most aesthetic visual simulations of the spatial-temporal dynamics of conversational groups.

Karimaghalou et al. (2014) also adapted Pedica and Vilhjálmsson's (2008) model, to control the movement of characters in to and out of conversational groups, based on a measure of interestingness. The model had one subject character participate in and leave conversational groups involving other characters. Each character was assigned an interestingness value at the start of simulations. The subject character moves towards the group that has the highest interestingness value. Once the subject character joins a group, the interestingness score of that group starts to vary as a function of time. The continued reassessment of the interestingness score causes the subject character to reevaluate its motivations to stay in or leave an existing conversational group. Despite the proposed enhancement, the model handles the spatial arrangement of characters in a similar fashion to the approach used in Jan and Traum (2007) and Pedica and Vilhjálmsson (2008, 2010, 2012) – agents in a group are instructed to stand at an equal distance from the centroid point and facing towards it. Figure 3.4 shows the spatial arrangements of groups resulting from Karimaghalou et al.'s (2014) model. Notice the arrows around characters pointing towards the centroid position of the group.

Karimaghalou et al. (2014) note that the duration a subject character spends within one group, and its transition between different groups, can be controlled by adjusting the thresholds set for interestingness in their model. However, Karimaghalou et al. (2014) did not implement a baseline scenario to measure the duration the subject spends within a group with just the centroid-based implementation, i.e. without the added motivation of interestingness scores. Without such a baseline, it is difficult to understand how the interestingness score influences the duration a character

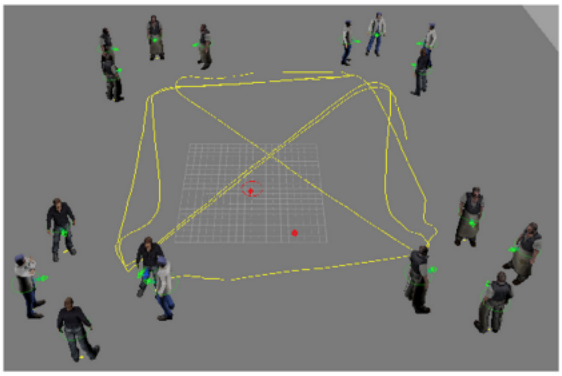


Image reproduced with kind permission of the publishers.

Figure 3.4: Spatial arrangements of characters in Karimaghalou et al.'s (2014) model.

spends within a group, and its transition between groups.

Salem and Earle (2000) proposed a model in which avatars are explicitly organised around the perimeter of a ring. Avatars are also evenly distributed along the circumference of a circle for the sake of affording equal spatial importance to all characters. The resulting spatial configurations of agent clusters are shown in figure 3.5. The explicit instructions in Salem and Earle's (2000) model, asking agents to form circular spatial arrangements, defeats the purpose of autonomy and emergence. Furthermore, as seen in figure 3.5, the resulting configurations appear too perfect to be real. Lastly, Riesman et al. (1960) suggested that six is a reasonable average for the size of conversational groups, but Salem and Earle (2000) forcefully restrict the size of conversational groups in their model to a maximum of six avatars.

Figure 3.5: Spatial arrangement of characters in Salem and Earle's (2000) model

Rojas et al. (2014, 2016) proposed a model of pedestrian behaviour, where agents intermittently form conversational groups, e.g. at side-walks of roads. They used the centroid-based approach to simulate F-formations. The model also triggers scripted gestural exchanges between agents participating in the F-formations. One of the agents begins to walk away from an F-formation after a prefixed time, causing other agents to follow, thereby ending the F-formation. Rojas et al. (2014) conducted a study in which participants were questioned about their perceptions of F-formations in the pedestrian simulation. The simulation was projected in a head-mounted display. Participants are said to have responded more favourably towards pedestrian simulations that generated F-formations, than the ones that simulated groups waiting at a traffic light.

There are two issues with Rojas et al.'s (2014) evaluation study. Firstly, from the description of the study, it is not clear what sort of group arrangements, i.e. other than F-formations, did agents maintain at the traffic lights, i.e. were agents not engaged in a conversation at all, or did they maintain non-centroid facing arrangements. Even if this information were provided, comparisons can only be made between simulations that used the centroid-based approach and the ones that did not use this approach, to simulate agent clusters. Otherwise, it is as if the centroid-based approach is assumed to be the standard way of generating F-formations.

The model proposed in Brom et al. (2012) is another example that uses the centroid-based approach. In this model, three characters (a boyfriend who goes out with two girlfriends at the same time) try to engage in a face-to-face quarrel. The model extends Reynolds's (1999) steering behaviours, and adopts the centroid-based approach using a finite state automata, to govern the dynamic positioning and orientation of characters trying to engage in a face-to-face quarrel.

Thus, as seen so far, there are several models which have used the same strategy to simulate the spatial patterns of conversational groups – instruct a group of agents to stand at an equal distance from, and facing towards, the centroid point of the group. The configurations of agent clusters resulting from this approach appear to be predominantly circular, and this seems to satisfy the assumption of models that multi-party conversational groups are indeed circular in shape. Furthermore, although Hall's (1966) proxemics and Kendon's (1990) F-formation systems are cited as the motivation for these models, it is evident that the centroid-based approach does not deal with conceptualisations of interpersonal distance zones or transactional segments. And yet, circular arrangements are identified as F-formations, and claimed to be realistic. There is also no evidence about the spatial patterns of dyadic agent clusters resulting from the centroid-based approach, i.e. we do not know what sort of dyadic arrangements (H, V, I, L, N, C-arrangements) emerge from it, as existing models have not focused on this aspect.

Very few models use a non-centroid approach to control the positioning and orientation of agents engaged in face-to-face encounters. Vilhjalmsson et al. (2007) proposed *Social Puppets*, a software engine to generate socially and culturally (Arabic and Pashto) appropriate non-verbal behaviour. The engine was intended to serve as an interface at the intermediate level, between a higher level cognitive model, and a lower-level physics engine which generates animations corresponding to cognitive states. The engine uses a manager that assigns one social puppet to each agent in the model. The manager then controls the puppets, at a global level, to coordinate the spatial behaviour of agents. Despite the realism that Social Puppets seeks to achieve, it defies the purpose of autonomy, because the manager explicitly controls the coordination of the spatial behaviour of individual puppets.

Laga and Amaoka (2009) proposed a model that allows an agent to control the extent of its personal space based on the agent's relationship with other agents in the model, and as a consequence of the evolution of relationship between agents. The model uses only two parameters, an agent's position and orientation, to compute the spatial extent of the agent's personal space. It is assumed that the personal space region is bigger at the front of an agent than at the sides and back. An agent uses the extent of its personal space as a means for calculating the most comfortable distance to maintain from other agents in the environment.

Ricks and Egbert (2012) proposed a method for pedestrian agents, which bump into one another, to engage in F-formation systems. Agents that are close enough, approach one other, and start interacting. The duration for which characters interact is determined as a function of the relationship between agents and their previous history of conversations. After this time, agents disengage from their conversation, and proceed towards their respective destinations. The model is claimed to be capable of simulating conversational groups in different pedestrian scenarios, such as the market place, party hall, or school campus, just by changing the kind of relationships defined between agents. Although implemented on a large-scale, i.e. in a crowd simulation of 3000 to 4000 agents, the model only implements dyadic F-formations.

Damian et al. (2011) proposed an Advanced Agent Animation (AAA) framework, which allows specifying parameters that will allow agents to get into specific configurations of F-formations, e.g. the C,V and L-shaped arrangements. In other words, the model does not specify individual behavioural rules that allow F-formations to emerge naturally, instead, it allows specifying parameters for each individual agent, such that the spatial patterns of F-formations are explicitly controlled.

Nguyen and Wachsmuth (2011, 2013) proposed a model of peripersonal space to help virtual characters reach out for objects in near proximity. The range of peripersonal space is intended to fall in between the intimate and personal distance zones proposed by Hall (1966). A model of interaction space was also proposed, to allow a couple of virtual characters to work collaboratively, while handling the objects in between them. The model implementation visualised a couple of agents working collaboratively with objects randomly placed on a two-dimensional table surface. Both peripersonal and interaction spaces were modelled as potential force fields – each distinct point in space is characterised by a force vector. The nature of collaboration between agents could be cooperative or competitive, depending upon the attractive or repulsive force fields used to model the peripersonal and interpersonal spaces. The model did not simulate o-spaces *per se*, rather it simulated a collaborative work space called the interaction space, and the resulting spatial configuration of agents was considered to be supplementary to F-formation systems (Nguyen and Wachsmuth, 2011, 2013).

3.6 Evaluation of Simulated Conversational Group Behaviour

While, on the one hand, we reviewed models used to generate the spatial patterns of conversational groups, on the other, it is important to understand how participants experience simulated conversational groups. It is important to know if participants react to simulated conversational groups in the same way as they would respond to real conversational groups. In this section, we review human-agent interaction experiments, avatar-mediated interaction experiments, and other evaluation methods, intended to analyse and gauge participants' reactions to simulated conversational groups.

Kistler et al. (2012) conducted an experiment where participants had to control their avatars to establish and participate in conversational groups involving other avatars. Participants also had to control the exit of their avatars from ongoing conversational groups. Kistler et al. (2012) had programmed the avatars in their experiment (except for the one controlled by the participant) with different cultural backgrounds, and consequently, with variations in proxemics and F-formation system behaviour. The spatial behaviour of avatars were not automated, but hand-coded for each agent in the system. The goal of Kistler et al.'s (2012) evaluation study was to assess if participants belonging to different cultures were able to recognise the culture-specific spatial behavioural traits of avatars. Results showed that participants considered the spatial behaviour of avatars belonging to their own culture as more appropriate than avatars belonging to other cultures.

Rehm et al. (2005) implemented a model of social group dynamics, which causes an agent to wander around until it detects other agents, and forms conversational groups. A de-centralised behaviour control mechanism manages the interpersonal distances and orientation of agents within groups. The relationships between agents are pre-defined, based on which they choose different spatial arrangements for the encounters. During an avatar-mediated interaction study, participants were asked to join one of the two groups of interacting dyads. Outcomes showed that participants preferred to join open arrangements (L, C and I-shaped arrangements) over closed arrangements (N, H and V-shaped arrangements) of dyads. It was also found that participants preferred to position themselves at the social distance zone.

Friedman et al. (2007) used customised software bots that wander around, engaging in simple interactions with the avatars of other (unknown) users, in the online virtual world Second Life. The bots collected data on the interpersonal distances maintained during encounters. The purpose was to assess if participants showed similar proxemic behaviour in Second Life as they would in real world conversational encounters. The criteria used for measuring interpersonal distances was that characters should be in the same area for at least a minute, and should not be out of orientation by more than 90°. Satisfying these conditions were considered as valid encounters, in which case, the interpersonal distances maintained between bots were measured. Like Rehm et al. (2005), the results of Friedman et al.'s (2007) study suggested that participants preferred the social distance for interactions. Results actually showed two peaks in the distribution of interpersonal distances among interacting bots. One was around the 1 - 1.5m mark and the other was around the 4.5m mark. These ranges fall under Hall's (1966) classification of personal and social distances respectively. Friedman et al. (2007) also found that a proxemic move from one interactant bot often sees a corresponding proxemic move from another interactant bot.

Assessments in Friedman et al.'s (2007) study did not include intimate encounters, which could be a reason why no intimate distances were recorded in their experiments. Jeffrey and Mark (1998) argued that avatars engaged in private conversations tend to stand closer to one another in face-to-face formations. They also found that violations of personal space resulted in compensatory moves, such as characters stepping back, or away from one another. It was also noted that the average distance between avatars increased as the number of avatars in the group increased.

The findings from Ennis and O'Sullivan's (2012) study slightly differ from the distance ranges suggested by Friedman et al. (2007). Ennis and O'Sullivan (2012) used motion capture technology to record the interpersonal distances and head orientations of real people in real world conversations. They then replayed these patterns on virtual characters to generate visual stimuli, under different distance ranges, and different orientation patterns. The distance patterns considered were close (0.4m - 0.5m), middle (1.65m - 2.1 m) and far (7m) ranges. The orientation conditions were: (1) two listeners facing one speaker in a triadic conversation, (2) one listener standing about 30° from the speaker, (3) one listener standing about 60° from the speaker, and lastly, (4) one listener standing about 60° from the speaker, but having its head oriented towards the speaker. Participants were played short conversation excerpts (7 seconds) between the virtual characters in each of these conditions, and then questioned about their perceptions of the spatial behaviour of agents. The findings from the study were that: (1) interpersonal distances between 0.4m - 2.1m were realistic, (2) participants were okay with a listener standing 30° apart from the speaker but not 60°, (3) when characters were out of sync by 60°, it helped improve the realism of interactions, at least if the listener had its head oriented towards the speaker.

Bailenson et al. (2003) report an experiment, where participants had to approach a virtual character, first from behind and then from the front. The study found that participants honoured the personal space bubble, both when at the back of the character, as well as when standing in front of it. Participants even seemed to honour the shape of the personal space boundaries. They stood close to the character when approaching from behind, but when approaching from the front they tended to stand at about 0.5m from the character.

Although Friedman et al. (2007) and Ennis and O'Sullivan (2012) report slightly different distance preferences, we must recognise that in the former study participants interacted with one another via their avatars, whereas in the latter, participants watched agents interact with one another. Hence, we appreciate the possibility that the mode of assessment could have influenced the differences in interpersonal distance preferences. Bailenson et al. (2003) and Ennis and O'Sullivan (2012) found that the type of agents used in a study also influences participants' evaluations of simulated social spatial behaviour, i.e. participants react differently towards avatars and autonomous agents. The difference is attributed to the possibility of participants being aware that, in case of avatars, it's a human controlling the characters, and therefore, they are more likely to relate to other avatars in similar ways as they would to other humans. On the other hand, when participants know that the agents are controlled by a computer, they may not endeavour to engage as much. Bailenson et al. (2003) also found that, when approached by characters at close interpersonal distances, participants avoided computer-controlled agents more than they avoided user-controlled avatars. The reason given was that participants considered it less likely that avatars would collide with or pass through them.

Llobera et al. (2010) conducted an experiment to assess how the electrodermal activity of participants was affected when approached by virtual characters at one of the four different distances: intimate, personal, social, and public. The virtual characters were either female or average human-sized cylinder shapes. The distance of approach, the number of characters that approached the participant, and the type of characters, were varied between trials. Results showed that participants displayed heightened physiological arousal when they were approached closely, but it did not seem to make a difference if the characters were female or cylinders. Participants showed more arousal when characters were approaching at closer distance ranges and in large numbers. The experiment also showed that arousal rates decreased with time due to adaptation influences, i.e. participants became familiar with the characters. In a long-term human-robot interaction study, where participants got to acclimatise with a robot over several weeks, Koay et al. (2014) also found that participants who were familiar with the robot preferred closer interpersonal distances, than those who had just encountered the robot.

Cafaro et al. (2012) conducted a study to assess the first impressions of participants when they interacted with a virtual human playing the role of a museum presenter. The virtual human tried to engage participants using one of the following strategies: proximity (approach or do not approach), smile, and gaze (for longer duration or for shorter duration). The study revealed that participants assessed the virtual human's approaching behaviour as a sign of extroversion. Another finding was that the virtual human's gaze and smile behaviour influenced perceptions regarding its friendliness, but proxemics had no influence on judgements concerning friendliness. In a recent study, Cafaro et al. (2016) showed that when virtual characters engaged in small-group conversations smile and orient their body towards an approaching user avatar, there is an increased sense of perceived social presence and believability.

Drawing together these findings, it is evident that people tend to react to simulated social spatial behaviour in similar ways as they would in real life. However, it has been shown that people perceive fully autonomous agents and user controlled avatars differently. Furthermore, the mode of interaction, i.e. direct interaction, mediated interaction, or third-person perspective, also influences participants' perceptions of and reaction to simulated spatial behaviour. Standing too close is considered to be a violation of personal space, just as in the case of real-world interactions. Apart from standing at the far phase of the intimate distance (Bailenson et al., 2003; Ennis and O'Sullivan, 2012), or the personal-social distance (Rehm et al., 2005; Friedman et al., 2007), facing one another within 30°, 60° (Ennis and O'Sullivan, 2012), or 90° (Friedman et al., 2007), are perceived as valid interactional encounters.

3.7 Models of Spatial Behaviour for Robots

Research in the field of robotics has also focused on generating social spatial behaviour. The main purpose of generating social spatial behaviour for robots is to enhance their day-to-day interactions with humans. Kuzuoka et al. (2010) proposed a model that allows a robot, acting based on a pre-defined script, to re-configure its position and orientation while presenting information about products to people. The model automates the positional-orientational readjustments of the robot only when it presents information to the users, but not when initiating interactions, because it is assumed that people will naturally establish an F-formation with the robot at the start of conversations. Contrary to this assumption, Shi et al. (2011) suggested two strategies for initiating conversations. One is the proactive strategy, which corresponds to the robot initiating a conversation, and the other is a reactive strategy, which corresponds to humans initiating the conversation and the robot responding to it.

In Shi et al.'s (2011) model, which of the two strategies is chosen for initiating interaction depends on whether the robot sees the user first (proactive), or if the user sees the robot first (reactive). In the reactive case, the robot was programmed to arrive at the sight zone of the user, i.e. a visual arc extending about 270° in front of the user, stand at about 1.5m from the user, and then to readjust the body to face the user in a vis-à-vis arrangement. The strategy used in the proactive case was to arrive at the focus zone of the user, i.e. a visual arc extending about 30° in front of the user, stand at about 1.1m - 1.5m from the user, at a position that affords visual access both to the user and to the object about which information is presented. Both strategies were found to perform better than a baseline scenario, which did not have a strategy for initiating conversations.

Yamaoka et al. (2010) proposed a model of proximity control for robots presenting information about objects at a museum. Rules were implemented in an incremental fashion to allow the robot to position itself at a suitable spot with respect to the visitor and the object. The first iteration had the robot position itself at an almost equal distance from both the visitor and the object. In the second iteration, the robot was positioned either close to the object, or close to the visitor. In the third iteration, both the robot and the visitor had unobstructed visual access to the object, because the positions of the robot and the object fell within a 90° field of view of the visitor, while the visitor and the object were within a 150° field of view of the robot. An evaluation of Yamaoka et al.'s (2010) model showed that participants preferred the arrangement in the third condition over the other two. However, in the first two conditions the robot was facing the visitor at all times, whereas in the third condition the robot was facing both the visitor and the object. Therefore, it may be possible that participants found conditions 1 and 2 less desirable than the third one, either because the robot was standing too close or too far from them, or because the robot never looked at the object when talking about it.

Mead and Matarić (2011) investigated an experimental setup involving two participants and a robot artefact. They instructed participants to interact among themselves and to include the robot in their interactions. A general trend found among the 11 pairs who participated in the study was that participants preferred vis-à-vis arrangements when their conversation did not involve the robot. The L-shaped arrangement was preferred when one designated participant was talking to the other about the robot. The side-by-side arrangement was preferred when both participants were talking about the robot as they continued looking at it. The spatial arrangements of participants were deduced qualitatively and not assessed quantitatively. Nevertheless, the intention was mainly to show that the spatial patterns of dyadic encounters could be influenced by a robot artefact.

Lindner and Eschenbach (2011) wanted a robot to have the ability to think and act when it came to space usage. They proposed a situation calculus approach that allows a robot to use high-level cognitive reasoning, to proactively assess the usage of five different zones of space around its body, under different social circumstances. The *personal space* is treated as the equivalent of Hall's (1966) definition of personal space. The *activity space* is considered to cover the o- and p-spaces of an F-formation system. The *affordance space* spans the area needed to perform activities within an environment. The *territory* signifies the space demarcated by markers. Lastly, the *penetrated space* indicates the extent of the entire spatial region where the robot performs its activities.

Mead and Mataric (2012) briefly describe a probabilistic framework for proxemic behaviour production in a human-robot interaction system. The framework is based on Hall's (1966) notion that different proxemic zones afford different sensory channels for communication. Consequently, Mead and Mataric's (2012) framework considers a set of desired sensory features that a robot wishes to experience during face-to-face encounters, which is then conditioned over spatial factors, such as the distance between a human and the robot, the angle from the robot to a person, and vice-versa.

Koay et al. (2014) conducted a human-robot interaction experiment, in which they assessed perceptions of practicality and comfort, when a service robot approached participants from different angles and distances. The robot approached a participant seated at a couch, either from a vis-à-vis angle, a 45° angle, or a 90° angle. The robot positioned itself either at 0.5m or at 1m from the participant. Handing a hat or serving drinks in a tray were the services offered by the robot. The experiment showed that, in terms of practicality, closer spatial approaches (at 0.5m) were preferred, whereas in terms of comfort, farther spatial approaches were preferred (at 1.0m). The side approaches of the robot were more preferred than the vis-à-vis approach in terms of both practicality and comfort. On the other hand, participants accustomed with the robot were found to be comfortable even with closer interpersonal distances and vis-à-vis approaches.

Programming the legged locomotion of robots in a physical world is more challenging than generating navigation behaviour for characters in a virtual world. Nonetheless, the aim of our review was to highlight the ways in which the field of robotics has dealt with modelling the socio-spatial behaviour of robots. Firstly, we found that positional-orientational readjustments during face-to-face encounters are regarded as an important aspect of human-robot interactions, too. Secondly, existing models have predominantly focused only on dyadic interactions between a robot and a user, or triadic interactions involving a robot, a user and an object. Vroon et al. (2015) acknowledge the difficulty in modelling human-robot interactions in a group scenario. They conducted an experiment involving groups of four participants solving a murder mystery in a face-to-face encounter. The conversational group involved three participants, and one robot, which was controlled by a fourth participant in a telepresence mode. Findings from the study suggest that participants preferred the robot standing at about 1.25m in front of them and facing the center of the group. But the study did not evaluate if the instructions to participants at the start of the experiment, i.e. to stand in circular or semi-circular arrangements, influenced their observations during the study.

3.8 Models for Automatic Detection of Conversational Groups

Research in the field of computer vision, particularly vision-based surveillance, has focused on using the concepts of proxemics and F-formation systems to automatically detect conversational groups in still images. Cristani et al. (2011b) proposed an approach referred to as the Hough Voting for F-formations (HVFF). The HVFF model takes as input the positions (with respect to the ground plane) and orientations of individuals in a scene, and votes for potential F-formation centres, a.k.a. o-space centres. One of the o-space centres is eventually chosen as a valid F-formation centre based on the following three step process. (1) Sampling: parameters of uncertainty are attached to the position and orientation estimates of people. (2) Voting: candidate o-space centres are weighted based on the uncertainties identified in the previous step. Voting also takes into account the 45° transactional segment ranges. (3) The o-space validation: verifying that the channel of space between any two o-space candidates are free of intrusions. The last condition aims to satisfy Kendon's (1990) suggestion that members of F-formations should have unobstructed access to the o-space. Based on the chosen F-formation centre, the interactants associated with it are retrieved from a list, which is recursively sorted through the three-step o-space estimation process.

Cristani et al. (2011b) suggest that the HVFF model performed well in detecting F-formations in both synthetic and real-world datasets involving conversational groups. Performance was measured by comparing the F-formations detected by the HVFF model with the F-formations identified by human annotators. A notable aspect about the HVFF model is that it does not assume that multi-party conversational groups have to be circular (Cristani et al., 2011a). Also, whereas most computational models have only aimed to simulate multi-party conversational groups of four to six agents, the HVFF model is capable of detecting F-formations of bigger groups. Vázquez et al. (2015) adapted Cristani et al.'s (2011a) Hough voting scheme to detect F-formations by tracking the lower body orientations of people, instead of head orientations. Vázquez et al.'s (2015) model is inspired by Kendon's (1990) suggestion that the spatial extent of individual transactional segments and o-spaces are determined by the orientation of the lower body and feet placement rather than head orientations.

Another model for automatic F-formation detection which uses a different approach, is called the Graph Cuts for F-formation (GCFF) discovery (Setti et al., 2014, 2015). Unlike HVFF, which requires both the position and orientation of individuals in a scene, the GCFF approach is capable of working only with position information. Because automatic head orientation estimation is a challenging problem, the ability to work with only position information is deemed to be an evident advantage of the GCFF approach. A comparison of the HVFF and GCFF models described in Setti et al. (2013) suggests it is best to use HVFF when clear, and incontrovertible, data regarding the position and orientation of participants is available. On the other hand, if only the position of individuals is available, because most modern trackers offer only that information, using the GCFF model is recommended. Setti et al. (2013) emphasise that both HVFF and GCFF models correctly identify F-formations, and at least two-thirds of its members, while also ensuring that less than one-third false judgements regarding membership are made.

Bazzani et al. (2013) proposed another social interaction discovery algorithm based on identifying the overlaps of visual focus of attention of people in a room. The visual focus of attention (VFOA) is considered as the visual field of a person in a three-dimensional representation of a scene. It represents where and what a person is looking at and is determined based on tracking the head position and orientation as well as the eye gaze dynamics. The relationship between people based on the overlap of VFOAs is deciphered based on computing an Inter-Relation Pattern Matrix (IRPM), which is a statistical technique to decipher a matrix summarising the social exchanges that may have occurred between participants. The IRPM works on the assumption that participants engaged in face-to-face conversations tend to stand at least within 2m of one another, their VFOAs overlap, and their head orientations fall within it. Outcomes of this model were found to work effectively, detecting conversational groups in still images from a coffee break dataset, and an airport lobby area dataset. Similar to the notion of VFOA, Zhang and Hung (2016) modelled people's frustum of attention to detect F-formations in an image. The notable aspect of Zhang and Hung's (2016) model is that, in addition to detecting members, the associates of an F-formation are also detected (cf. section 2.6.1.3 for definition of associates).

Gan et al. (2013) also proposed to automatically detect F-formations, and the participants in F-formations, by using a heat map approach to compute interaction spaces (akin to individual transactional segments) and global interaction spaces (akin to o-spaces) based on the positions and orientations of individuals in a scene. The approach differs from the HVFF and GCFF models by taking into account temporal information, i.e. assessing frames in a time-linear sequence to identify F-formations. It is noted that such a timeline approach tends not to assume that two people who are walking side-by-side (captured statically in a still frame) are in an F-formation. The model is intended to serve as an automatic photographing system to capture conversational groups in gatherings, thereby saving the cost and time for employing a photographer. Like the HVFF, GCFF and the IRPM models, Gan et al.'s (2013) model does not assume that multi-party F-formations are circular, or that its members are spaced out at equal distances.

3.9 Discussion and Summary

In this chapter we identified four distinct areas of research, and the work done in these areas, centred around the computational modelling of the spatial patterns of conversational groups. First, we reviewed computational models that aim to simulate the spatial patterns of conversational groups, by controlling the movement, position and orientation of individual agents. We found that most of the existing models have used the centroid-based approach to simulate the spatial patterns of conversational groups. We provided examples of the circular spatial arrangements of agent clusters resulting from these centroid-based models.

Secondly, we reviewed studies that focused on assessing participants' evaluations of simulated conversational groups. We found that evaluations of simulated spatial behaviour and conversational groups could vary based on the mode of assessment, i.e. depending on whether participants engage with the agents directly (via avatars or mediated interactions), or just witness agents interacting among themselves (third-person perspective). Thirdly, we reviewed models of social spatial behaviour in the context of human-robot interaction. We understand there is interest in modelling the social spatial behaviour of robots, to equip robots with the means to detect the social spatial behaviour of humans, and to detect ongoing F-formations. Lastly, we reviewed computational models that aim to detect F-formations in a still image, by taking as input the position and orientation (of the head or the lower body) of the people in the image.

We found that Hall's (1966) theory of interpersonal distances and Kendon's (1990) theory of F-formation systems are unifying themes across the four different areas of research. Since our research questions are framed around modelling the emergence of the spatial patterns of conversational groups (see chapter 1), we consider the models reviewed in section 3.5 as the motivation for our work. At the same time, since our research questions focus on the relationship between individual spatial behaviour and the emergent group patterns, we hope that the findings of our work will contribute to all four areas of research.

Our main concern about existing models is that most of them have used the centroid-based approach to simulate circular conversational groups. In the past, we have argued that the regularity of the spatial patterns of conversational groups resulting from the centroid-based model is questionable and does not correspond to reality (Narasimhan and White, 2014). However, even recently, Cafaro et al. (2016) have used the centroid-based approach by extending their earlier work in Pedica and Vilhjálmsson (2010, 2012). They argue that formations, which are irregular and yet compliant with Kendon's (1990) theory, can be generated by adapting the centroid-based model to vary interpersonal distances as a function of interpersonal attitudes. Nonetheless, the spatial arrangements of agent clusters shown in Cafaro et al. (2016) were still regularly shaped. This relates back to the question we raised concerning Jan and Traum's (2007) model in section 3.5 – $\overrightarrow{F}_{circle}$, which causes agents to organize themselves at an equal distance from the centroid point, was the least weighted force in Jan and Traum's (2007) model, and yet, the resulting spatial patterns of agent clusters were circular.

Therefore, our first question is why use the centroid-based approach to simulate the spatial patterns of conversational groups? There has been no empirical validation of the spatial patterns of agent clusters resulting from the centroid-based model, except for a few evaluations regarding the perceived believability of the arrangements. Hence, it remains unclear why the centroid-based approach is preferred for simulating conversational groups. One reason could be an assumption that multi-party F-formations are invariably circular in shape. However, although Kendon (1990) suggested that F-formations can be observed in the circles of free-standing conversational groups, he also emphasised that a circular pattern is not the only possible arrangement. In fact, neither Kendon's (1990) study, nor any other study, has estimated or measured the degree of roundness of F-formations. However, it appears that most models have over-generalised the notion of circular arrangements for multi-party conversational groups, and consequently used the centroid-based approach.

Our second question is regarding the applicability of the centroid-based approach for simulating dyadic arrangements. In chapter 2, we found that there are several types of dyadic arrangements of F-formations (see section 2.7). However, the influence of the centroid-based approach on dyadic arrangements has not been analysed. Our last question concerns the spatio-temporal dynamics of arrangements when using the centroid-based approach. In chapter 2, we found that F-formations are spatio-temporally regulated systems, i.e. they change over time as members join and exit the group (see section 2.6.1.4). There has been some focus on the rearrangement of agent clusters when using the centroid-based model. Jan and Traum (2007), Pedica and Vilhjálmsson (2012), Cafaro et al. (2016), etc., all provide examples of rearrangement of agent clusters. The examples focus on specific instances, i.e. when a new member joins the group or when an existing member leaves, but they do not explain the spatio-temporal dynamics of groups over longer durations.

In our work, we will take a step back to replicate the centroid-based model, to address the aforementioned concerns in a systematic way. Wilensky and Rand (2007) note that replication of agent-based models is a scientific process that involves implementing a model described and already implemented by another scientist. It has been recommended that the model replicated should differ from the original model in one or more of the following factors: time, authors, algorithm, toolkits, hardware and languages. Our replication of the centroid-based approach differs from the existing implementations in all six aspects. After replicating the model, we will analyse in a systematic way whether the spatial patterns of agent clusters resulting from a centroid-based approach correspond to reality, i.e. the spatial patterns of real-world conversational groups.

There is another important issue. In chapter 2, we noted that it might not be practical, or even possible, to form circular spatial arrangements all the time, due to constraints imposed by the physical environment where conversational groups occur (see section 2.7.1). Therefore, there is a need for an alternative to the centroid-based approach. In this thesis, we propose an alternative model, in which agents try to position themselves at reasonable interpersonal distances, while trying to maximise the number of agents covered within a 180° frontal view. As with the replication of the centroid-based model, we will systematically analyse the results of our alternative model, to verify and validate its performance.

Lastly, we believe that the spatial patterns of conversational groups will vary depending on whether agents readjust their position and orientation with respect to every other agent in their cluster, or only with respect to immediate neighbours. We refer to this concept as the neighbourhood effect. Kendon (1990) suggests that interactants in an F-formation tend to readjust their position and orientation with respect to one another. The examples of F-formation systems he provides involve only three to four participants, in which case it might be easy for each person to readjust their position and orientation with respect to others in the group. However, in bigger groups, it may not be practical, or even possible, for each individual to readjust their position and orientation based on every other member in the group. Consequently, we test the influence of neighbourhood effect on the spatial-orientational readjustments of agents, and consequently, the spatial patterns of agent clusters. In the following chapter we present the methodology used for implementing our models and describe each of the models in detail.

Part II

Method

Chapter 4

Party World: Our Agent-based Models and Simulations

> "Contrasting alternative theories, or hypotheses, is the foundation of what is simply called *scientific method*".
> — Steven F. Railsback and Volker Grimm, *Agent-based and Individual-based Modeling: A Practical Introduction*

The objective of our research is to explore and understand the role of interactants' spacing and orientation behaviour in the functioning of conversational groups. We are particularly interested to examine the effect of interactants' spatial-orientational relationship on the spatial form and social dynamics of conversational groups. To this end, we use an agent-based modelling approach to examine the emergence and maintenance of realistic patterns of conversational groups, from simple, individual-level rules of movement, positioning and orientation behaviour. In this chapter we describe the agent based models we have developed to simulate human conversational groups in a two-dimensional virtual world.

We propose three different agent-based models in the chapter. Each uses a different strategy to control the positioning and orientation behaviour of agents engaged in face-to-face encounters. The first is a simple model, where agents do not readjust their position or orientation during a face-to-face encounter. The second model allows agents to readjust their position and orientation relative to other agents within the group. The third model is our adaptation of the centroid-based approach, where agents readjust their position and orientation relative to the centroid point of the group. We describe each of these models and their implementations in this chapter. We analyse and compare the spatial form and social dynamics of agent clusters resulting from these models in subsequent chapters.

The rest of this chapter is organised as follows. We begin by defining the relevant agent-based modelling terminology in section 4.1. We then describe the conceptual basis of the three agent-based models in section 4.2, followed by a description of the implementation of models in section 4.3. We provide a description of the outputs obtained from the simulations in section 4.3.5 and a glimpse of the agent clusters resulting from the respective models in section 4.4. Lastly, we conclude the chapter by providing a summary of the models, a discussion highlighting the need to systematically validate the model outputs, and a closing discussion about the advantages and limitations of the Party World models.

4.1 Agent-based modelling: a review of the terminology

There is some terminology involved in specifying and operationalising agent-based models. In this section, using literature evidence, we describe the terminology relevant to our models presented in later sections.

Agent-based modelling is a recognised approach to model, analyse and understand the behaviour of real-world phenomena. The process or phenomenon that an agent-based model seeks to explain is referred to as the *target* of the system (Gilbert and Troitzsch, 2005, p. 14). For example, the flocking of birds and the schooling of fish are the target phenomena explained by the flocking model (Reynolds, 1987) and the movement of fish schools model (Huth and Wissel, 1992). The target behaviour we seek to explore and understand is the social and spatial dynamics of human conversational groups.

Agent-based models have two key entities, one of which is *agents* that represent social actors, such as humans, animals, companies, etc. (Gilbert, 2008). Agents are defined as computer programs with *attributes* and *methods* (Macal and North, 2010). Attributes define the characteristics of an agent. Attributes may be static and unchangeable during the simulation (e.g. gender) or dynamic and changeable as the simulation progresses (e.g. location). Attributes can also be *state variables* that reflect the state of an agent at any time during the simulation (Railsback and Grimm, 2011). Methods govern the behaviour of agents and are made up of simple rules that determine how agents sense and interact with one another and their environment. Specifying the attributes and methods allows agents to act autonomously during the simulation (Macy and Willer, 2002).

The *environment* is the virtual world where agents reside, act and interact. The environment may be an entirely neutral medium with no effect on the agents, or may simply be used to provide information on the spatial location of an agent relative to other agents, or a sophisticated medium providing a rich set of information to influence and constrain the actions of agents in the system (Macal and North, 2010). The environment can represent a geographical space (spatially explicit models) or an abstract space (Gilbert, 2008). In spatially explicit models, agents have a dynamic attribute to keep track of their location. The value of this attribute changes as the agents move in space.

Relationship is the aspect that links agents and their environment by defining how and with whom agents interact (Macal and North, 2010). It is often unrealistic to assume that all agents interact with every other agent in the system. Hence the notion of interactions within a neighbourhood is used. An agent's neighbours are a subset of the agent population with whom the agent interacts at a particular time step. An agent can interact with other agents in its immediate physical environment (i.e. physical neighbourhood) or with those in its social circle comprised of relatives, family, friends, etc. (i.e. social neighbourhood). An agent's neighbourhood and set of neighbours changes as the simulation progresses and agents move through space (Macal and North, 2010).

Interactions enable agents to exchange information localised to their neighbourhood. The flow of information can be either direct or indirect. For example, spoken dialogue represents a direct flow of information, whereas agents observing one another signifies an indirect exchange of information (Gilbert, 2008). Interactions can alter the dynamic attributes and/or behaviour of agents. For example, in the flocking model, each boid alters its speed and heading direction based on observing other boids and obstacles in its local neighbourhood (Reynolds, 1987, 2001).

Sensing provides the means for agents to perceive other agents and entities in their neighbourhood. Agents are programmed to respond to the information they perceive of their neighbourhood. Hence, depending on the target phenomenon being addressed, it is important to decide what information agents perceive and how they perceive it (Railsback and Grimm, 2011). The type of information held by agents can be: (1) complete or incomplete and (2) correct or incorrect. Complete information means an agent has access to all the variables of its neighbours, whereas incomplete information means limited access to neighbours' characteristics. Correct information means an agent has accurate information regarding its neighbours, whereas incorrect information is imperfect. In terms of the strategy used to obtain information, the assumption may be that agents have ready access to the variables of their neighbours, or that they make a formal request to obtain the information.

In some models, *memory* is a dynamic attribute of agents, which allows them to remember details of their past actions and interactions. An agent's memory can also hold information about its neighbourhood, e.g., an agent can maintain and update a list of its neighbours.

Lastly, an agent-based model is implemented as a computer simulation, which is run to generate data about the target phenomenon that the model intends to represent virtually. The simulated data is then compared with real-world data about the target phenomenon. This suggests if the model produces outcomes similar to the real-world processes. If it does, we can reach a better understanding of how agent-level behaviours lead to the emergence of group-level (or system-level) behaviour that corresponds to reality. This logic of using agent-based modelling as a method of scientific inquiry is described in Gilbert and Troitzsch (2005) (see figure 4.1). In the following section we present the conceptual design of the agent-based models we have developed to simulate human conversational groups.

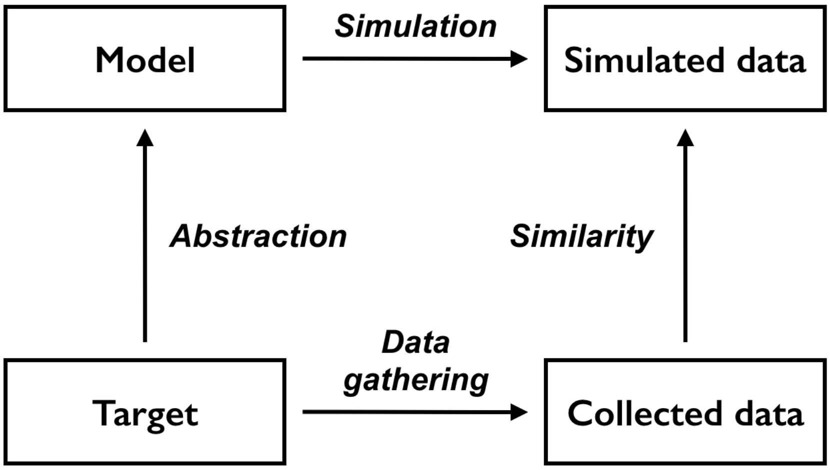


Figure 4.1: Gilbert and Troitzsch (2005) logic of simulation as a method

4.2 The conceptual design of models

We have conceptualised three different models, as noted before – Model 1, Model 2a and Model 2b, to simulate the emergence of conversational groups in a two-dimensional virtual world. At the start of the simulation agents begin to wander around the virtual world. They only move in the forward direction. When agents come close to one another, they stop moving. After remaining stationary for some time, agents then resume movement, either in the same heading direction, or in a new direction. The continuous interval of time during which a cluster of agents remain stationary signifies a face-to-face encounter. Each of the three models described here uses a different strategy to control the positioning and orientation of agents engaged in face-to-face encounters.

4.2.1 Model 1

Model 1 has the simplest approach: agents do not readjust their position or orientation in a face-to-face encounter. The formulation of Model 1 is as follows. Agents wander around in a two-dimensional space starting from random initial locations. Agents need to maintain a certain distance between one another to keep moving.

If the relative distance between an agent and other agents falls below a predetermined separation distance, the agents stop moving, and remain stationary for a predetermined stop time.

When the stop time is reached, agents will resume movement either in the same direction (i.e., along the current heading direction) or in a new direction. An agent will continue moving as long as the relative distance between itself and other agents does not fall below the separation distance. If it does, the agent will stop moving, again. In this manner, agents exhibit alternating phases of movement and stationary behaviour, for the entire duration of the simulation. Figure 4.2 is a flowchart summarising the sequence of actions performed by each agent in Model 1 at every time step of the simulation.

Model 1 is a conceptually simple approach to simulate conversational groups: when agents arrive close to one another, they stop moving and remain stationary for some time. During the encounter, agents do not change their position or orientation, but prior to the encounter, movement is in the forward direction. The purpose of Model 1 is to assess if the spatial patterns of agent clusters, which arise naturally from the forward-steering movement of agents, resemble human conversational groups.

4.2.2 Model 2a

Model 2a is similar to Model 1 in that agents demonstrate alternating phases of movement and stationary behaviour. But in Model 2a, agents also readjust their position and orientation, while they are in a stationary state[1].

The formulation of Model 2a is as follows. To begin with, agents start moving around from random initial locations. Agents stop moving when they come closer than the separation distance. When the stop time is reached, agents resume movement, either in the same direction or in a new direction. Up to this point, the logic is the same as in Model 1. The additional assumption is that, when agents stop

[1] The Oxford dictionary defines stationary as 'Not moving or not intended to be moved'. We use the term to refer to agents that are not moving in the sense of being engaged in a face-to-face encounter, but can still modify their position and/or orientation.

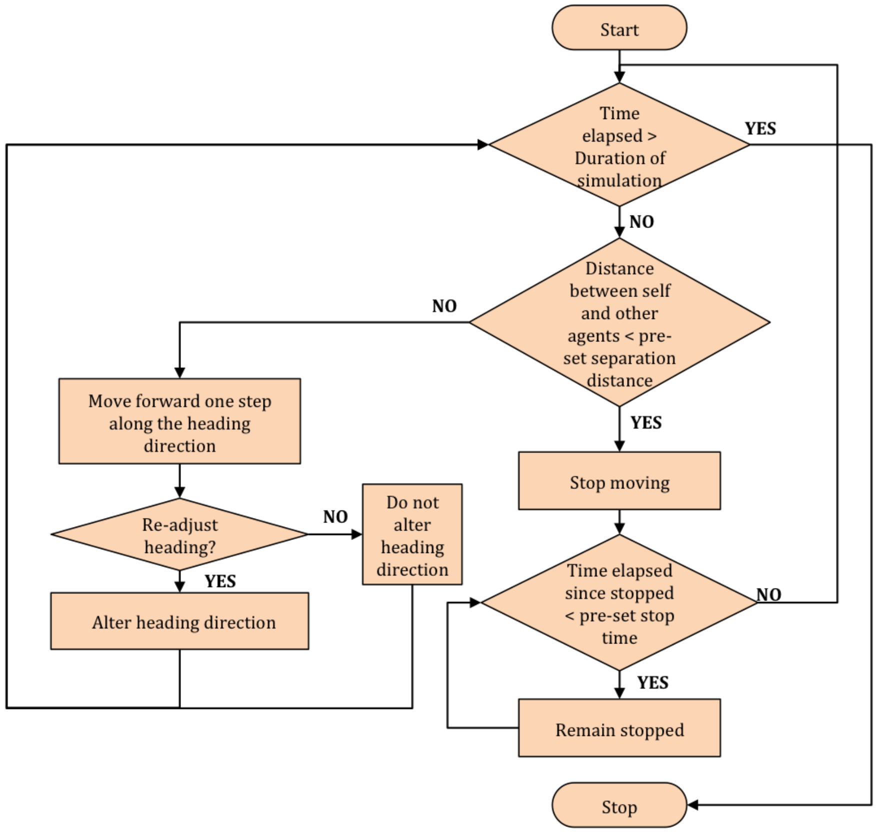


Figure 4.2: Flowchart of Model 1

moving, they alter their position and orientation to perceive as many neighbours as possible, within a 180° *field of view*. Here, field of view is calculated by taking as reference the line extending outwards from the centre of an agent's body, along its current heading direction. The field of view extends 90° on either side (left and right, L-R) of the line of reference.

The important aspect of Model 2a is that, an agent's goal is to move to a position and an orientation that enables perceiving as many neighbours as possible, within its field of view. Consequently, the extent to which an agent alters its position and orientation at a particular time step, will depend on who and where the agent's neighbours are at the time. An agent performs fewer readjustments if the relative position and orientation of its neighbours fall within its field of view. On the other hand, if neighbours are outside the field of view, or worse, somewhere behind, the agent then needs to perform a lot of changes.

The model has no global control over the coordination of readjustments – each agent acts autonomously, and on an egocentric basis, trying to perceive as many neighbours as possible at any time. Furthermore, it has to be remembered that an agent, in a stationary state, alters its position and orientation relative to its neighbours, which are also in a stationary state. Consequently, each agent within a cluster will be readjusting its position and/or orientation at the same time as its neighbours. The result will be a wave of changes, not only in the position and orientation of each individual agent, but also in the overall spatial pattern of the cluster itself. The purpose of Model 2a is to examine if the social dynamics and the patterns of agent clusters, resulting from the field-of-view based readjustment behaviour of agents, resembles human conversational groups.

In addition to the readjustment behaviour, Model 2a is applied in two scenarios, each with a different definition of neighbourhood. Earlier we described that an agent's neighbours are a subset of the population with whom it interacts on a particular occasion. Model 2a is applied in two scenarios: one where an agent's neighbours are only those that are immediately adjacent to it (i.e. those that are closer than the separation distance), and another where an agent's neighbours and their neighbours form an extended neighbourhood. The former is called the local neighbourhood scenario (Model 2a LA) and the latter is called the extended neighbourhood scenario (Model 2a EA).

In agent-based modelling, *scenario contrast experiments* are a way of testing the results of two or more scenarios of the same model. Railsback and Grimm (2011) compare scenario contrasts to the experimental design used to test real systems, where the target is subject to two or more different treatments, to determine the effects of each treatment. In the context of computer simulations, scenario contrast experiments are used to explore and analyse if, and how, the results vary based on the difference in agent behaviour across scenarios. The two scenarios presented here – Model 2a LA and Model 2a EA – are intended to provide insight on the influence of neighbourhood effects on the social dynamics and spatial patterns of agent clusters.

The flowchart of Model 2a in figure 4.3 summarises the sequence of actions performed by each agent at every time step of the simulation. To begin with, agents start moving from random initial locations. They stop moving if the separation distance condition is violated. While agents are in a stationary state, they readjust their position and orientation relative to their neighbours. In Model 2a LA, agents readjust their position and orientation to perceive as many local neighbours as possible, within their field of view. On the other hand, in Model 2a EA, agents readjust to perceive as many extended neighbours as possible. Agents will resume movement once the stop time is reached. They will keep moving, either in the same direction or in a new direction, as long as they do not get closer than the separation distance to any other agent in the system. In this manner, the phases of movement, stationary and positional-orientational readjustment behaviour will continue for the entire duration of the simulation.

4.2.3 Model 2b

Model 2b is our adaptation of the centroid-based approach to control the positioning and orientation of agents engaged in face-to-face encounters. As with Model 2a, agents demonstrate alternating phases of movement, stationary and readjustment behaviour, but the strategy used for governing the changes in the position and orientation of agents is different.

In model 2b, each agent in a stationary state calculates a centroid point, by averaging the (x,y) position values of itself and its neighbours. The agent then self-regulates its position to arrive at an equidistant point from the centroid. The agent also adapts its orientation to face the centroid point. The equidistant point from centroid is calculated by averaging the distances between an agent and the centroid point and the agent's neighbours and the centroid point. An agent moves towards the centroid point if its current position is beyond the equidistant point, whereas it moves away, if its current position lies inside the equidistant point.

Like Model 2a, Model 2b is also applied in two scenarios: a local neighbourhood scenario (Model 2b LA) and an extended neighbourhood scenario (Model 2b EA). In Model 2b LA, an agent calculates the centroid point by averaging the positions of itself and its immediate neighbours, whereas in Model 2b EA, the centroid point is calculated by averaging the positions of an agent, its neighbours, their neighbours and so on.

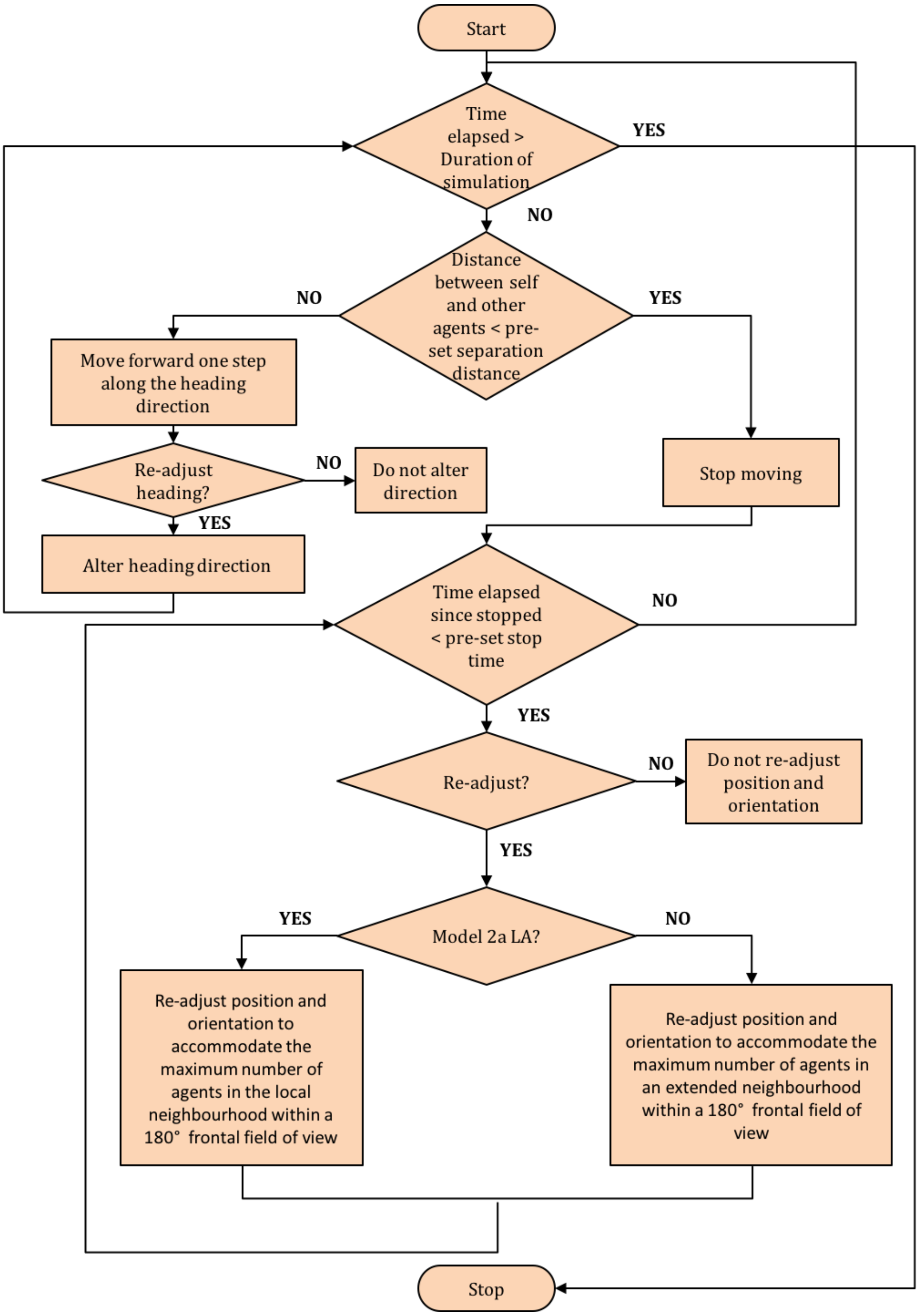


Figure 4.3: Flowchart of Model 2a

The sequence of actions performed by agents in Model 2b is as follows. Agents start moving around from random initial locations. They stop moving when they encounter other agents at a distance closer than the separation distance. When agents are in a stationary state, they readjust their position and orientation based on the centroid point. In Model 2b LA, agents calculate the centroid point based on a local neighbourhood, whereas in Model 2b EA, agents calculate the centroid based on an extended neighbourhood. As long as agents remain in a stationary state, they continue readjusting their position and orientation relative to the centroid point. Agents resume movement, either in the same direction or in a new direction, once the stop time is reached. They continue moving until when the separation distance condition is violated. The alternate phases of movement, stationary and readjustment behaviour continues for the entire duration of the simulation. Figure 4.4 shows the sequence of actions performed by each agent in Model 2b.

4.3 Model implementation

In this section we describe the implementation of the five models introduced above. First we describe the visual appearance and the characteristics of agents in sections 4.3.1 and 4.3.2, respectively. We then describe the environment and the Party World context of the simulations in section 4.3.3. We describe the software used for implementing the simulations in section 4.3.4. We end the section by reporting the details of the simulation runs.

4.3.1 The visual appearance of agents

In our simulations, agents represent humans. Kendon (1990) used bird's eye view depictions of the human body to explain the concept of F-formation systems – e.g., figure 4.5b is the bird's eye view representation of the F-formation and its interactants shown in figure 4.5a. Although the two figures denote the same F-formation, the individual frames themselves are not a one-to-one match. That is to say, frame 1 in figure 4.5b is not an equivalent representation of frame 1 in figure 4.5a. Nonetheless, the images are two different representations of the same F-formation.

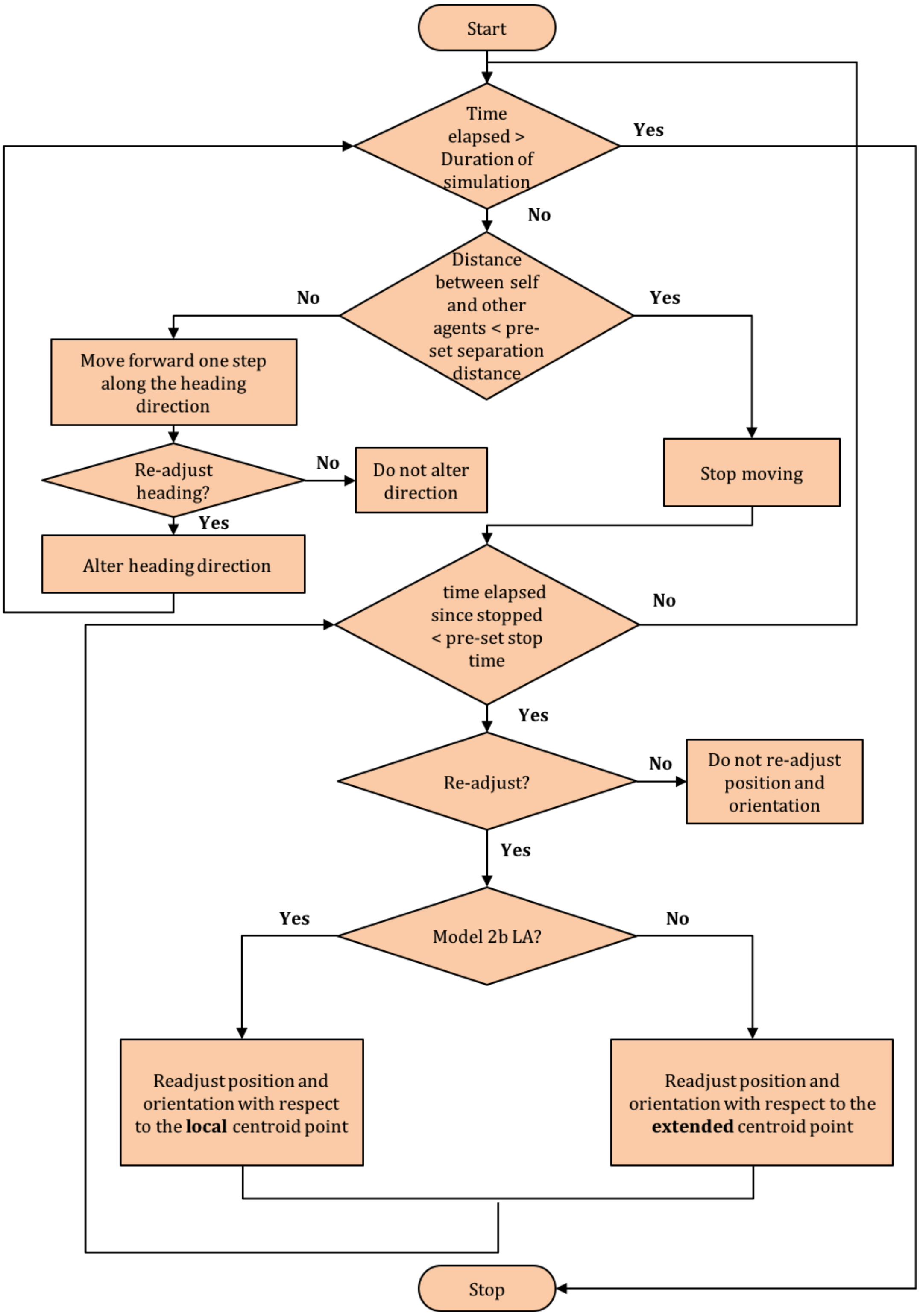
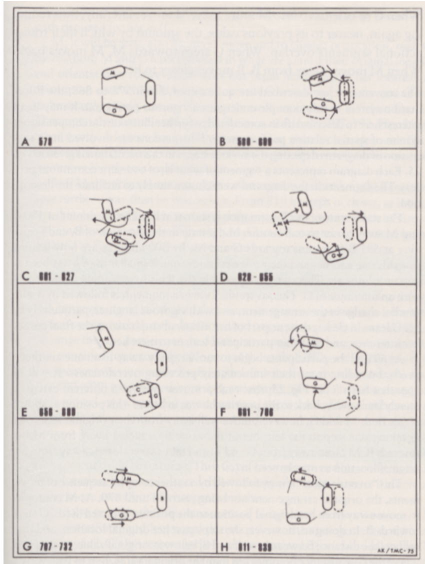


Figure 4.4: Flowchart of Model 2b

(a) Photographs showing an F-formation (Deutsch, 1979)



(b) Bird's eye view drawings of the F-formation in figure 4.5a (Kendon, 1990)

Figure 4.5: Photographs and bird's eye view drawings of an F-formation

The former is a series of photographs, which clearly identifies the participants in the F-formation, as well as aspects such as gaze, gesture and stance, in addition to body position and orientation. The latter is a series of drawings that specifically directs our attention to interactants' body position and orientation, and thereby, to the spatial manifestation of the F-formation.

Comparing figures 4.5a and 4.5b, it may be argued that the bird's eye view drawings provide a much clearer picture of the spatial patterns of the F-formation than the photographs. As our objective is to investigate the causal role of body position and orientation in the spatial arrangement of conversational encounters, we have chosen to represent agents using a bird's eye view appearance of the human body in our simulations.

Figure 4.6a shows the visual appearance of an agent in our simulations. There is a circle at the centre, which displays the unique id of each agent (a number). The elongated structure at the front centre is the nose, which indicates an agent's heading direction. The ellipsoidal black colour projections on either side of the circle are the shoulders. The shoulder-to-shoulder width of an agent is roughly 30 pixels.

The width was chosen so that the agent appears neither too small nor too large in the virtual world. Other dimensions such as the height of the shoulders and the height and width of the nose were chosen proportional to the shoulder-to-shoulder width.

We chose to draw a trapezoidal transactional segment extending outwards from an agent's body whenever it stops moving (see figure 4.6b). In chapter 2 we saw that every individual has access to a use space or a transactional segment, extending outwards from the body, to which the individual has uninterrupted access to carry out a desired activity. The trapezoidal shape is inspired by Deutsch's (1979) denotation of a transactional segment, where a shorter edge, called the base, is parallel to a longer edge, called the crown, and there are two legs connecting the base to the crown, on either side, at a 45° angle (see chapter 2). In our simulations, the dimensions of the trapezoidal transactional segment were chosen to be proportional to the agent's body. The agents don't have a transactional segment when moving, but have one when they are in a stationary state.

(a) Head and body (b) Transactional segment

Figure 4.6: The visual appearance of agents in our simulations

4.3.2 The characteristics of agents

Each agent in the system has the following state variables:

1. **ID:** A unique number to identify the agent. ID is a static variable; an agent's ID does not change as the simulation progresses.

2. **LOC:** A dynamic vector variable denoting the agent's (x,y) position with respect to the origin (0,0) in the virtual world. At the start of the simulation, LOC is assigned

a random (x,y) value within the bounds of the simulation window. The LOC value changes as the agent moves in the virtual world. The (x,y) coordinates of the centre of the circle bearing the ID of an agent (e.g. see figure 4.6a) is the LOC value.

3. **ROT:** A dynamic variable denoting the agent's heading direction with respect to the 0° orientation. At the start of the simulation, ROT is assigned a random value in the range of [-180°, 180°]. The ROT value changes whenever the agent alters its orientation. The pointing direction of an agent's nose is the ROT value.

4. **STATE:** A dynamic variable to keep track of the movement and stationary states of the agent. The STATE variable is updated depending on whether an agent is moving (STATE =“none”) or not moving (STATE =“stopped”) at a particular time step. At the start of the simulation, the value of the STATE variable is set to “none” for all the agents – i.e. no agent will be in a stationary state at the start of a simulation.

Sensing is a key characteristic of agents in our simulations. The sensing mechanism used allows an agent to perceive other agents in the environment, to decide whether to continue moving or to stop moving. At every time step, an agent queries the environment to obtain the LOC values of other agents, based on which it calculates the Euclidean distance between itself and others. The agent stops moving if the distance calculated is less than the predetermined separation distance. As an example, figure 4.7 shows agent no. 10 sensing the presence of other agents (12 and 14), within the radius of the separation distance. Agents 10, 12 and 14 all stop moving as soon as the distance separating them falls below the separation distance.

Memory is another characteristic of agents in Models 2a and 2b. In these models, agents have an additional dynamic attribute called INTERLOCUTORS, which is a list variable to keep track of an agent's neighbours. The list is initially empty, and it remains so when agents are moving. But when an agent is stationary, the list stores the immediate neighbours of an agent (in the LA scenarios) or its extended neighbours (in the EA scenarios). The information in this list will be used to govern

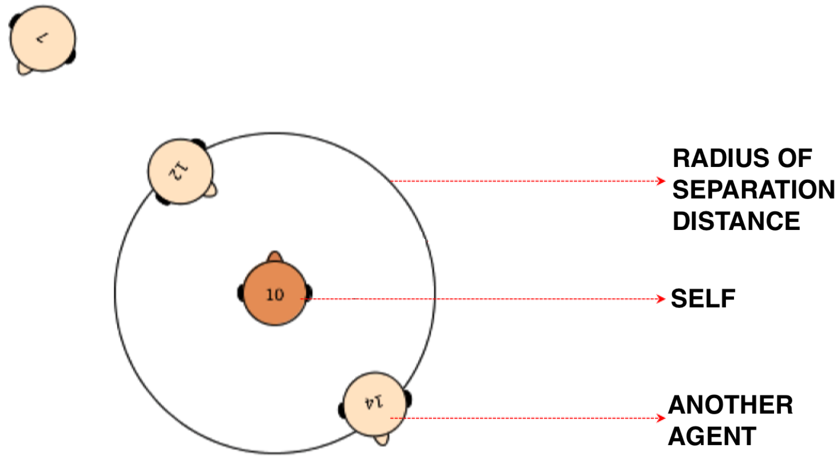


Figure 4.7: Demonstration of an agent sensing other agents near it

the readjustment behaviour of agents in Models 2a and 2b[2]. The INTERLOCUTORS list is a dynamic variable, because it will be updated whenever an agent's neighbours change, i.e., when new agents join or current neighbours leave. The list will be emptied when the agent resumes movement, and populated again when it stops moving. In Model 1, agents do not maintain a list of interlocutors, because readjustment behaviour is not included in the model.

4.3.3 The environment and the Party World context

The environment in our simulations is a two-dimensional Euclidean space that is 1200 pixels wide and 700 pixels high. The width and height were chosen on a trial and error basis so the environment is neither too big nor too small. We have used an off-white shade as the background colour for the environment. The space is bounded, i.e. the edges do not wrap around. Hence, when agents reach the boundaries of the simulation area, they alter their heading direction to remain within the bounds of the environment.

We needed a meaningful physical and social context for the simulations, because

[2] The models could function without using the INTERLOCUTORS variable, i.e. by detecting who the neighbours are at every time step of the simulation. But using this variable utilises the concept of memory (see section 4.1) and improves the program speed.

the size, shape and the number of conversational groups, are influenced by the physical settings of the environment, and the surrounding social context (see chapter 2). For example, in a narrow corridor or a street pavement, we will usually find fewer and smaller conversational groups. On the other hand, in a party-type social setting, several conversational groups may coexist. We consider the latter context for our simulations. We implemented a Party World environment that is free of furniture, or any other objects, so the agents can move around and interact freely. Due to the party context, our simulations are referred to as *Party World* simulations.

4.3.4 The simulation platform

Choosing an appropriate simulation platform is an important step in agent-based modelling (Gilbert, 2008, p.22 and p.46). We initially implemented the Party World simulations using Greenfoot[3], an open-source Java based programming environment used to build games, agent-based simulations and other graphical visualisations. However, we encountered some limitations when using Greenfoot, consequently, we chose the Processing[4] open source programming language and integrated development environment (IDE) for implementing the simulations. Each Party World model simulation is implemented as a Processing sketch.

4.3.5 The model outputs

The outputs and displays required to understand and test an agent-based model are chosen depending on the target phenomenon, which the model is designed to replicate (Railsback and Grimm, 2011). The Party World models and simulations are intended to replicate the spatial-orientational behaviour of interactants engaged in face-to-face encounters, and the spatial form and the social dynamics of encounters. In terms of display, looking at our simulations, where agents exhibit alternating phases of movement, stationary and readjustment behaviour, is intended to emulate the effect of observing a real-world situation, from a bird's eye view, where people move around and engage in face-to-face encounters. In terms of output, screenshots

[3]http://www.greenfoot.org/home
[4]https://processing.org/

of the simulation window, and the position and orientation of agents engaged in face-to-face encounters (referred to as *interactant agents* are recorded). The screenshots and data are used in subsequent chapters to validate the emergence of realistic conversational group arrangements.

4.3.6 Specifics of the simulation runs

Table 4.1 lists the parameters and values used as input to the Party World simulations. The first parameter, *agent population*, reflects the total number of agents in the simulations. The value of this parameter is set to 20, so there are neither too many nor too few agents within the 1200 pixels x 700 pixels boundary of the simulation area[5]. The value of the second input parameter *separation distance* is set to 45 pixels. The value has been set at 1.5 times the shoulder-to-shoulder width of an agent (i.e. 30 pixels). Agents stop moving if the distance separating them from any other agent(s) falls below 45 pixels at any point during the simulation.

The *stop time* parameter, which denotes the duration for which agents remain stationary before resuming movement, is set as 30 seconds (real time). The following calibration procedure was used to determine the value. Encounters were brief, and the spatial patterns of agent clusters did not stabilise, when stop time was set to a value much smaller than 30. On the other hand, values much higher than 30, produced similar clustering effects as seen when the value was set at 30. Due to these outcomes, we initialised stop time as 30 seconds in the Party World simulations.

In Models 2a (LA and EA) and 2b (LA and EA), the *probability of readjustment* controls the frequency at which agents alter their position and orientation in the stationary state. We assigned a 3% probability that each agent will readjust its position and orientation at every time step of the simulation. The value was calibrated on a trial and error basis, but the same value was used in all the models, and we ran the simulations multiple times to account for the uncertainties of the input parameters.

The screenshot sampling interval is 5 seconds – i.e. screenshots of the simulation window showing the agent clusters are recorded once every 5 seconds. The data

[5]We ran the simulations with 10 and 30 agents, it appeared empty with only 10 agents and crowded with 30 agents, so we set the agent population to 20.

sampling interval is also 5 seconds, i.e., the LOC and ROT values of agents engaged in face-to-face encounters are also exported at 5-second intervals, at the same time as the screenshots are made.

As noted before, the LOC and ROT values of each agent are assigned randomly, at the start of the simulation. The x-component of the LOC variable is assigned a random value between [0, 1200] (i.e. width of the simulation area) and the y-component is assigned a random value between [0, 700] (height of the simulation area). The ROT variable is assigned a random value between [-180°, 180°]. The LOC and ROT variables change as and when agents move in the two-dimensional virtual world. The Party World models use a steering algorithm to control the motion of agents. A steering force acts on each agent causing it to move at a particular speed in a particular direction. Changes in the steering force, which may be induced randomly or when agents reach the boundaries of the simulation area, will cause agents to alter their state, speed and/or heading direction.

The Party World simulations are given an initial *settling period* of 100 simulation time steps during which agents do not engage in face-to-face encounters. The settling period is allowed at the start of the simulation so that agents are able to stabilise their course of motion. For example, agents might just happen to be closer than the separation distance due to the randomly initialised LOC values, which in turn induces the clustering process right at the start of the simulation. We have the settling period to avoid such situations. During the settling period, agents keep moving in the chosen (ROT) direction, but if they come closer than the separation distance, they will steer away from one another to avoid bumping. We have adapted Reynolds's (1987) concept of separation force to implement the effect of steering away.

Using the list of input parameters specified in table 4.1, we ran each Party World model 5 times. The number of runs were chosen based on the fact that the spatial patterns and dynamics of agent clusters generated were similar across runs but varied across models. However, we still performed five runs of each model, to take into account the effect of the stochastic input parameters. The duration of each

Input parameters	**Values**
Agent population	20
Separation distance	45 pixels
Stop time	30 seconds
Probability of readjustment	p = 0.03
Screenshot sampling interval	5 seconds
Data sampling interval	5 seconds
Duration of simulation	30 minutes
LOC (x,y)	$x \in \{0, ..., 1200\}, y \in \{0, ..., 700\}$
ROT	$\{-180°, ..., 180°\}$
Settling period	100 simulation time steps

Table 4.1: Input parameters to the Party World simulations

run was set to 30 minutes. The temporal extent of agent-based models are normally chosen based on the system-level phenomena or observations of interest (Railsback and Grimm, 2011). The target phenomena we wish to analyse are the dynamics and patterns of agent clusters resulting from the different models. As all the five Party World models produced agent clusters, almost immediately after the settling period, we limited each run to 30 minutes. While determining the length of the simulation runs, we also factored in the observation that 15-minute runs were too short in a party context, and the outcome of 60-minute runs was not radically different from the outcome of 30-minute runs.

4.4 Party World simulations: A glimpse of agent clusters

In this section we sample one screenshot from each Party World model simulation to provide: (1) an example of the simulated world and (2) an overview of the spatial arrangement of agent clusters resulting from the different models. While looking at the screenshots, the trapezoidal transactional segments can be used to distinguish interactant agents from the ones that were moving at the time the screenshots were made.

Figure 4.8 is a screenshot obtained from Model 1. A brief visual inspection indicates there are about 6 to 7 agent clusters in the image. In some cases, the spatial patterns delineated by the positional-orientational arrangements of interactant agents are readily recognisable as conversational groups. For example, the three dyadic clusters, one at the bottom center with agents {2,8}, one at the top right with agents {7,19}, and another at the bottom left {4,17}, show agents facing one another in C- or V-shaped arrangements (see classification of arrangements in section 2.7). Another dyadic group at the top center, comprised of agents {9,18}, depicts a side-by-side arrangement.

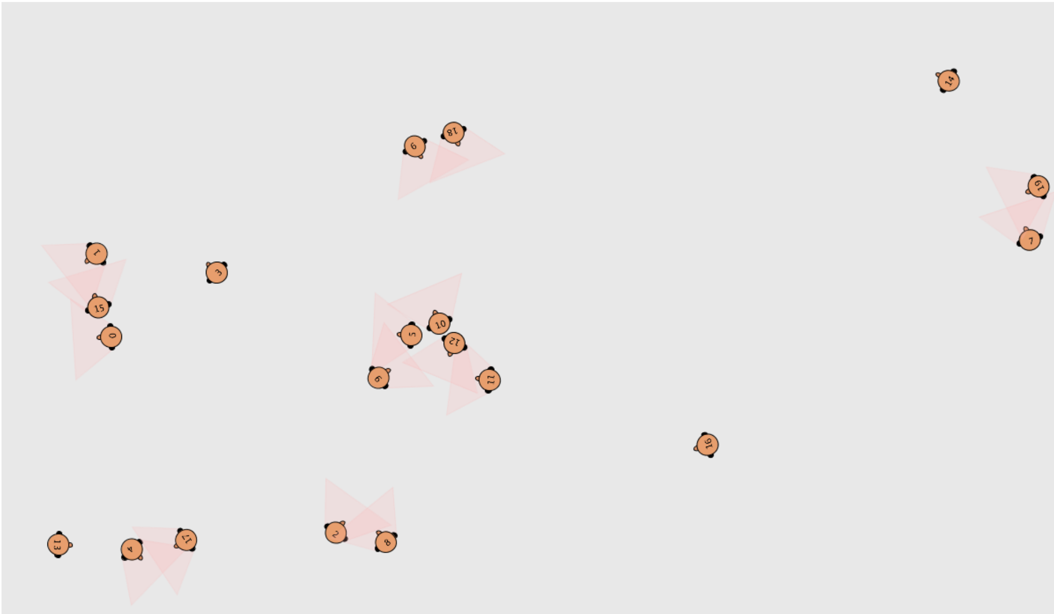


Figure 4.8: Screenshot of Model 1

While the spatial patterns of dyadic clusters are quite identifiable, the ones of multi-party arrangements are not as straightforward. For example, in the triadic cluster at the center left of the image with agents {0,1,15}, 1 and 15 are facing one another, but agent 0 is facing away from the group. It comes across as if neither agent 0 nor agent 15 have acknowledged the presence of one another despite standing next to each other. The arrangement of the big group at the center of the image with agents {5,6,10,11,12} is also not clear. Either it could be perceived that agents 5 and 6 are a dyad, as are agents 11 and 12, while agent 10 is standing near but facing away from both groups. Or it could be perceived that the agents are all in one big group, with agents 6, 11 and 12 facing one another more or less, whereas agents 5 and 10 are clearly facing away from the group.

Unlike the case of Model 1, the screenshot from Model 2a LA (figure 4.9) does not show agents standing near, and yet facing away from one another. A brief inspection of the image indicates about 5 to 7 agent clusters in the image. There is a triadic arrangement of agents {1,3,10} at the top left of the image. Agents 1 and 3 are seen facing inwards, while agent 10 is facing slightly away from the group, but the transactional segments of all three agents are overlapping. Agents 8 and 9 are in a seemingly vis-á-vis arrangement at the top right of the image. Agents 7 and 19 are in a L-shaped arrangement at the center right of the image.

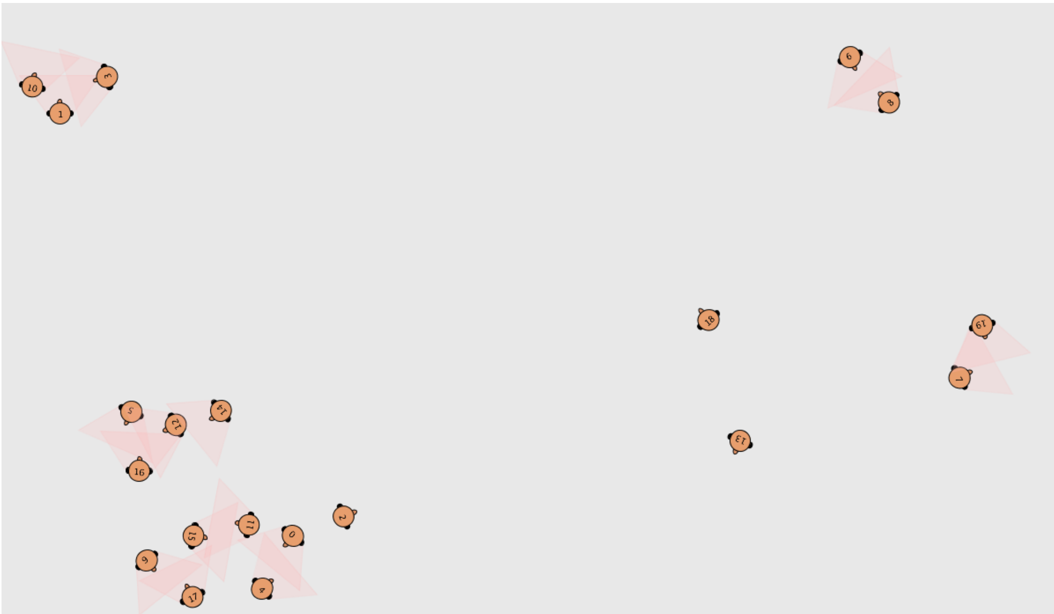


Figure 4.9: Screenshot of Model 2a LA

There are two big agent clusters in figure 4.9. In the first one, agents 5, 12 and 16 are all facing inwards and their transactional segments are overlapping, while agent 14 is in the outsider position (see section 2.6.1.7). The big cluster at the bottom left of the image with agents {0,4,6,11,15,17} appears to be a trio of dyads – {0,4}, {6,17}, {11,15}.

The screenshot obtained from Model 2a EA in figure 4.10 shows five distinct agent clusters. The three dyadic arrangements at the bottom of the image – {10,17}, {5, 19} and {3,8} are in what could be perceived as C-, V- and L-shaped arrangements. Agents {1,6,7,9} are in a semi-closed arrangement, where agents are facing one another, and their transactional segments are overlapping. Agents {0, 2, 11, 13, 15, 16, 18} are a big group in a closed arrangement where agents are all facing inwards.

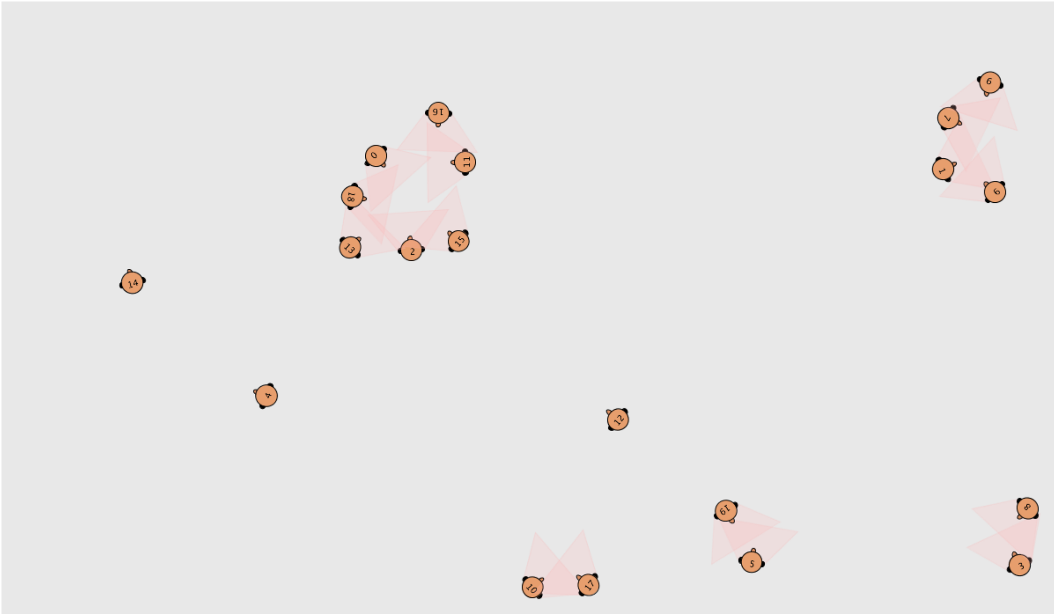


Figure 4.10: Screenshot of Model 2a EA

The screenshot obtained from Model 2b LA shows about 7 to 8 agent clusters (figure 4.11). Depending on how it is looked at, there is a one four member cluster or two dyadic clusters at the far left of the image. Agents 0 and 2 are facing one another vis-á-vis, as are agents 4 and 15. Similarly agents {3,5} are in a vis-á-vis arrangement towards the right center of the image, as are agents {6,9} at the top right corner of the image. Agents {8,10,18} are in a perfectly triangular arrangement towards the right center of the image. Agents {7,14,16,19} are all in one group at the top right of the image. In both groups, {8,10,18} and {7,14,16,19}, agents are all facing inwards and their transactional segments are overlapping. The big cluster of agents at the top center of the image appears to have two sub-groups: one dyadic cluster in a vis-á-vis arrangement ({12,17}) and another triadic cluster in a triangular arrangement ({1,13,11}).

The screenshot obtained from Model 2b EA shows 7 distinct agent clusters in

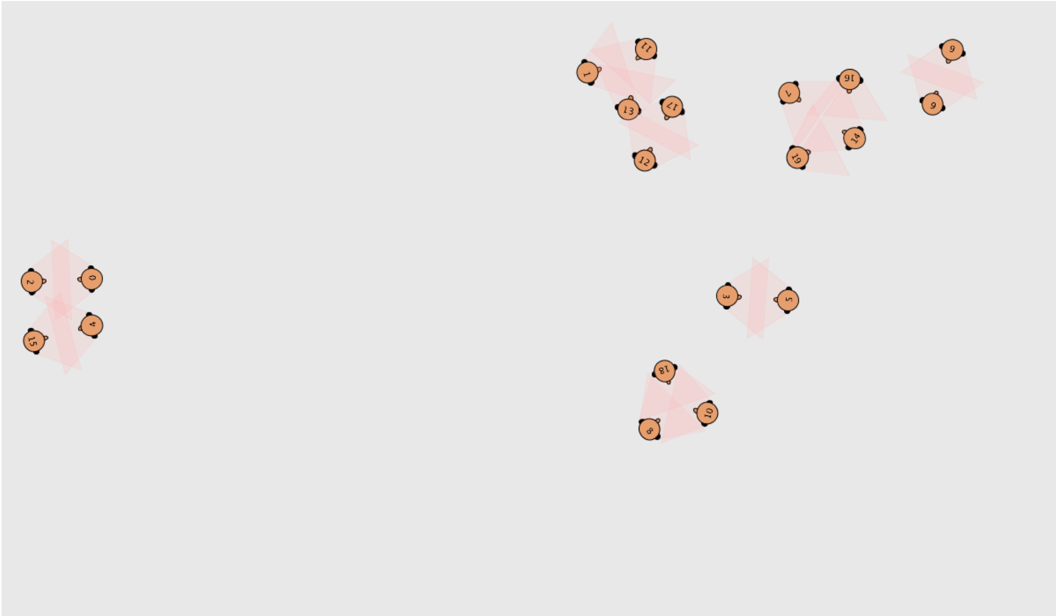


Figure 4.11: Screenshot of Model 2b LA

figure 4.12. Agents {0,18}, {1,2}, {5,8},{12,13} and {14,19} are in dyadic vis-á-vis arrangements. Agents {4,7,10}, at the center of the image, are in a triadic arrangement. Just next to the triadic arrangement are agents {3,6,9,11,15,16,17} in a circular arrangement. Interactant agents are facing one another and their transactional segments are overlapping in all seven clusters in figure 4.12.

The screenshots reviewed in this section are an overview of the different types of agent clusters resulting from the different Party World models. However, more importantly, the images raise further questions regarding the performance of models. For example:

1. We need to understand the reasons for the ambiguities seen in the multi-party arrangements of agent clusters resulting from Model 1 (figure 4.8). There could be two potential explanations for the ambiguities. Either the spatial arrangement of agent clusters were not fully formed at the time of the screenshot. Or given that the dyadic arrangements in the image look plausible, it could be that Model 1 just does not produce plausible multi-party arrangements, because interactant agents do not adjust their position or orientation.

2. The dyadic arrangements in figure 4.11 and figure 4.12 from Model 2b LA and Model

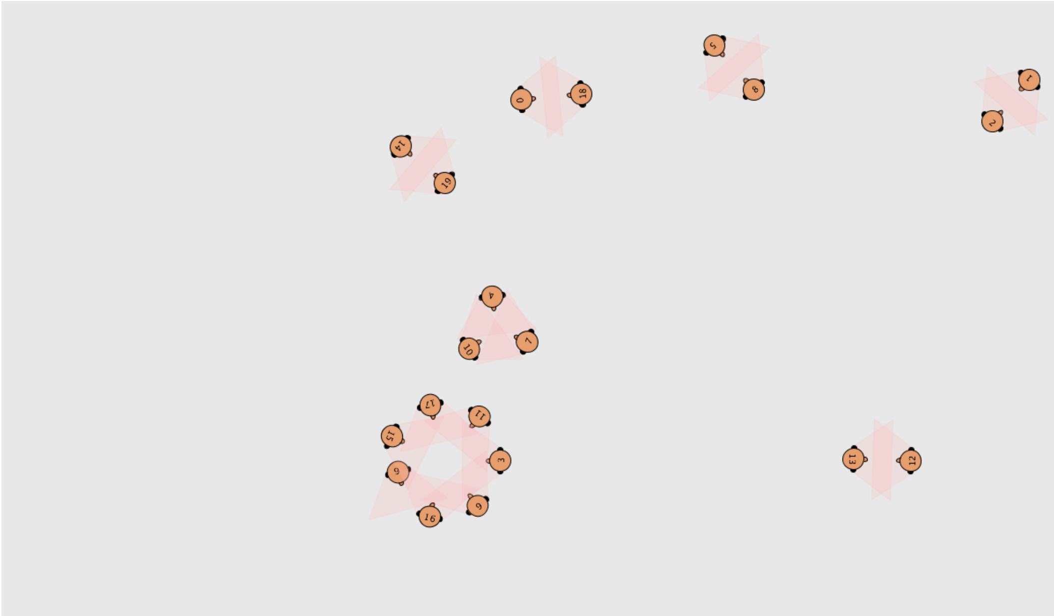


Figure 4.12: Screenshot of Model 2b EA

2b EA are all vis-á-vis, whereas the dyadic arrangements resulting from Models 1, 2a LA and 2a EA assume different shapes, such as C-, V-, L- and side-by-side arrangements (cf. figures 4.8, 4.9 and 4.10). Hence we need to determine if and why Model 2b (LA and EA) always produces vis-á-vis arrangements, while others produce different types of dyadic arrangements.

3. The spatial pattern of the seven member agent cluster seen in figure 4.10 and 4.12 are different – the former is not a perfect circle whereas the latter is. We need to examine the dynamics of agent behaviour that resulted in these patterns, and determine if these dynamics are realistic. We also need to find out which of the two closed group patterns correspond more closely to reality.

4. We need to analyse how group formations end. The stop time value used in the simulation runs is 30 seconds. We need to find out if the predetermined stop time causes all agent clusters to last only for 30 seconds or longer than that. If agent clusters last longer than 30 seconds, we need to assess how and when agents break off and steer away from their groups.

These sort of questions have never been addressed before, but are essential for understanding the ability and comprehensiveness of models to simulate realistic conversational group patterns. In the subsequent results chapters 5, 6 and 8, our objective is to address these questions, by using a systematic approach to validate the spatial patterns and dynamics of agent clusters resulting from the different models

4.5 Summary

In this thesis, we use an agent-based approach to model, explore and understand the formation and maintenance of conversational groups from the constantly adapting positional-orientational relationship of interactants. Our approach is inspired by the work of Reynolds (1987), where realistic flocking patterns were shown to result from three simple rules of boid behaviour: separation, alignment and cohesion. The models are built in an incremental fashion just as Braitenberg built increasingly complex vehicles (Braitenberg, 1986). More importantly, our approach draws inspiration from Railsback and Grimm (2011), who recommend starting from a simple null theory, and then testing alternative theories of behaviour, to identify the one that most closely reproduces the characteristics of the target phenomenon. In this chapter, we proposed and described three different agent-based models to simulate the movement, positioning and orientation of agents. Model 1 is the simplest approach to simulate the target phenomenon, while Model 2a and Model 2b are the field-of-view and centroid based approaches, respectively.

In all simulations, agents move around randomly until they encounter another agent or agents at close distances, at which point they stop moving. Different models follow different strategies to control the positioning and orientation behaviour of agents that are stopped. Model 1 does not alter the position and orientation of agents. Model 2a causes an agent to readjust its position and orientation relative to the positions and orientations of its neighbours. Model 2b causes an agent to readjust its position and orientation relative to a centroid point. Model 2a and Model 2b are both implemented in two scenarios: one where agents interact with and respond to a local neighbourhood, and another where agents interact with and respond to an extended neighbourhood. Taken together, we consider the following models:

1. Model 1: Baseline
2. Model 2a LA: Field-of-view based readjustments in a local neighbourhood
3. Model 2a EA: Field-of-view based readjustments in an extended neighbourhood
4. Model 2b LA: Centroid based readjustments based in a local neighbourhood
5. Model 2b EA: Centroid based readjustments based in an extended neighbourhood

The models are implemented in a party-type social setting, to explore the possibility of simulating multiple coexisting conversational groups found in social gatherings, such as a cocktail party or drinks reception. The visual appearance and characteristics of agents, the results obtained from the models, and the input parameters were described in section 4.3. A brief overview of the various kinds of spatial arrangements of agent clusters resulting from the models was presented in section 4.4. In subsequent chapters we will present a detailed analysis of the spatial patterns of agent clusters.

4.5.1 Benefits of our agent-based modelling approach

Our agent-based modelling approach has the following benefits. Firstly, it affords the opportunity to understand how a real world target process works by building smaller, less detailed and less complex models of the target (Gilbert and Troitzsch, 2005, p. 2). There are different aspects of verbal and non-verbal interaction involved in the functioning of conversational groups, e.g., speech, eye gaze, head and body position and orientation, gesture, etc. Among these different behaviours, we particularly wish to examine if and how the spatial position and orientation of interactants' bodies influences the spatial arrangements of conversational groups. Consequently, we have implemented agent-based models and simulations, where only the movement, positioning and orientation of agents are considered.

Second, agent-based modelling provides an electronic laboratory (Macal and North, 2010), to carry out experiments that may be expensive, difficult or infeasible to carry out in the real world (Gilbert, 1993, 2008). In our agent-based models, we isolate, manipulate and analyse the role of movement, position and orientation behaviour in the functioning of conversational groups. It may be difficult to perform such analysis in the real-world because human interactions are multimodal in nature. Even in contrived laboratory-based interaction experiments, participants might find it challenging to interact only using movement, position and orientation behaviour. Furthermore, agent-based models and simulations are inexpensive to develop and run, whereas real-world experiments can be expensive.

Third, agent-based modelling allows us to explore and analyse how realistic dynamics emerge from few simple agent behaviours, in unpredictable and non-linear ways (Railsback and Grimm, 2011). For example, the interactions between boids governed by just three simple rules (separation, alignment and cohesion), leads to the emergence of realistic flocking motion (Reynolds, 1987). An F-formation system is referred to as an emergent phenomenon, resulting from the coordination of spatial-orientational behaviour of interactants (Kendon, 1990). In this thesis, we use the agent-based modelling approach to address questions about the emergence of believable conversational group arrangements (like F-formations), from simple rules that govern the movement, positioning and orientation of agents.

Fourth, agent-based modelling enables accurate theorisation of a target process. An agent-based model must be specified clearly using a suitable programming language, toolkit or development environment (Gilbert, 1993). Consequently, unlike textual descriptions of a theory, agent-based model specifications allow less room for ambiguity or misinterpretation (Gilbert, 2008). Each of the three models described in this chapter uses a different theory to control the movement, positioning and orientation behaviour of agents. The models are implemented as computer simulations that are run to observe the resulting behaviour of agents. By doing this, a clear understanding of how the different models influence the emergence of believable arrangements of conversational groups is gained.

Lastly, validation is a crucial stage in the development of agent-based models. Validation ensures that a model is a good model of the target (Gilbert and Troitzsch, 2005; Gilbert, 2008). Validation is done by comparing the output of the simulation with real-world data on the target behaviour. One of the limitations of the current approaches to simulating conversational groups is that they have not been validated (see chapter 3). However, the agent-based models described in this chapter will be validated (in subsequent chapters) using a systematic approach. By so doing, we aim to provide a clearer understanding of how the different models influence the emergence of conversational group arrangements.

4.5.2 Other advantages of our simulations

While reviewing the literature on automatic F-formation detection tools (see chapter 3), and based on emails exchanged with the researchers developing these tools, it became evident that there is a need for simulated datasets of conversational groups. The meta-data set resulting from the Party World simulations, containing images of simulated conversational groups and data on the position and orientation of interactants, may be a useful resource in this regard.

4.5.3 Limitation of our Party World models

Our exclusive focus on agents' position and orientation, and the blob-shaped appearance used to represent agents, may be a paradoxical limitation of the Party World simulations. Kendon (1990) notes that an individual's transactional segment is governed by the position and orientation of the lower body. However, there may be other factors that indirectly influence the position and orientation of the lower body, and consequently the manifestation of transactional segments. For example, if the head and upper body are positioned outside the transactional segment for a significant period of time, interactants would normally readjust their position and orientation to diminish the misalignment (Kendon, 1990). Furthermore, the use of eye gaze and gesture could also affect interactants' position and orientation. Since our research objective is to examine the outcomes of specific models of positional-orientational behaviour, we have not considered the influence of other forms of bodily interaction in controlling the position and orientation of agents. Nevertheless, in chapter 10, we describe a semi-immersive virtual reality Party World system with non-player humanoid characters. In this system, the appearance of agents allows us to analyse the influence of other forms of bodily communication on the positional-orientational behaviour of agents, and subsequently, the spatial arrangement of agent clusters.

We have also not implemented any adaptive or learning behaviour for agents. Hence agents cannot improve their readjustment behaviour as the simulation progresses or consciously engage in a face-to-face encounter with the same set of interaction partners. But our objective is to start by identifying a simple yet robust model to simulate conversational groups, and then extend the model to include more advanced behaviour, either in the two-dimensional or in the semi-immersive Party World. The work in this thesis thus lays the groundwork for future research in a number of directions.

Part III

Results: Evaluation and Validation of Party World Models

Chapter 5

Analysis of the Emergence of F-formations

In this part of the thesis, we will evaluate and validate the Party World models, by analysing the spatial, temporal and social characteristics of agent clusters produced by the different models. There are five chapters in this part; the first two are concerned with analysing and comparing how different the results are between the different Party World models, while chapters 7 and 8 are concerned with evaluating if the outcomes of the Party World models correspond to reality, i.e. real-world conversational groups. In chapter 9, we present findings from a user evaluation study, conducted before developing the Party World models.

The purpose of this chapter is to confirm if the Party World models generate F-formations, and to analyse and compare the characteristics of the F-formations resulting from the different models. We begin the chapter by describing how we consider F-formations as the unit of analysis. We also describe the data collected from the Party World simulations and the methods used for detecting F-formations in the data. We then recall our research questions, and present our analysis of the F-formations resulting from the different models in sections 5.2 and 5.3, respectively. In section 5.4, we perform a qualitative analysis of changes in F-formations, in terms of participants and spatial arrangements. In section 5.5, we consider scenarios where the Party World environment has obstacles in it, and how that affects the emergence of F-formations. We conclude the chapter with a summary of findings.

5.1 F-formations as the Unit of Analysis

As noted before, an F-formation denotes the spatial arrangement of a conversational group, delineated by the positional-orientational arrangement of members in the group (Kendon, 1990). We use F-formations as the unit of analysis in this chapter – we determine if the spatial patterns of agent clusters resulting from the different Party World models qualify as F-formations. We explore the differences in the size and the spatial distribution of F-formations resulting from the different models. We define size as the number of agents per F-formation, and spatial distribution refers to the spread of coexisting F-formations in the Party World environment. We also examine if, and how, the social, spatial and temporal characteristics of F-formations differ between models.

5.1.1 Data from the Party World simulations

We ran each Party World model for 30 minutes, and recorded screenshots of the simulation window and the position and orientation of stationary agents, once every five seconds. Figure 5.1 is a screenshot from the simulation of Model 2a LA. The agents with a (trapezoidal) transactional segment in figure 5.1 are the ones which were stationary at the time the screenshot was captured.

We recorded 360 images for one run of each model. We ran each Party World model five times, which resulted in a dataset consisting of 1800 images for each model, along with the position and orientation data of stationary agents.

We fed the position and orientation data to the F-formation detection models described in the following section. The models were used to automatically detect the F-formations in the screenshots. We recorded the position and orientation of only the stationary agents to avoid the possibility of models detecting F-formations among agents that were moving at the time the screenshots were made. We used the built-in functions of the Processing programming language to save the screenshots and the positional-orientational data of agents.

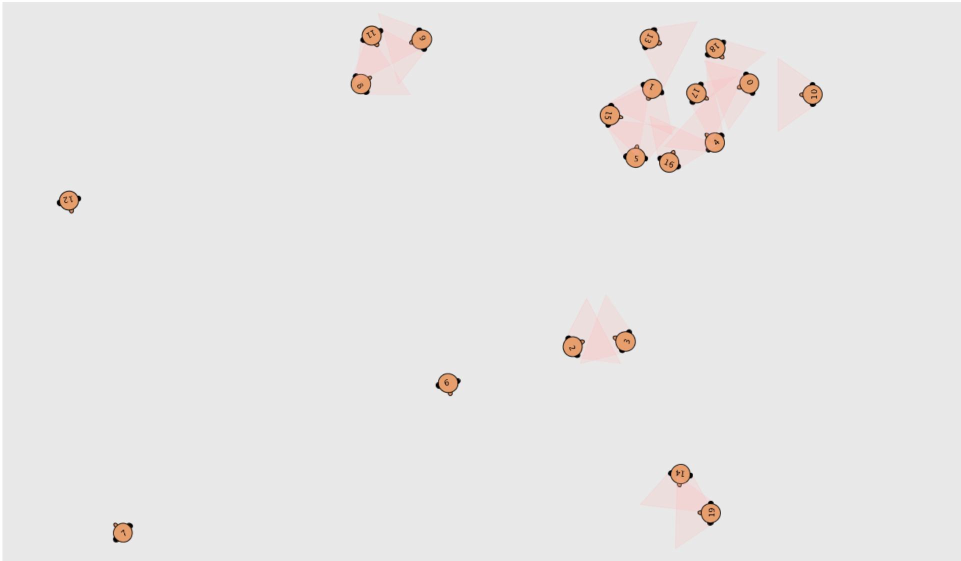


Figure 5.1: A screenshot from the simulation of Model 2a LA

5.1.2 Automatic detection of F-formations

We used three different methods to automatically detect F-formations in the screenshots saved from the Party World simulations. The Hough Voting for F-formation (HVFF) detection and the Graph-Cuts for F-formation (GCFF) detection, whose working was described in section 3.8, were used. The HVFF and GCFF methods take the positions and orientations of individuals in an image, and detect the F-formations in the image, along with identifying the members in each F-formation. Based on email communications with the developers of HVFF and GCFF[1], we were able to adapt these methods to work with the data from our Party World simulations. The adapted HVFF and GCFF methods were implemented in MATLAB[2]. Figures 5.2 and 5.3 are the outcomes generated by the HVFF and GCFF methods by taking as input the positions and orientations of the stationary agents in figure 5.1.

In figures 5.2 and 5.3 agents appear to be closer than they actually are in figure 5.1 because of the image processing techniques used in the F-formation detection methods. Among the two, HVFF took the longest processing time to detect F-

[1]The investigator is thankful to the developers of HVFF and GCFF, particularly Marco Cristani and Francesco Setti, for their help and guidance in adapting the HVFF and GCFF methods for our investigation.

[2]http://uk.mathworks.com/

formations. We ran the model on an iMac PC (Intel Core 2 Duo processor with speed 2.66 GHz and 4GB RAM) for 300 hours to process the 1800 images from each Party World model. We ran GCFF on a laptop (Intel Core i7 processor with speed 2.8 GHz and 16GB RAM) for about 7 hours to process the 1800 images from each Party World model.

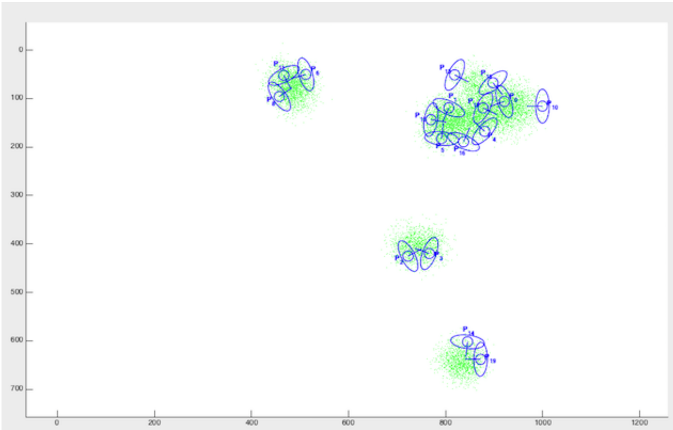


Figure 5.2: F-formations in figure 5.1 detected by HVFF

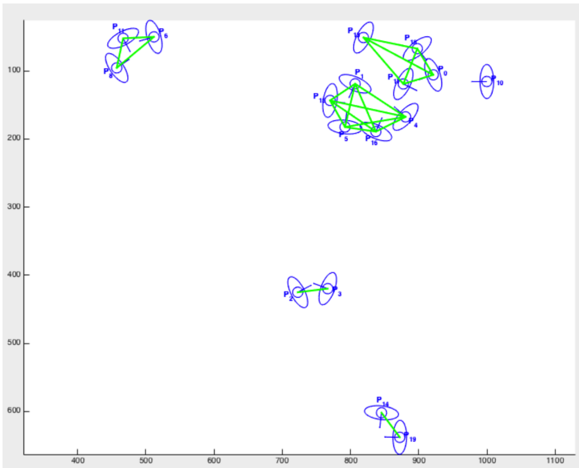


Figure 5.3: F-formations in figure 5.1 detected by GCFF

In addition to using the HVFF and GCFF methods, we implemented our own version of an F-formation detection algorithm. As described in chapter 3, the HVFF and GCFF methods use recursive voting strategies to compute the o-space centres, and its affiliate members. Our Party World F-formation Detection (PWFD) method does not use any sophisticated recursive voting strategies like HVFF and GCFF. Our PWFD method identifies the agent clusters in an image, where the transactional segment of every agent within a cluster overlaps at least with one other agent in the cluster, as F-formations. The difference between the methods is that, HVFF and GCFF combine rules concerning interpersonal distances, individual transactional segments, and unobstructed o-spaces to identify F-formations, whereas PWFD considers the overlap of transactional segments as the essential condition to identify F-formations.

The working of our PWFD method is shown in figure 5.4. Unlike the HVFF and GCFF methods, PWFD runs and detects the F-formations while the models are running, hence there is no requirement for post-processing the data. We implemented the PWFD method for three reasons: (1) HVFF and GCFF have long run times, (2) PWFD detects F-formations in real-time requiring no post-processing, and (3) the PWFD method can be used in extending the Party World models such that agents can identify and react to ongoing F-formations.

Figure 5.5 shows the F-formations detected by PWFD in figure 5.1. The F-formations and their members detected by the different models are summarised in table 5.1. The table also shows the slight differences in the performance of the different models. For example, agent 4 is in the middle of two groups in figure 5.1. While HVFF and GCFF detected agent 4 as a member of the F-formation comprised of agents {1,5,15,16}, PWFD detected agent 4 as belonging to the F-formation comprised of agents {0,17,18}. Also, GCFF and PWFD detected agents 0 and 17 as members of an F-formation comprised of two other agents, whereas HVFF identified 0 and 17 as a dyadic F-formation.

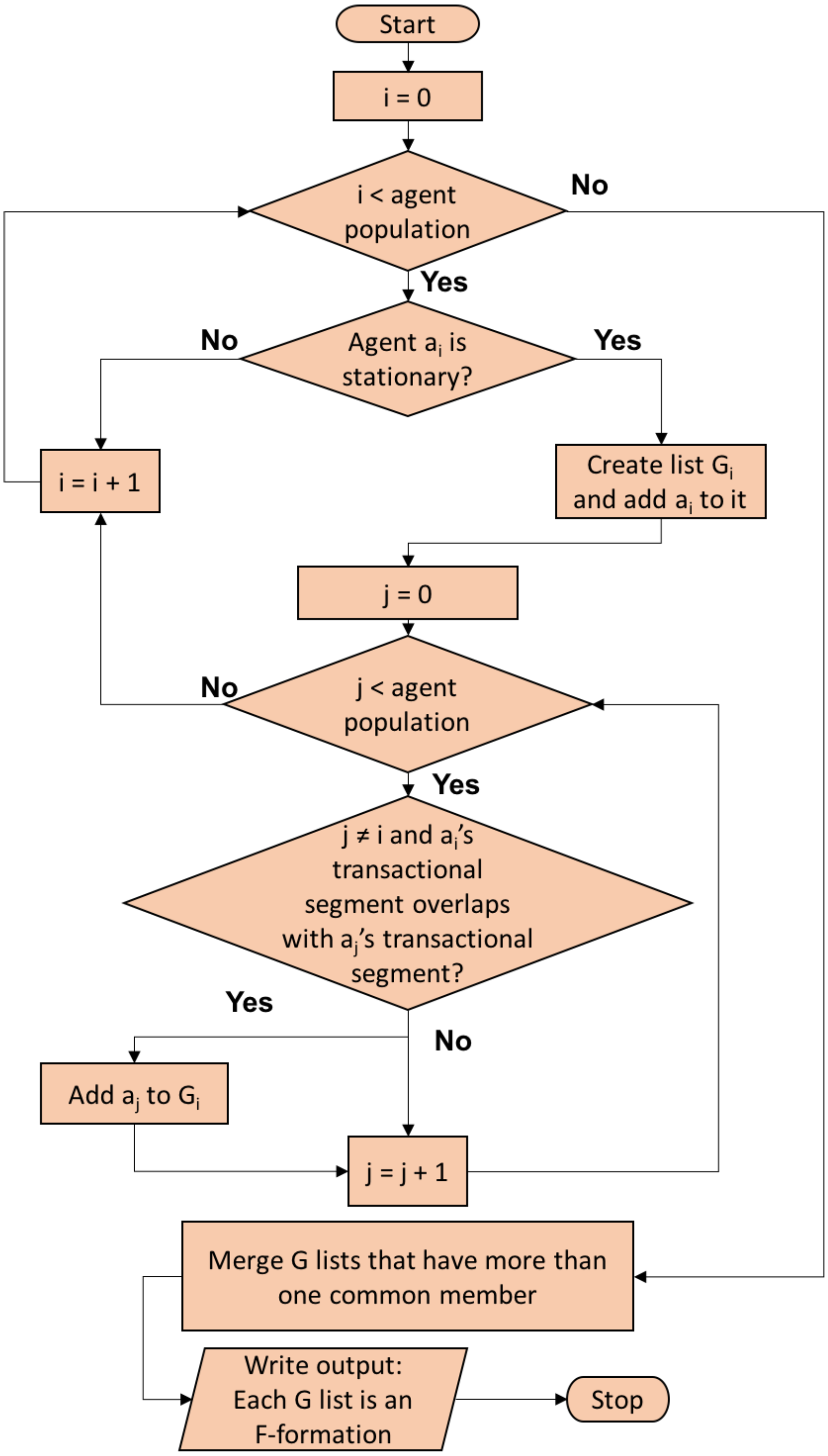


Figure 5.4: The Party World F-formation Detection (PWFD) method

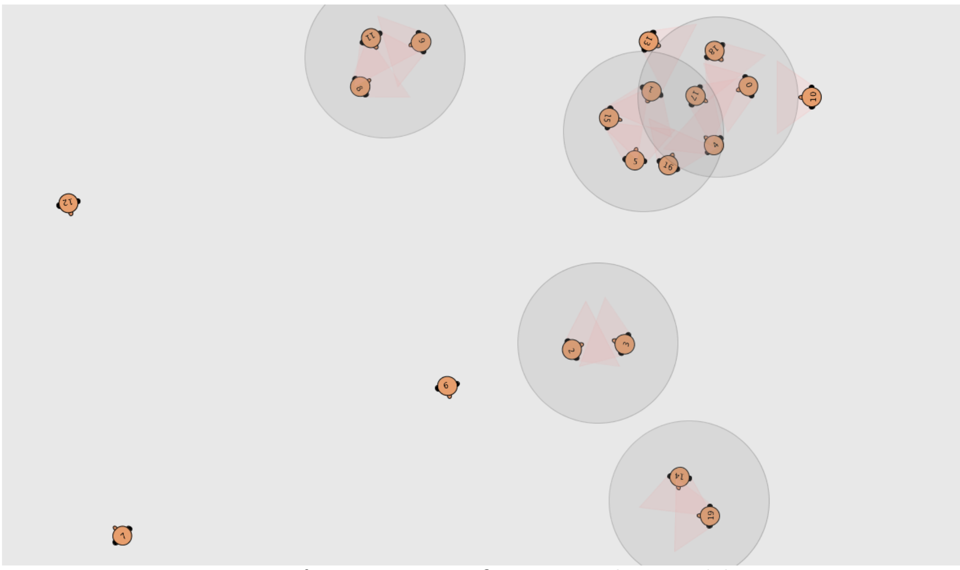


Figure 5.5: F-formations in figure 5.1 detected by PWFD

HVFF		GCFF		PWFD	
F-formation	Members	F-formation	Members	F-formation	Members
1	{1,4,5,15,16}	1	{1,4,5,15,16}	1	{1,5,15,16}
2	{6,8,11}	2	{6,8,11}	2	{6,8,11}
3	{2,3}	3	{2,3}	3	{2,3}
4	{10,18}	4	{0,13,17,18}	4	{0,4,17,18}
5	{14,19}	5	{14,19}	5	{14,19}
6	{0,17}	-	-	-	-

Table 5.1: F-formations in figure 5.1 detected by HVFF, GCFF and PWFD

5.2 Research Questions

As noted in chapter 1, we aim to analyse the outcomes of the Party World simulations to answer the following research questions:

1. What are the differences between the Party World models in terms of the F-formations they generate?

2. What are the differences in the spatial patterns and durations of agent clusters resulting from the different Party World models?

3. Which Party World model produces agent clusters whose spatial patterns most closely resemble to those of human conversational groups?

We address the first question in this chapter, while the second and third questions are addressed in the subsequent chapters. The first question seeks to determine if the Party World models generate F-formations, and if yes, to identify the differences between the models in terms of the characteristics of the F-formations they produce. We perform the analysis in four stages: (1) to examine if the spatial patterns of agent clusters resulting from the different Party World models qualify as F-formations, (2) to explore the characteristics of the simulated F-formations, (3) to examine the reshaping of F-formations due to changes in group membership, and (4) to examine the emergence and spread of F-formations in a non-empty Party World environment. The following sections describe the details of each stage, including the hypothesis tested at each stage.

5.3 Analysis of F-formations Generated by the Party World Models

Hypothesis 1: The estimates of the number of F-formations resulting from the Party World simulations will be different for the different models.

Kendon (1990) suggests that an F-formation arises whenever two or more individuals organise their positional-orientational behaviour with respect to one another such that it leads to the emergence of an o-space. The total number of agents in the system was kept constant (at 20) for all the models. However, because agents use different readjustment strategies in the centroid-based and field-of-view approaches, and no strategy in Model 1, we hypothesised that the estimates of the number of F-formations resulting from the Party World simulations will be different for the different models. We expected that the field-of-view models and the centroid-based models would produce more F-formations than the baseline model, which does not have rules to control the orientation behaviour of stationary agents.

5.3.1 Valid F-formations

Figure 5.6 shows the median number of F-formations detected over five runs of the different Party World models. Model 2b LA, i.e. centroid-based approach in a local neighbourhood, generates the most F-formations (HVFF=2365, GCFF=2440, PWFD=2267). Model 2a EA, i.e. field-of-view approach in an extended neighbourhood, generates the least F-formations (HVFF=1847, GCFF=1907, PWFD=1372). Modeb 2b EA, i.e. centroid-based approach in an extended neighbourhood, also generates fewer F-formations (HVFF=1917, GCFF=1901, PWFD=1532). Model 1 and Model 2a LA perform in the middle range, (HVFF=2060, GCFF=2063, PWFD=1844) and (HVFF=1980, GCFF=1999, PWFD=1855), respectively. We performed further tests to determine the statistically significant variations between models.

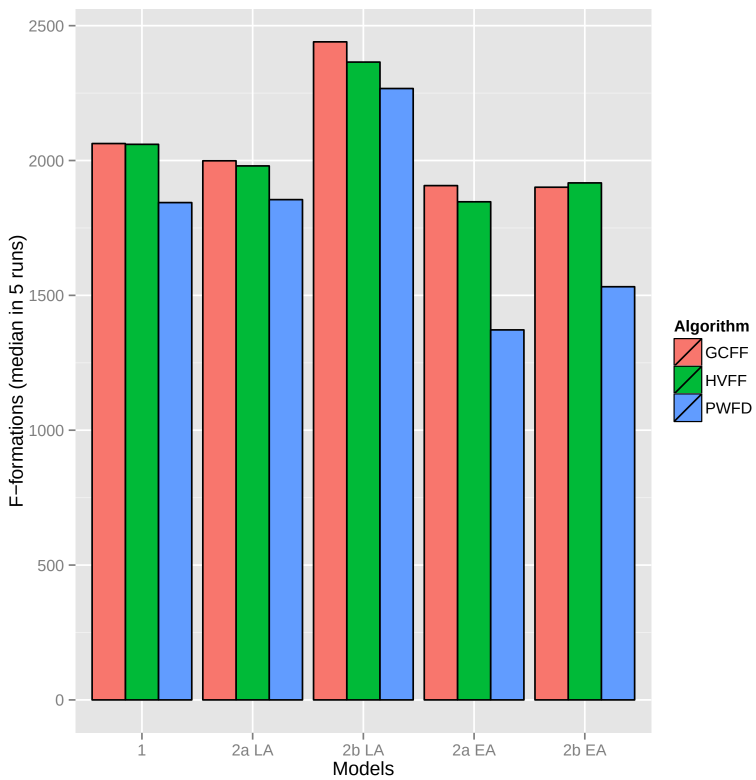


Figure 5.6: Median number of F-formations detected by HVFF, GCFF & PWFD in the Party World dataset

5.3.1.1 Variation between runs in estimates of F-formations

There are three random factors in the Party World models: (1) agents start moving from random initial positions and orientations in each run, (2) agents can randomly change their direction of motion, (3) the frequency at which agents, in the *stopped* state, readjust their position and orientation differs between agents. These factors could cause variations in agents' behaviour across model runs, and hence affect the emergence of F-formations. Figure 5.7 shows the variation in the F-formations detected over five model runs. The extended neighbourhood influence in Model 2a EA and Model 2b EA result in high variation between runs. Model 2a LA shows less variation with an outlier at 2076 (HVFF) and 2089 (GCFF) respectively. Model 1 also shows less variation, while Model 2b LA performs in the middle range.

The trend suggests that models where agents readjust their position and orientation based on an extended neighbourhood, show high variation between runs. On the other hand, models where agents are only influenced by their immediate neighbours show less variation between runs in terms of generating F-formations. The finding is also supported by the coefficient of variations in the number of F-formations detected over five runs, summarised in table A.1 in the appendix chapter A.

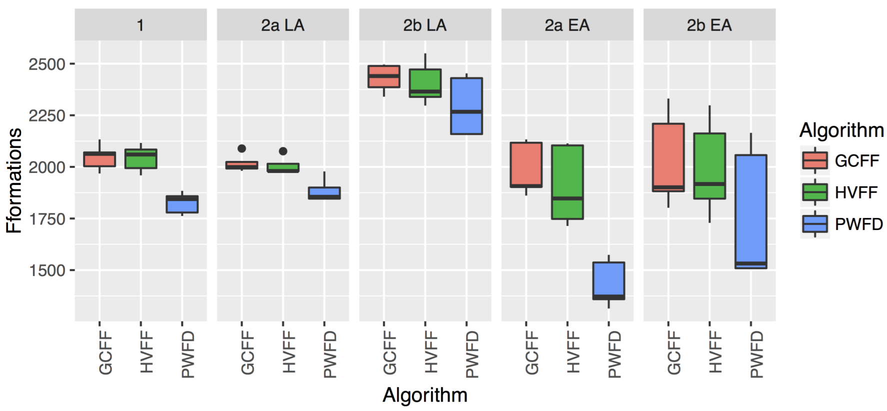


Figure 5.7: Variation in F-formation measures over five runs

5.3.1.2 Variation across models in estimates of F-formation

As the data used for analysis was collected over five runs of each model, we had a balanced design, but with a small sample size (N = 5). We used Shapiro Wilk's test for normality (see table A.2 in appendix chapter A) and found the samples (measures of the total number of F-formations detected in each run) were mostly drawn from a normal distribution (i.e. significance value > 0.05). We could have used the Kruskal Wallis non-parametric test to assess the significance of variation in the F-formations detected across models, but heteroscedasticity was an issue, i.e. there was a large variation in the standard deviation across models. Hence, with a balanced design of a small sample size, where samples of the measurement variable were shown to be drawn from a normal distribution, we performed a one-way ANOVA instead, with Welch's correction applied to address heteroscedasticity (McDonald, 2009). The results of the analysis are reported in table 5.2 suggesting that there are significant variations in the total number of F-formations detected across the different Party World models.

HVFF		GCFF		PWFD	
$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value
13.803	$p < 0.001^a$	30.471	$p < 0.001^b$	24.017	$p < 0.001^c$

[a] 5.984e-04; [b]1.947e-05; [c]5.68e-05

Table 5.2: Summary of one-way ANOVA to assess variations in F-formations across models

To further analyse the variation, we performed the Tukey Kramer test. Results are shown in figures 5.8, 5.9a and 5.9b, and the measures of significance among pairs of means are summarised in the appendices chapter in tables A.3, A.4, and A.5. Results of the Tukey Kramer test on F-formations detected by HVFF and GCFF show that Model 2b LA produces a significantly higher number of F-formations than all the other models. Results of the post hoc analysis on the F-formations detected by PWFD also shows that Model 2b LA varies significantly from all the other models, as well as showing that Model 2a EA produces significantly fewer F-formations than Models 1, 2a LA and 2b EA.

The main purpose of the analysis in this section was to confirm whether the Party World models generated F-formations, and to understand the differences between models in terms of the number of F-formations they produce. We found that all the

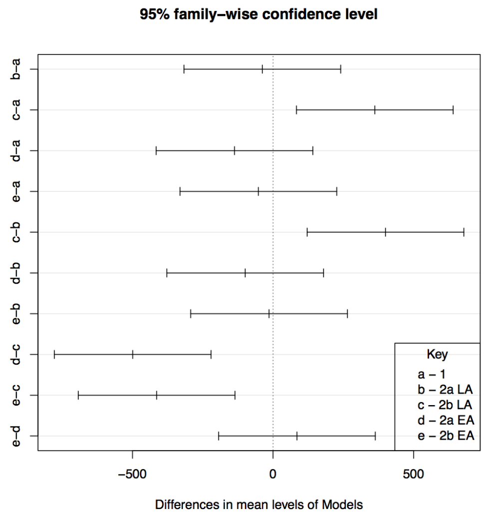
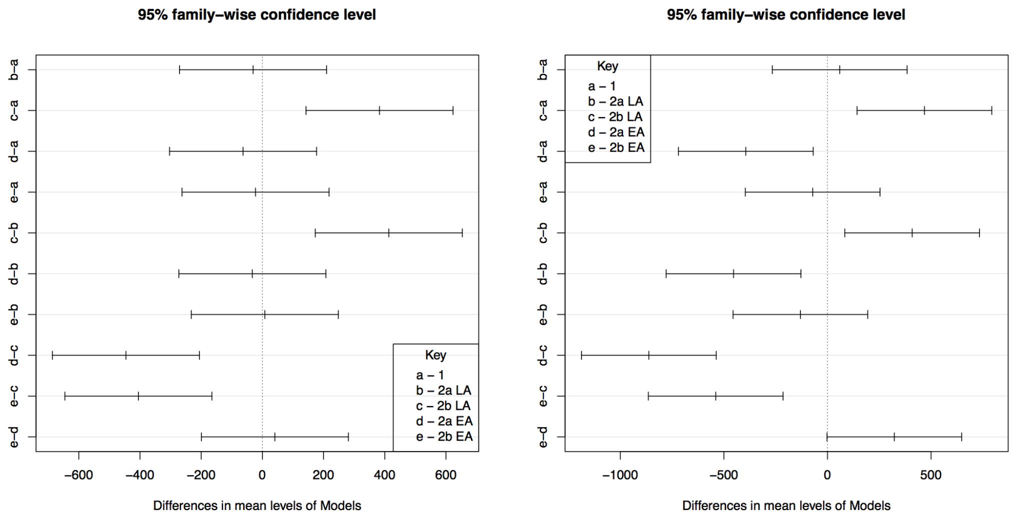


Figure 5.8: Tukey-Kramer test on HVFF outcomes

Party World models generate agent clusters that qualify as F-formations. The results showed that Model 2b LA produces a significantly higher number of F-formations than all the other Party World other models. We expected that the baseline approach would generate fewer F-formations than all the other models, but it was interesting to find that Model 1 performs on a par with Model 2a LA, Model 2a EA and Model 2b EA. In the next section, we analyse the size of F-formations resulting from the Party World models, to better understand the differences between models in terms of the number of F-formations they produce.



(a) GCFF (b) PWFD

Figure 5.10: Tukey Kramer test on GCFF and PWFD outcomes

5.3.2 The size of F-formations

Hypothesis 2 (H2): The size of F-formations resulting from the simulations will vary across the different Party World models. We expected that the field-of-view approach and the centroid-based approach, and likewise the local and extended neighbourhood influences, will impact the size of F-formations resulting from the models.

Existing computational models have not investigated how the rules used in the model influence the size of agent clusters. Salem and Earle (2000) made an assumption of six members per cluster. Riesman et al. (1960) suggest that six members per conversational group is a realistic assumption. Nonetheless, in Salem and Earle (2000), the group size was prefixed and not emergent. Many of the other models reviewed in chapter 3, e.g. Jan and Traum (2007), Karimaghalou et al. (2014) and Pedica and Vilhjálmsson (2012), generate agent groups of four to six members, but there is no consensus on how rules of individual spatial behaviour influence the size of resulting agent clusters. It is also not known how the local and extended neighbourhood influences affect the size of F-formations. Hence, we proposed hypothesis H2 to analyse if, and how, the Party World models differ in influencing the size of F-formations.

The data used for analysis came from the HVFF, GCFF and PWFD methods, which in addition to identifying F-formations, also identify the members of the F-formations. We used the membership information to derive the size of F-formations. Figure 5.11a shows the median number (over the same 5 runs as the previous analysis) of F-formations according to their size, detected by HVFF, across the different Party World models. Figures 5.11b and 5.11c show the median number of F-formations and their size, detected by GCFF and PWFD respectively, across the different models.

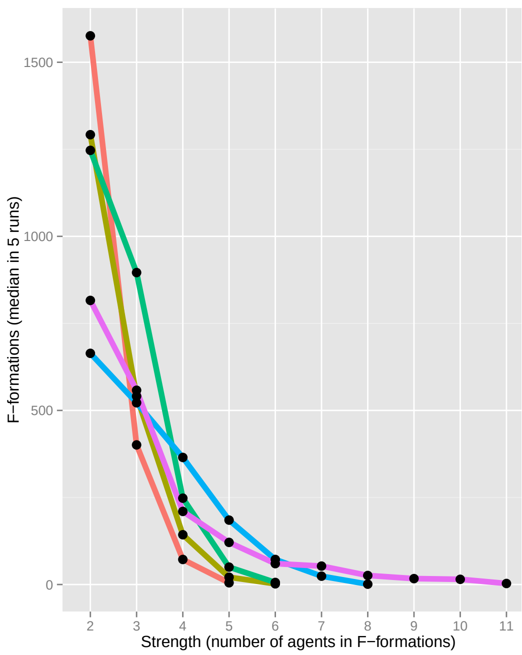
The overall trend seems to be that the models are all capable of generating F-formations of 2 to 5 agents, but F-formations of 6 or more agents are uncommon in some models. Model 1, Model 2a LA and Model 2b LA occasionally generate F-formations of 6 or 7 agents, but groups of 8 or more members are not generated by these models. Model 2a EA generates F-formations of two to nine members, while Model 2b EA generates F-formations of 11 to 12 members, and occasionally, even up to 13 or 14 members.

Since F-formations of two to five members were generated by all the five Party World models, we conducted further analysis to compare the frequencies of occurrence of dyadic, triadic, four-member and five-member F-formations, across the different models. Once again, because the samples were shown to be drawn from a normal distribution (Shapiro-Wilk's test) but with non-homogeneous standard deviations, we conducted a one-way ANOVA for comparing the variations across models for each size. Table 5.3 shows the results of the one-way ANOVA, suggesting there are significant differences between the Party World models in terms of generating F-formations of different sizes.

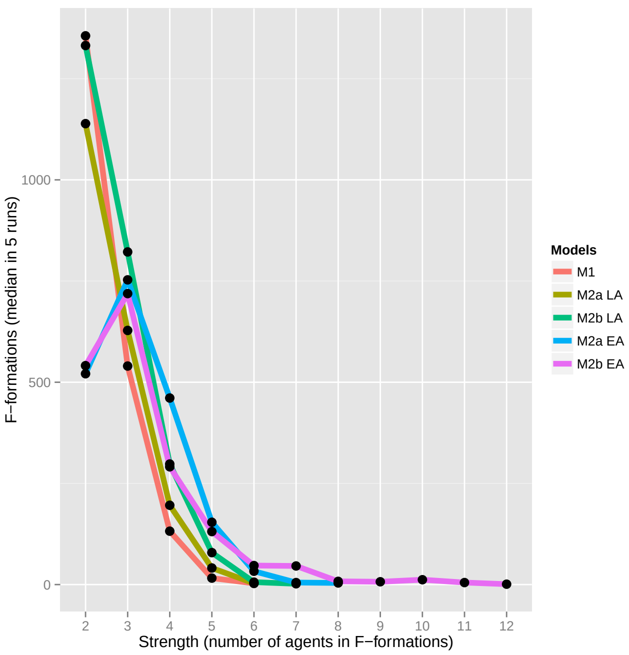
Size	**HVFF**		**GCFF**		**PWFD**	
	$\mathbf{F_{4,20}}$	**p value**	$\mathbf{F_{4,20}}$	**p value**	$\mathbf{F_{4,20}}$	**p value**
Two	33.6328	$p < 0.001$ (1.392e-05)	44.0916	$p < 0.001$ (6.312e-06)	53.3752	$p < 0.001$ (1.987e-06)
Three	33.534	$p < 0.001$ (1.894e-05)	19.3223	$p < 0.001$ (1.259e-04)	8.5435	$p < 0.05$ (3.474e-03)
Four	287.9257	$p < 0.001$ (8.077e-10)	46.9553	$p < 0.001$ (3.436e-06)	6.2032	$p < 0.05$ (9.877e-03)
Five	20.8587	$p < 0.001$ (1.787e-04)	32.811	$p < 0.001$ (4.187e-05)	44.0192	$p < 0.001$ (5.376e-06)

Table 5.3: Summary of one-way ANOVA to assess variations in the size of F-formations

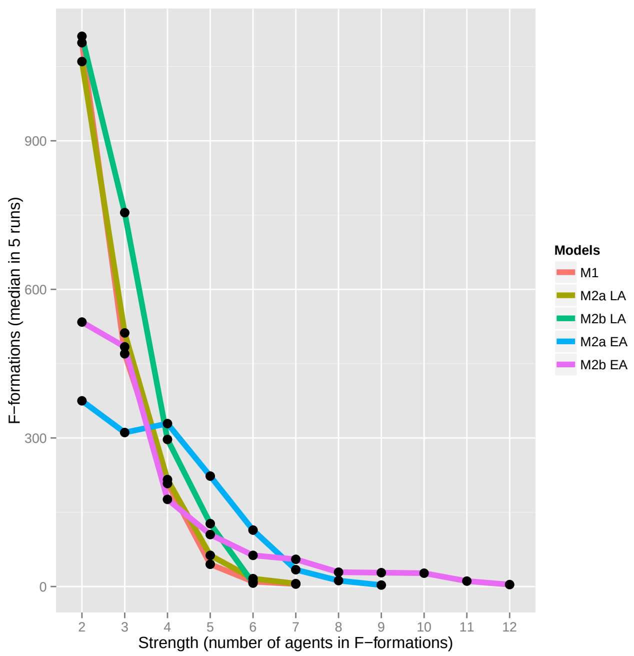
We performed the Tukey Kramer test to further examine how the size of F-formations varied across models. Results showed that the extended neighbourhood influence in Model 2a EA and Model 2b EA resulted in significantly fewer dyadic F-formations than the other Party World models. Results also showed that Model 1 produces significantly fewer triadic F-formations than the centroid-based models and Model 2a EA. Furthermore, Model 2b LA produces significantly more triadic F-formations than the other Party World models. The results of the Tukey Kramer test on the outcomes of PWFD showed a slightly different pattern of variation across models in terms of generating triadic F-formations – Model 1 produces significantly fewer triadic F-formations than Model 2b LA, while Model 2a EA produces sig-



(a) HVFF outcomes



(b) GCFF outcomes



(c) PWFD outcomes

Figure 5.11: Size of F-formations

nificantly fewer triadic F-formations than Models 2a LA, 2b LA and 2b EA. The results of the Tukey Kramer test for dyadic and triadic F-formations are also shown in figure A.1 in appendices chapter A.

Results of the Tukey Kramer test on the outcomes of HVFF and GCFF methods showed that Model 2a EA produces significantly more four member F-formations than all the other Party World models. On the other hand, Model 2b LA and Model 2b EA, are statistically similar in terms of generating four member F-formations. Analysis of the GCFF data also showed similarities between Model 1 and Model 2a LA, and between Model 2a LA and Model 2b LA, in terms of generating F-formations of four members. Analysis of the PWFD data showed that Model 1 and Model 2a EA vary significantly from Models 2a LA, 2b LA and 2b EA.

Lastly, the results of the Tukey Kramer test on the outcomes of the HVFF and GCFF methods showed that Model 2a EA and Model 2b EA produce significantly more five member F-formations than all the other models. Analysis of the PWFD data showed that Model 2a EA produces significantly more five member F-formations than all the other models. These results are also shown in figure A.1 in appendices chapter A.

Drawing together these findings, it is apparent that there are variations among the different Party World models, in terms of the size of the F-formations they produce. These results support hypothesis 2: the size of F-formations does vary across the different Party World models. The trend observed is that F-formations of two to five agents are generated by all the five Party World models, although there are variations among models in terms of how many of these arrangements each one produces. Model 2a EA and Model 2b EA generate bigger F-formations – the former generates F-formations of up to 9 members and the latter generates F-formations of even up to 14 members. Model 1, Model 2a LA and Model 2b LA occasionally generate F-formations of six or seven members.

In summary, we found that models using an extended neighbourhood influence produce bigger F-formations and fewer dyadic F-formations. On the other hand, models using a local neighbourhood influence produce fewer F-formations that have more than five or six members.

5.3.3 The number of F-formations per image

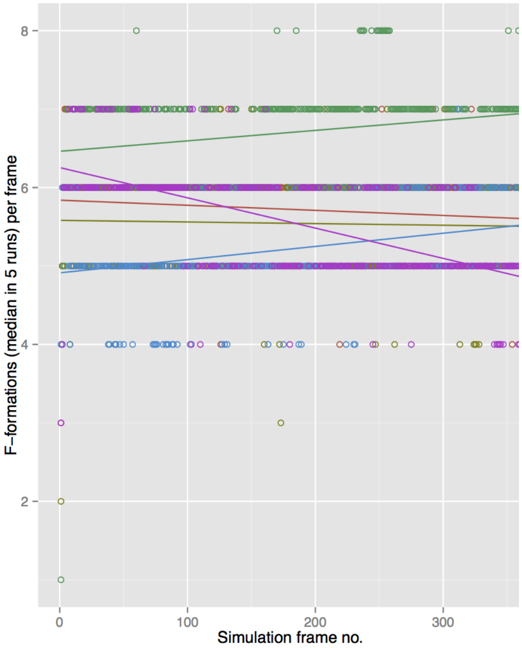
We found that the average number of F-formations per image varies over time across the different Party World models. Plotting the median number (over five runs) of F-formations detected per image by the HVFF method shows that for Model 2b LA and Model 2a EA the number of F-formations per image increases over time, whereas for Model 2b EA there is an evident fall in F-formations (see figure 5.12a). There is a marginal decrease in F-formations for Model 1 and Model 2a LA. Except for minor differences, a very similar trend is seen in the case of GCFF and PWFD assessments, too (see figures 5.12b and 5.12c).

The marked decrease in the number of F-formations per image over time for Model 2b EA is consistent with our previous finding that this model generates F-formations of 11 to 12 members, or even up to 14 members. That is to say, more agents per group means fewer F-formations per image, because the total number of agents in the simulation remained fixed. The marked increase in the number of F-formations per image over time for Model 2b LA is also consistent with our previous finding that this model generates the most F-formations, wherein the size of groups varied between 2 to 7 members. In other words, having fewer agents per group means it is possible to have more F-formations per image.

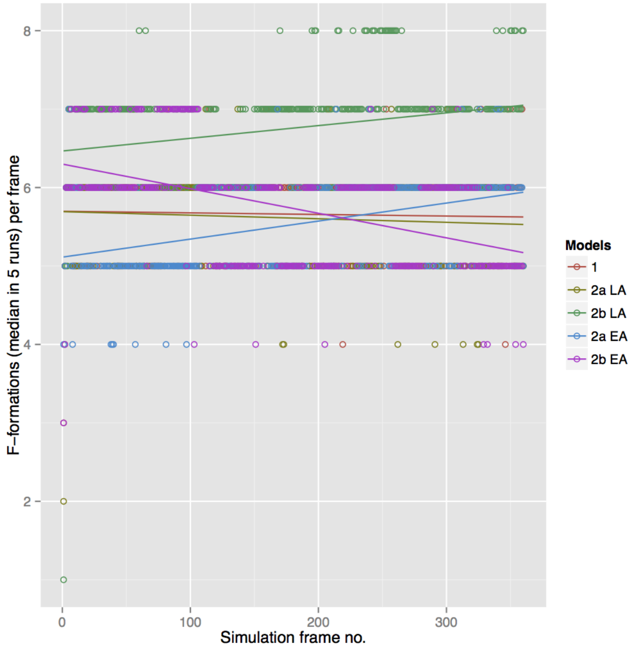
5.4 Analysis of Changes in F-formations

Hypothesis 3: The dynamics of changes in the social, spatial and temporal characteristics of agent clusters will vary across the different Party World models. We expected that the characteristics of agent clusters in Model 1 will be less realistic than the other models. We also expected a difference between the field-of-view and centroid-based models, and likewise, the local and extended neighbourhood influences.

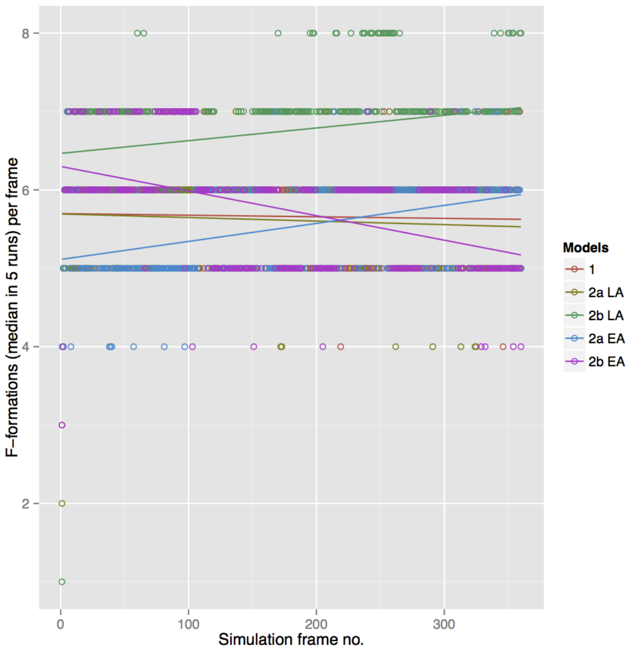
The process of establishing, maintaining and ending F-formations is a dynamic one, where the spatial arrangements of groups are likely to change when new mem-



(a) HVFF outcomes



(b) GCFF outcomes



(c) PWFD outcomes

Figure 5.12: Number of F-formations per image over time

bers join and existing members leave. To examine how the social, spatial and temporal characteristics of agent clusters produced by the different Party World models change over time, we perform a timeline-based analysis similar to the ones done by Deutsch (1977) and Kendon (1990) (cf. see figure 4.5a). Since the Party World models use different strategies to control the positional-orientational readjustments of agents, we hypothesised that the dynamics of changes in agent clusters will vary across the different Party World models. The segments used for analysis in this section were selected by visual inspection of the images in the Party World dataset. The segments illustrate the characteristics of agent clusters unique to each model.

5.4.1 A timeline analysis of Model 1

Figure 5.13 shows the changes in the spatial pattern of an agent cluster resulting from Model 1. What started out as a dyadic F-formation at 21:08:29 (detected by HVFF, GCFF and PWFD models) transformed to two F-formations as of 21:10:19 – the GCFF and PWFD models detected two F-formations, one comprised of agents {0,9} and the other comprised of agents {8,14,17}, while HVFF detected one F-formation comprised of agents {0,9} and another comprised of agents {8,17}.

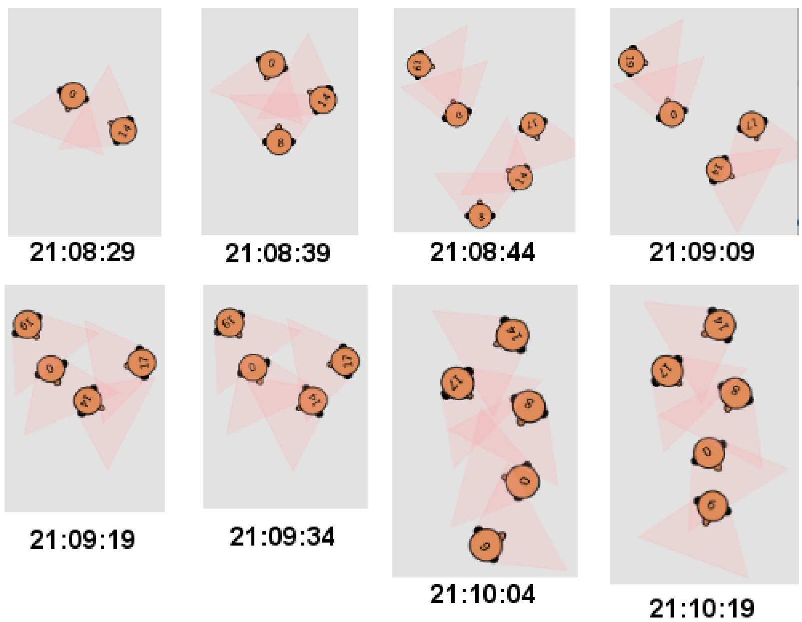


Figure 5.13: Timeline of an agent cluster produced by Model 1

Figure 5.13 also shows the changes in the spatial arrangement of the cluster corresponding to changes in participants. Agents 0 and 14 are in an L-shaped arrangement at first (21:08:29). Ten seconds later, agent 8 has joined the group and the trio are in a plausible face-to-face arrangement, and their individual transactional segments are overlapping. At 21:08:44, agents 8 and 14 have not changed their position or orientation, but agent 0 has left the group and is seen interacting with agent 19. On the other hand, agent 17 has joined agents 8 and 14, and the trio are in a plausible F-formation. Twenty five seconds later, agents 0 and 19 have not altered their position or orientation, while in the other group, agent 14 has changed its position and orientation and agent 8 has left the group.

In Model 1 agents do not readjust their position or orientation in the stationary state. However, because agents move in the forward direction, occasionally they do end up facing one other when they stop moving. Due to this reason, the spatial patterns of agent clusters, at least the ones we have considered so far, appear to be plausible F-formations. This also explains why a considerable number of F-formations were detected in the Model 1 dataset (cf. section 5.3). Nonetheless, Model 1 does not always produces plausible F-formations. For example, in figure 5.13, it can be seen that the agents are not always facing one another from 21:09:19 onwards. Some agents are in a queue-like arrangement, e.g. agents 0, 14 and 19 at 21:09:19.

At 21:09:34, agents 0 and 14 are facing one another, whereas agents 17 and 19 are behind agents 14 and 0, respectively. The spatial arrangement and the membership of the group changes quite a bit in the next 30 seconds. At 21:10:04, agents 0, 14 and 17 are still in an F-formation, although the position and the orientation of the agents has changed. Furthermore, as of 21:10:04, agent 19 has left the group, whereas agents 8 and 9 have newly joined. At this point, agents 0, 8, 9 and 17 appear to be facing one another, while agent 14 seems to be in a bystander position.

At 21:10:19, agents 8, 14 and 17 have not changed their position or orientation, but agents 0 and 9 are no longer facing one another or the group. This happened because the separation distance condition was violated almost as soon as agents 0 and 9 tried to disengage from the group and start moving in a different direction. Model 1 does not allow agents to readjust their position or orientation, and so, despite the physical proximity, agents 0 and 9 are facing away from one another and the group.

Just as the example shown in figure 5.13, we found that queue-like patterns, and arrangements where agents are facing away from one another, are common in Model 1. A general trend observed is that the spatial arrangement of the first two or three agents forming a group is plausible in Model 1. But as more members join the group, or when existing members alter their position or orientation, the arrangements become less plausible.

5.4.2 A timeline analysis of Model 2a LA

Figure 5.14 shows the timeline of an agent cluster generated by Model 2a LA. The cluster at 15:57:31 shows agents 1, 6 and 10 facing one another in a plausible arrangement. Ten seconds later, agent 11 has joined the group and agent 10 is facing it almost directly, while agents 1 and 6 continue facing agent 10. It can also be seen that agent 11 is in a slightly open position, facing both agents 1 and 10. The arrangement of agents 1, 6, 10 and 11 at 15:57:41 conveys an impression that, agent 10 is the speaker, agent 11 is the primary addressee, while agents 1 and 6 are listeners.

In Model 2a LA agents are capable of perceiving and reacting to other agents in their immediate neighbourhood. Agents readjust their position and orientation, in order to perceive as many immediate neighbours as possible within a 180° field of view. Agent 10 is the immediate neighbour of agents 1, 6 and 11, and so, they are all directly facing it. On the other hand, agent 11 doesn't readjust its position or orientation with respect to agent 1, who is not its immediate neighbour.

At 15:57:56, agent 3 joins the group, and twenty seconds later, agent 13 has also joined the group. At this point, each agent in the group readjusts its position and orientation with respect to its immediate neighbours. The resulting spatial configuration at 15:58:16 is not a proper circle, and yet, the agents are all facing one another.

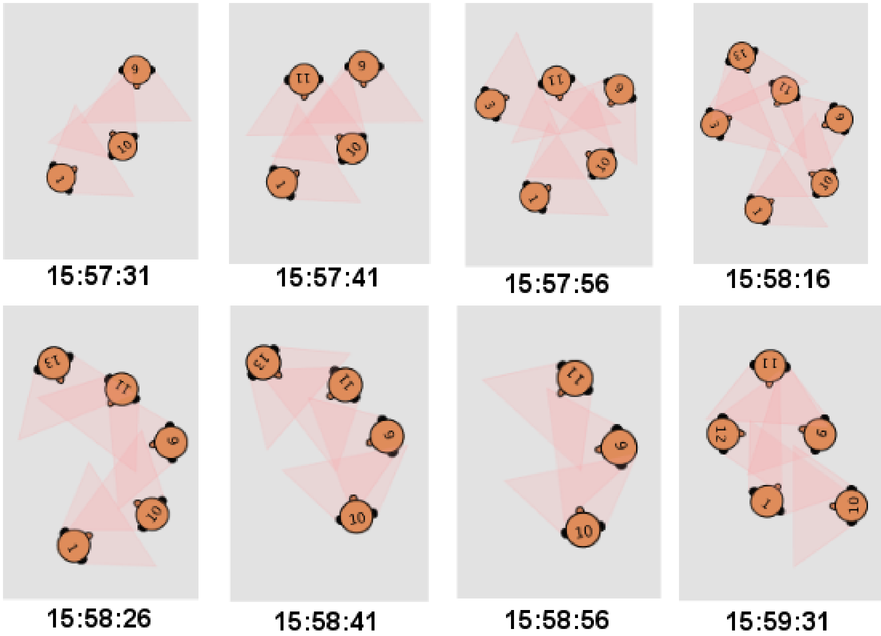


Figure 5.14: Timeline of an agent cluster produced by Model 2a LA

At 15:58:26, the cluster assumes a semi-circular arrangement, which continues at 15:58:41 and 15:58:56. During this time, agents 1, 3 and 13 have left the group, and the six member cluster has become a triadic cluster as of 15:58:56. Sometime later, agents 1 and 12 join the group, and the spatial pattern of the group has changed from a semi-circular open arrangement (at 15:58:56) to a closed ellipse-like arrangement (at 15:59:31).

Similar to the example described here, a general trend observed in Model 2a LA is that the spatial organisation of an agent cluster keeps changing as new members join and existing members leave. But more importantly, unlike Model 1, the spatial arrangement of agent clusters remain realistic, even when the size of the group grows. This behaviour corresponds with Kendon's (1990) suggestion that an F-formation continues to exist in the face of changes, such as when existing members leave the group, or even if there is a complete turnover of participants.

5.4.3 A timeline analysis of Model 2b LA

This section presents the timeline analysis of an agent cluster generated by Model 2b LA, where agents readjust their position and orientation based on a centroid point, calculated from the (x,y) positions of agents within a local neighbourhood. Each agent calculates a centroid point based on the position of itself and its immediate neighbours. The agent then keeps readjusting its position and orientation relative to its local centroid point. Figure 5.15 shows an agent cluster produced by Model 2b LA. At 18:19:33, agents 14 and 16 are in a vis-à-vis formation, almost a minute later, agent 6 joins the group and the three agents (6, 14 and 16) begin to form a triangular arrangement. The spatial arrangement of the triad continues to evolve through 18:21:03, and eventually, results in the triangular shape seen at 18:22:18.

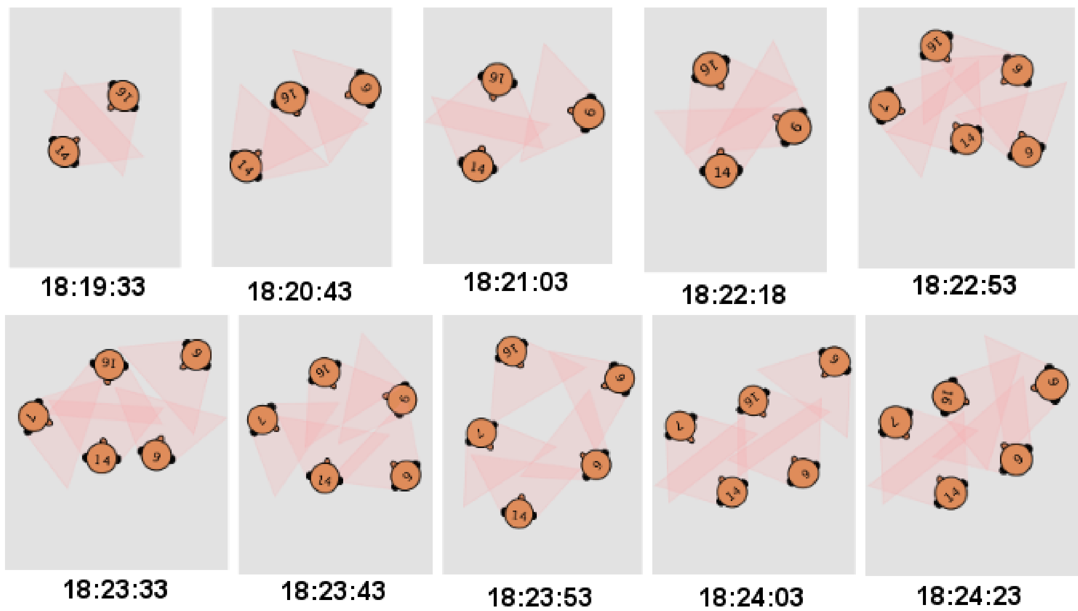


Figure 5.15: Timeline of an agent cluster produced by Model 2b LA

At 18:22:53, agents 7 and 9 join the group, causing all the agents to readjust their positions and orientations. The rearrangement continues until 18:23:53. In between, at 18:23:33, 18:23:43, and 18:23:53, agents are seen moving close to, or away from, their centroid point in attempts to maintain equal distance from the centroid point. The rearrangements eventually cause agents to spread out so much that the group splits into two – at 18:24:03, agents 7 and 14 are in one group, agents 9 and 16 are in another group and agent 6 is just outside. At 18:24:23, agent 6 has readjusted its position and orientation to join agents 9 and 16, as a full-fledged member of the group. The triangular arrangement of agents 6, 9 and 16 is close to the vis-à-vis formation of agents 7 and 14.

The spatial dynamics of agent clusters produced by Model 2b LA, which follow the same trend as this example, can be summarised as follows. Big groups split into smaller groups, causing agents to alternate between being full-fledged members, and bystanders/outsiders of a group. Occasionally, small groups merge and form a big group, but the formation does not hold for long. This is again due to the constant positional-orientational readjustments undertaken by agents, which eventually causes groups to split. The splitting phenomenon observed in Model 2b LA explains the considerably higher numbers of F-formations generated by this model than any other Party World model (see section 5.3.1). Another unique characteristic of Model 2b LA is that agents keep moving towards or away from the centroid point in an oscillatory manner. This type of motion was not found in Model 1 and Model 2a LA.

5.4.4 A timeline analysis of Model 2a EA

The logic underlying Model 2a EA is that agents readjust their position and orientation to perceive as many extended neighbours as possible within a 180° frontal view. Figure 5.16 shows the changes in the spatial pattern of an agent cluster generated by Model 2a EA. At 17:42:02, agents 7, 9, 16 and 19 are in a semi-circular arrangement, where there is a free bonding site (cf. section 2.5 in chapter 2) between agents 7 and 19. The arrangement becomes closed when agent 6 joints the group at 17:43:17. The arrangement remains closed as of 17:43:37 when agent 4 joins the group.

At 17:44:52 the group opens up and starts splitting due to changes in group membership. Old members (agents 4, 6 and 7) have left and new members (agents 5, 8, 11 and 12) have joined the group. Twenty seconds later, there is a big group of 8 agents (5, 6, 8, 9, 11, 12, 16 and 19), in a non-circular but closed arrangement.

Starting from 17:45:12 to 17:46:57, more agents join the group, and as of 17:54:32, all the twenty agents in the system have clustered together. Such big groups are a common outcome in Model 2a EA. This type of gathering behaviour was an unexpected outcome of Model 2a EA.

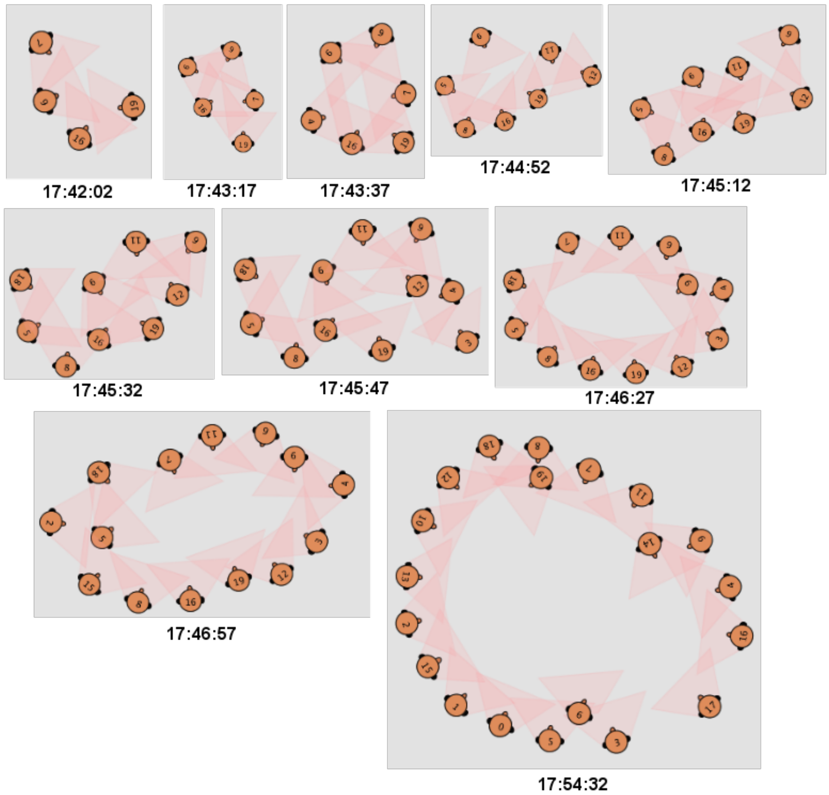


Figure 5.16: Timeline of an agent cluster produced by Model 2a EA

Although it looks like one big cluster at 17:54:32, the HVFF, GCFF and PWFD methods detected several small groups, and not one big F-formation. The GCFF model detected 7 F-formations ({0, 1, 15}, {4, 9, 14, 16}, {3, 5}, {8, 18, 19}, {7, 11}, {10, 12} and {2, 13}), HVFF detected 6 F-formations ({8, 10, 12, 18, 19}, {4, 14, 16}, {1, 2, 15}, {0, 5}, {7, 11} and {3, 17}), and PWFD detected 2 F-formations ({4, 9, 14, 16, 17} and {7, 8, 10, 11, 12, 18, 19}). This is an example of how an extreme case of clustering leads to different results from the different F-formation detection methods.

The HVFF, GCFF and PWFD methods behave differently because they each use a different strategy to detect F-formations. Both HVFF and GCFF use interpersonal distance as one of the filtering criteria. In a group that is as dispersed as the one at 17:54:32, agents that are farthest from one another are not identified as members of the same F-formation. On the other hand, the PWFD model checks if the transactional segments of agents are overlapping at a particular instance. At 17:54:32, the agents towards the bottom left of the group were trying to resume movement after a stationary phase, because of which PWFD did not identify agents

0, 1, 2, 3, 5, 6, 13 and 15 as members of the F-formation. The different results obtained using the HVFF, GCFF and PWFD methods in extreme cases was a motivating factor for using the timeline approach, to examine the changes in the spatial patterns of agent clusters.

Another characteristic of Model 2a EA is that, once established, the big group almost never splits or ends. When all the twenty agents in the system are clustered together, the separation distance condition becomes violated almost as soon as an agent tries to resume movement after a stationary phase. This makes it difficult for agents to break off from the group and resume movement. Even if agents manage to break off, they again bump into the big group after wandering around for a few seconds, and stop moving – perhaps joining the group from a different position and orientation. The formation of such big clusters does not happen in any other Party World model and is unique to Model 2a EA.

5.4.5 A timeline analysis of Model 2b EA

Figure 5.17 shows the timeline of an agent cluster resulting from Model 2b EA, where agents readjust their position and orientation based on the centroid point, calculated from the (x,y) positions of agents within an extended neighbourhood. First, at 20:37:08, we see a triangular arrangement of three agents (1, 4 and 19). The equal-sided triangular arrangement of agents 1, 4 and 19 is clearly different from the kind of triadic arrangements produced by Model 2a LA (e.g. 15:57:31 in figure 5.14) or Model 2b LA (e.g. 18:20:43 in figure 5.15).

Regular polygons with equal edges are a standard outcome of Model 2b EA. For example, we can see that as more members join the group, the equilateral triangle at 20:37:08 becomes a rhombus at 20:37:33, and then a pentagon at 20:38:18. The group then splits into two at 20:38:33, one pentagonal arrangement (agents 1, 2, 4, 8 and 19), and another semi-circular arrangement (agents 0, 3, 9 and 11).

The semi-circular arrangement of agents 0, 3, 9 and 11 was not fully established as of 20:38:33. The agents were in the process of moving in and out, trying to position themselves at an equal distance from, and facing towards the centroid point. In the process, they eventually stepped into the region occupied by the pentagon of agents 1, 2, 4, 8 and 19. Consequently, at 20:39:08, the two groups start merging into one.

The merging process takes a long time, especially as more agents continue to join the group as of 20:39:38, 20:40:13, 20:40:33 and 20:41:28. Collisions occur when the group size grows and agents continue moving in and out, trying to maintain the equilibrium position, e.g. agents 0 and 3 are seen colliding at 20:39:08, agents 6 and 8 and agents 1, 4 and 19 are seen colliding at 20:41:28. Such collisions often occur in Model 2b EA, especially when the group expands beyond six agents. This finding is important in light of the fact that most existing models, which use the centroid-based approach, have not investigated the nature of positional-orientational readjustments in groups involving more than 6 members.

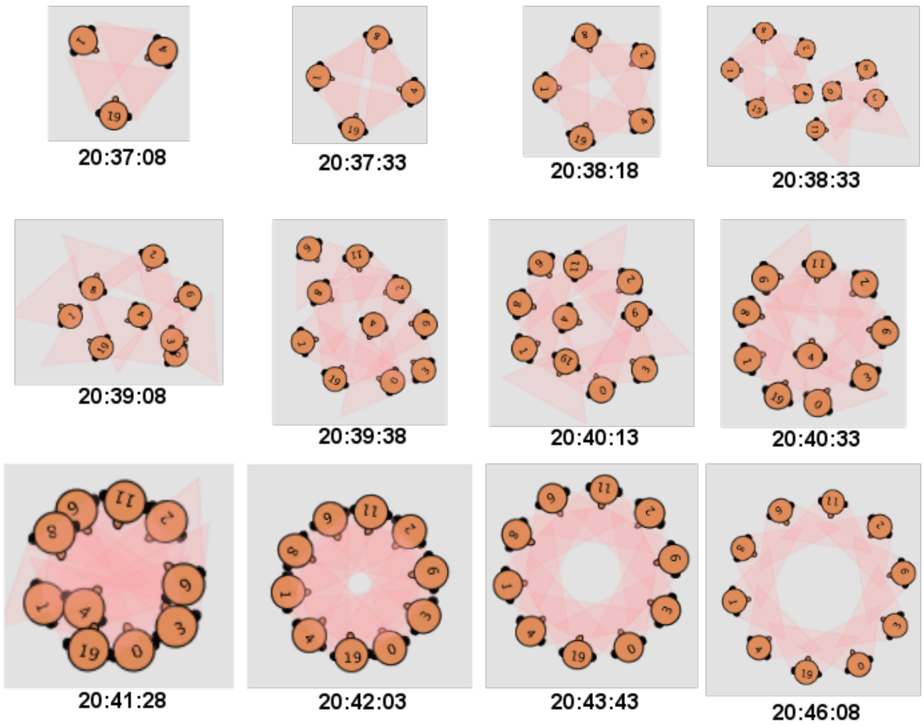


Figure 5.17: Timeline of an agent cluster produced by Model 2b EA

The agents are in a perfectly circular arrangement as of 20:42:03. Similar to Model 2a EA, the agents were not able to break off from the group, because the separation distance condition was violated almost as soon as agents resumed movement. However, unlike Model 2a EA, any change in the position or orientation of even one agent destabilises the equilibrium position of all the other agents, because the centroid point changes. Agents then begin readjusting their position and orientation to restore the equilibrium of the arrangement. The process of readjustment, where the circular arrangement expands, can be seen at 20:43:43 and 20:46:08.

Although perfectly circular arrangements are a standard outcome of Model 2b EA, the agents do not all end up in one big cluster like in Model 2a EA. At the end of each run, Model 2b EA resulted in two or three big groups, or one big group and two to three smaller groups. Lastly, as with Model 2a EA, the HVFF, GCFF and PWFD methods led to different results while processing the big clusters resulting from Model 2b EA.

5.4.6 Summary of timeline analysis

As illustrated in this section, the characteristics of agent clusters unique to each model, supports hypothesis 3: the dynamics of changes in the social, spatial and temporal characteristics of agent clusters does vary across the different Party World models. The baseline approach resulted in plausible F-formations when there were less than 4 members due to the forward movement of agents. However, the queue arrangements resulting from Model 1 are unrealistic.

The local neighbourhood influence in the field-of-view model (Model 2a LA) results in smaller groups that last for a shorter time. However, unlike Model 1, agents face one another in Model 2a LA. The local neighbourhood influence in the centroid-based model (Model 2b LA) also produces smaller agent clusters, and causes the splitting phenomenon – a bigger group breaks into co-located sub-groups when agents keep moving in and out relative to the local centroid point. The spatial patterns of agent clusters resulting from Model 2a LA and Model 2b LA are usually open or semi-open arrangements.

The extended neighbourhood influence produces bigger clusters both in Model 2a EA and in Model 2b EA. In the field-of-view model, it causes almost all the agents in the system to cluster together, and in the centroid-based model, it often results in one big group and two to three smaller groups. The agent clusters resulting from Model 2a EA and Model 2b EA last longer than the other Party World models. In the next chapter, we perform further analysis to understand the differences in the shapes and temporal characteristics of agent clusters resulting from the different Party World models.

5.5 Non-empty Party World Environments

The environment in the Party World simulations are empty, whereas in reality, face-to-face encounters occur in spaces with obstacles, such as walls, furniture, equipment, etc. We wanted to check whether adding obstacles to the Party World environment affects the emergence of F-formations. We generated three obstacles of random sizes at random locations in the Party World environment (e.g. see figure 5.18). Rules were programmed to prevent agents from colliding with obstacles. At every time step, each agent predicts its location 50 pixels ahead of its current position and orientation, assuming that its speed remains the same. If the predicted location falls within the boundary of an obstacle, the agent will alter its direction of motion to steer away from the obstacle. Except for this avoidance mechanism, all the other rules remain the same in all the Party World models.

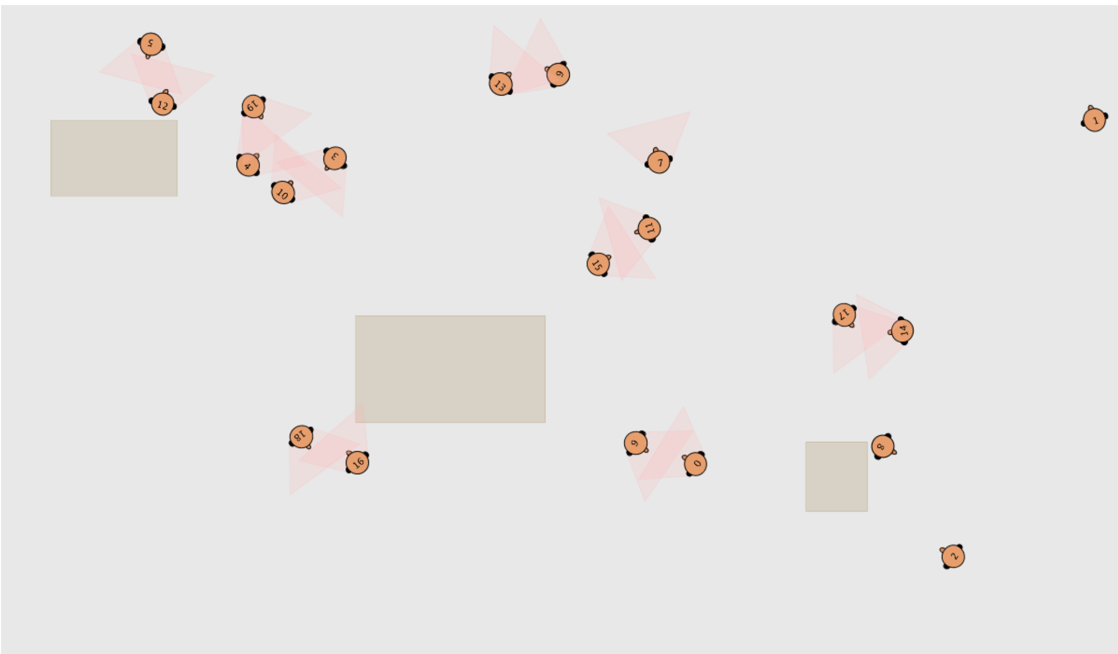


Figure 5.18: Screenshot of Model 2a LA (with obstacles in the environment)

We ran each Party World obstacle simulation model five times (30 minutes per run) to obtain the average measures of agent clusters. For each run, the height, width, and position of the three obstacles were randomly assigned for Model 1. The same measures were then used to recreate the obstacles for all the other models. We did this to ensure that for each run there were no variations in the obstacle placement across models. The Party World obstacle dataset was generated in the same way as the Party World dataset (see section 5.1.1). We used the HVFF, GCFF and PWFD methods to detect the F-formations in the dataset. The processing time taken by the HVFF and GCFF methods was 320 hours on an iMac PC, whereas PWFD did not require any extra processing time to detect F-formations (cf. section 5.1.2).

Figure 5.19 shows the median number of F-formations detected over five runs in each Party World model. Model 2b LA generated the most F-formations (HVFF=2391, GCFF=2388, PWFD=2187). The HVFF and GCFF methods detected the least F-formations in Model 2b EA (HVFF=1893, GCFF=1991), while PWFD detected the least F-formations in Model 2a EA (PWFD=1473). Model 1 and Model 2a LA were found to perform in the middle range.

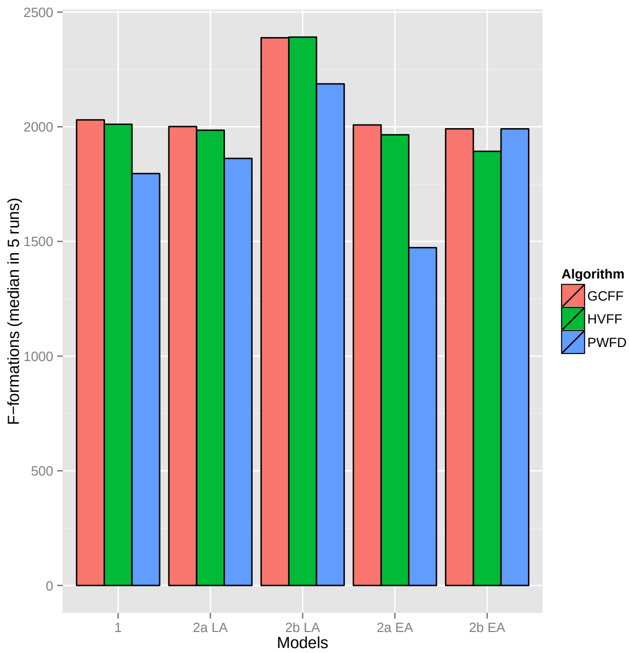


Figure 5.19: F-formations detected in the different Party World models with obstacles

The results of one-way ANOVA suggested significant variations across models. We followed up with the Tukey Kramer test to investigate specific variations and found that Model 2b LA varies significantly from all the other models in terms of producing the most F-formations. Model 2a EA also showed some variations. The results of the Tukey Kramer test are summarised in the appendix chapter in tables A.6, A.7 and A.8.

The findings reported here are similar to what we found in case of Party World models with no obstacles – the centroid-based approach with a local neighbourhood (Model 2b LA) and the field-of-view approach with an extended neighbourhood (Model 2a EA) differ significantly from all the other models (with obstacles) in terms of simulating agent clusters that qualify as F-formations.

We also performed a two-way ANOVA to compare the estimates of F-formations detected in the different Party World models with and without obstacles. The presence or absence of obstacles was one independent (nominal) variable, the models were another nominal variable, and the estimates of F-formations was the dependent variable. The results in table 5.4 show that there are no significant variations between models with regard to the presence or absence of obstacles. However, there were significant variations among the different models in terms of the F-formations they produce.

Factor	HVFF		GCFF		PWFD	
	F statistic	p value	F statistic	p value	F statistic	p value
Model	23.0792	6.077e-10	29.7647	1.634e-11	37.6250	4.663e-13
Obstacles	0.2477	0.6214	0.6056	0.4410	0.0859	0.7710
Model * Obstacles	0.5288	0.7152	0.1757	0.9496	0.3055	0.8726

Table 5.4: Summary of two-way ANOVA to assess variations in F-formations between models with obstacles and those without obstacles in the environment

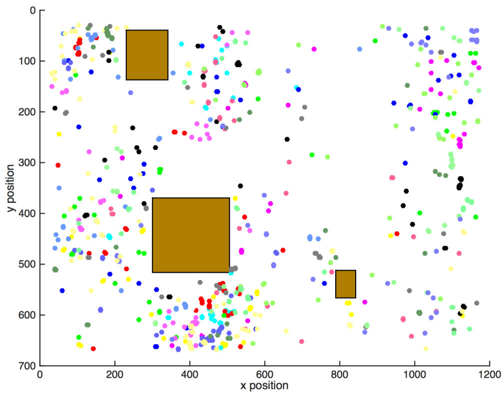
Figure 5.20 shows the spatial distribution of stationary agents from a single run of all the five Party World models with obstacles. Each distinctly coloured circle in the figure denotes the positions occupied by one of the twenty agents at different times during the 30-minute run of a simulation. The spatial distribution of agents is somewhat similar for Model 1 (figure 5.20a) and Model 2a LA (figure 5.20b), i.e. agent clusters use almost all of the available space in the Party World environment. On the other hand, in Model 2b LA and Model 2b EA, the spatial distribution of agent clusters is highly concentrated in specific areas (see figures 5.20c and 5.20e). That is to say, the centroid-based approach causes agents to utilise only a part of the available space to form agent clusters. The spatial distribution of agents in Model 2a EA is clearly different from all the other models – clustering is concentrated in one specific area, while the remaining space is almost unused (see figure 5.20d).

Lastly, although adding obstacles to the environment did not affect the emergence of F-formations, there were cases where agents collided with the obstacles while forming clusters. For example, see the violet circles on the obstacle at the bottom right in figure 5.20d, the red circles on the obstacle at the bottom left in figure 5.20e, and the yellow circles on the obstacle at the bottom left in figure 5.20c. In other words, the field-of-view approach with an extended neighbourhood influence, and the two centroid-based approaches cause agents to collide with obstacles when forming agent clusters.

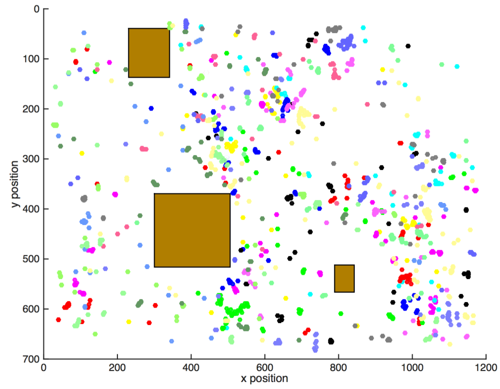
5.6 Summary of Findings

In this chapter we analysed the agent clusters resulting from the Party World models by considering F-formations as the unit of analysis. The HVFF, GCFF and PWFD methods were used to detect F-formations in the screenshots obtained from the Party World simulations. These results were used to determine the differences in the number, size and spatial distribution of F-formations generated by the different Party World models. We also used the results to perform a qualitative comparison of the dynamics of changes in F-formations resulting from a typical simulation run of each of the five Party World models.

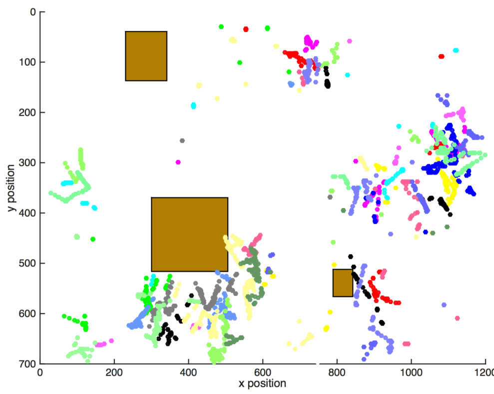
1. Differences between approaches: Comparing the agent clusters resulting from the baseline (Model 1), field-of-view (Models 2a LA and 2a EA) and centroid-based (Models 2b LA and 2b EA) approaches, we found that:



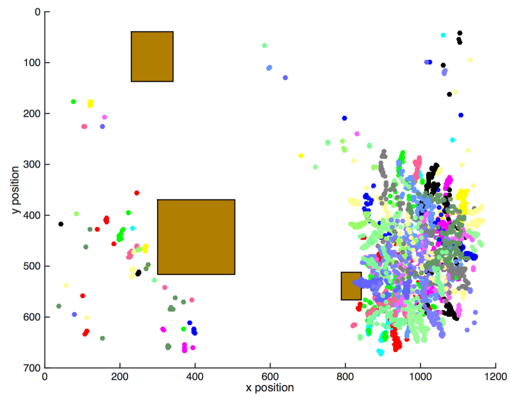
(a) Model 1



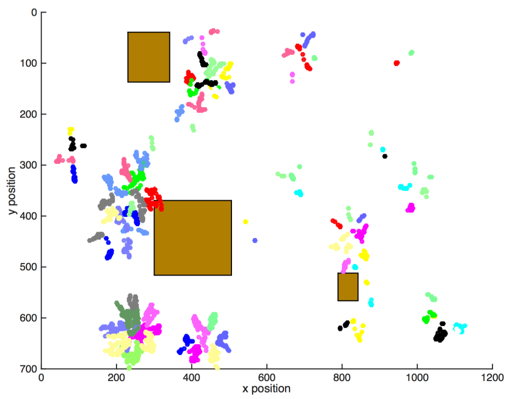
(b) Model 2a LA



(c) Model 2b LA



(d) Model 2a EA



(e) Model 2b EA

Figure 5.20: Spatial distribution of agents in a non-empty Party World environment

1a. Generation of F-formations: All the five models generate F-formations. The centroid-based approach, with a local neighbourhood influence, produces the highest number of F-formations. In the extended neighbourhood scenario, there is not much difference in the performance of the centroid-based and field-of-view approaches, in terms of generating F-formations – they both produce fewer F-formations than all the other models. Model 1 and Model 2a LA perform in the middle range in terms of generating F-formations.

1b. Changes in F-formations: The changes in the social and spatial characteristics of F-formations is distinctly different for the different Party World models. Model 1 causes agents to queue or face away from one another as the size of the group increases. The centroid-based approach causes agents to swing in and out, as if in reflex, in the process of trying to maintain an equilibrium distance from the centroid point. Furthermore, as the size of the group increases, agents end up colliding with another in the process of restoring the equilibrium position. The positional-orientational readjustments of agents in the field-of-view approach appears to be more deliberate and less reflexive than the centroid-based approach.

2. Differences between neighbourhood influences: Comparing the agent clusters resulting from the local neighbourhood influence (Models 2a LA and 2b LA) and the extended neighbourhood influence (Models 2a EA and 2b EA) of models, we found that:

2a. Differences in the number of F-formations: We found that the local neighbourhood scenarios (Model 2a LA and Model 2b LA) result in more F-formations than the extended neighbourhood scenarios (Model 2a EA and Model 2b EA).

2b. Differences in the size of F-formations: We found that extended neighbourhood influence results in clusters that have 10 or more agents. Bigger F-formations were detected in Model 2b EA than in Model 2a EA. The local neighbourhood influence results in F-formations that are comparatively smaller in size – 2 to 6 agents per group was an average measure.

2c. Differences in Open/Closed arrangements: The extended neighbourhood influence generates mostly closed arrangements, while the local neighbourhood influence generates a combination of open, closed and semi-open arrangements in Model 2a LA and Model 2b LA.

2d. Unique clustering patterns: In the timeline analysis, we showed that Model 2a EA results in a unique pattern – almost all the 20 agents in the system cluster together during the simulation. Agents move in and out of the group, but we found that the cluster itself never disintegrates. Model 2b EA did not result in such an obvious clustering as Model 2a EA, but by the end of each run, the model results in one big group and two or three smaller groups. The local neighbourhood influence does not produce any unique patterns as the extended neighbourhood influence.

2e. Differences in the number of F-formations per image: There is a marked increase in the number of F-formations per image over time in Model 2b LA and a marked decrease in Model 2b EA. This correlates with our finding that Model 2b LA produces a lot of F-formations, while Model 2b EA produces fewer F-formations. On the other hand, despite the unique clustering pattern observed in Model 2a EA, we found that the number of F-formations per image do not reduce over time in this model. This result was due to the limitation of the automatic F-formation detection methods in detecting F-formations in big clusters, which is described in more detail in chapter 8.

3. Party World models with obstacles: We modelled three randomly positioned, randomly sized, obstacles in the Party World environment and programmed behaviour rules that allow agents to avoid colliding with the obstacles. The purpose was to investigate the effect of the obstacles on the emergence of F-formations in the different models. We found that:

3a. Effect of introducing obstacles in the Party World environment: Our analysis showed that adding obstacles to the environment does not cause the models to breakdown in terms of generating F-formations. We also found that the presence or absence of obstacles does not cause significant differences in the number of F-formations resulting from the models in a 30-minute run. Similar to Party World models without obstacles, Model 2b LA produces most F-formations, Models 2a EA and 2b EA produce fewer F-formations, while Models 1 and 2a LA perform in the middle range. However, we found that agents violate the obstacle avoidance rule in the extended neighbourhood scenarios and in the centroid based approach.

3b. Spatial distribution of agent clusters: We found that the agent clusters resulting from Model 1 and Model 2a LA are spread across the Party World environment, unlike the ones resulting from Model 2b LA, Model 2a EA and Model 2b EA, which concentrate in specific areas.

We have not investigated the differences between models in terms of the exact shapes of the agent clusters they produce in this chapter, i.e. we did not investigate what type of dyadic arrangements, such as the H, V, C, I, N and L-shaped arrangements (cf. Ciolek and Kendon (1980)) result from the different models. Furthermore, the circular arrangement of agent clusters was the main focus of existing models reviewed in chapter 3. However, our qualitative analysis in section 5.4 showed that, except for Model 2b EA, none of the other Party World models produce perfectly circular arrangements. Therefore, we need to categorise and analyse the spatial patterns of multi-party agent clusters resulting from the different models. Lastly, we highlighted the fact that the bigger groups resulting from Model 2a EA and Model 2b EA do not disintegrate easily, even though agents move in and out of the groups. However, given that the predetermined stop time is 30 seconds for all agents in all models, we are yet to explore how the rules in the models affect the duration for which agents participate in clusters. To address these issues, in the next chapter, our analysis will focus on detecting, categorising, and analysing the differences in the spatial patterns and temporal characteristics of agent clusters resulting from the different Party World models in a systematic way.

Chapter 6

Analysis of the Spatial Patterns and Temporal Characteristics of Agent Clusters

---

In this chapter, we provide a systematic analysis of the different types of spatial arrangements of agent clusters resulting from the Party World simulations. First, we compare the spatial patterns of dyadic arrangements resulting from the different models. We classify the arrangements based on Ciolek and Kendon's (1980) categorisation of dyadic F-formations. We then compare if there are any significant differences between models in terms of the types of dyadic arrangements they produce. Secondly, we classify and compare the spatial patterns of multi-party agent clusters. Our classification is based on the degree of roundness of multi-party arrangements. Lastly, we compare the temporal trends of participation in agent clusters in the different models.

The rest of this chapter is organised as follows. We begin by reviewing the research question addressed in this chapter and the hypotheses we will be testing. In section 6.2, we describe the classification of the dyadic arrangements resulting from the Party World simulations under different categories, and then compare the differences between models. Similarly, in section 6.3, we discuss the basis for considering the roundness of shapes as a measurement attribute for multi-party F-formations, and then compare the differences between models. In section 6.4, we compare the temporal characteristics of agent clusters in the different models. We conclude the chapter with a summary of findings.

6.1 Research Question

The research question addressed in this chapter is: What are the differences in the spatial patterns and durations of agent clusters resulting from the different Party World models? To answer this question, we consider the following hypotheses:

Hypothesis 4: The spatial patterns of agent clusters resulting from the Party World simulations will be different for the different models. Based on the findings of the previous chapter, we expected that the local neighbourhood influence in models will result in more open arrangements than the ones resulting from models using the extended neighbourhood influence.
The spatial manifestation of an F-formation is determined by the arrangement of interactants' bodies in space (Kendon, 1990). Since each Party World model uses a different strategy to control the position and orientation of agents, we hypothesise that the spatial patterns of agent clusters will vary across the different models. We will be comparing the outcomes of dyadic and multi-party arrangements separately, so we consider H4 in two parts:

Hypothesis H4a: The spatial patterns of dyadic clusters resulting from the Party World simulations will be different for the different models.

Hypothesis H4b: The degree of roundness of multi-party F-formations resulting from the Party World simulations will be different for the different models.

Hypothesis 5: The duration for which agents participate in clusters will vary across the different models.
We described in chapter 4 that when participating in a group the duration for which agents have to remain stationary is prefixed at 30 seconds. This duration is the same for all agents in all the models. Hence, a likely assumption would be that the duration for which agents participate in clusters will be the same across all models. However, while analysing the timelines of agent clusters in the previous chapter,

we found that Model 2a EA and Model 2b EA result in unique clustering patterns. Model 2a EA results in one big cluster and Model 2b EA results in one big cluster and two to three smaller clusters, which last longer notwithstanding changes in participants. This result partly confirms that the temporal patterns of agent clusters vary across models. To further investigate the temporal differences between models, we hypothesise here that the duration for which agents participate in clusters will be different for the different models. Testing this hypothesis will provide further insights into the temporal dynamics of agent clusters brought about by the different models.

6.2 Analysis of Dyadic Patterns

Ciolek and Kendon (1980) identified six different categories of dyadic F-formations: (1) H-shaped (vis-à-vis) means individuals face one another almost directly; (2) N-shaped also means face-to-face arrangement, but the bodies are out of alignment at least by half the body-width; (3) V-shaped means the angle formed by the intersection of transactional segments is 45°; (4) L-shaped means the angle formed by the intersection of transactional segments is 90°; (5) C-shaped means the angle formed by the intersection of transactional segments is 135°; (6) I-shaped refers to side-by-side arrangements (see section 2.7 for a detailed review of these arrangements).

We classify the dyadic arrangements of agent clusters resulting from the Party World models based on these six different categories. We also consider two further categories: a *queue* arrangement and a *back-to-back* arrangement. Queue arrangement means agents are one behind the other. Back-to-back arrangement means agents have their backs to each other. Figure 6.1 shows the eight types of dyadic arrangements, based on which we classify the dyadic clusters resulting from the different Party World models.

6.2.1 Categories of dyadic spatial patterns

We programmed a module in the Party World simulation platform to detect and classify the dyadic arrangements of agents. If the transactional segments of two

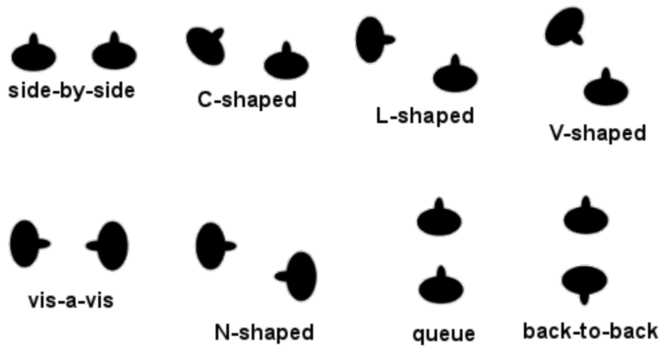


Figure 6.1: Classification of dyadic arrangements

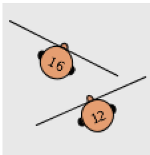
(and only two) agents overlap, then the type of arrangement is calculated based on the supplementary angle of the relative orientation of agents, which we call θ_{dyad}. The values of angles defining Ciolek and Kendon's (1980) H, N, V, C, L and I-shaped arrangements are only meant as a reference and not as exact measures. Therefore, instead of classifying the dyadic arrangements based on a single value for θ_{dyad}, we have considered a range of values for each of the six categories of arrangements (see table 6.1).

The dyadic pattern detection module runs at the same time as the simulation. Once every five seconds, the module detects all the active dyadic arrangements, identifies their shapes, and records the outcomes – the simulation frame number, the members and shapes of the dyadic arrangements are recorded. Our analysis of the spatial patterns of dyadic arrangements resulting from the Party World models are based on the data collected using the dyadic pattern detection module.

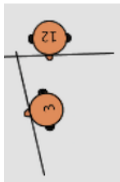
Arrangement	Ciolek and Kendon's (1980) reference values for relative orientation of dyads	The range of values we have considered for θ_{dyad}
H-shaped	$\approx 0°$	[170°..180°]
N-shaped	$\approx 0°$	[170°..180°]; but bodies are out of alignment
V-shaped	$\approx 45°$	[101°..169°]
L-shaped	$\approx 90°$	[80°..100°]
C-shaped	$\approx 135°$	[11°..79°]
I-shaped	$\approx 180°$	[0°..10°]

Table 6.1: Classification of dyadic arrangements based on θ_{dyad}

We have categorised the H, N and V-shaped patterns as closed dyadic arrangements, and the L, C and I-shaped arrangements as open dyadic arrangements (cf. Ciolek and Kendon (1980)). We have classified queue and back-to-back arrangements in a category called *Non F-formations.* Sometimes agents form a queue that is not exactly a straight line, e.g. figure 6.2. We have categorised these sort of arrangements as *other queues.*



(a) Example 1



(b) Example 2

Figure 6.2: Examples of other queue patterns

6.2.2 Data for analysis

The dyadic arrangements resulting from the Party World simulations were obtained over five 30-minute runs of each model. Following the same procedure described in section 5.1.1 of the previous chapter, screenshots of the simulation were recorded once every five seconds while the simulations were running. This produced 360 images per run of each model and 1800 images across five runs of each Party World model. When the screenshots were taken, the dyadic pattern detection module detected and classified the dyadic arrangements at the time, and saved the results.

6.2.3 The dyadic arrangements resulting from the different Party World models

Figures 6.3a, 6.3b and 6.3c show the median number of different types of closed dyadic arrangements detected over five runs of the different models. As seen in the figures, Models 2b LA and 2b EA, i.e. both the centroid-based models, mostly generate vis-à-vis (H-shaped) arrangements. The centroid-based models produced fewer V-shaped arrangements and no N-shaped arrangements. The field-of-view models show a different trend compared to the centroid-based models. Model 2a LA mostly generates V-shaped arrangements and comparatively fewer N-shaped and

H-shaped arrangements. Model 2a EA mostly produces V-shaped arrangements, followed by the N-shaped arrangements, and very few H-arrangements.

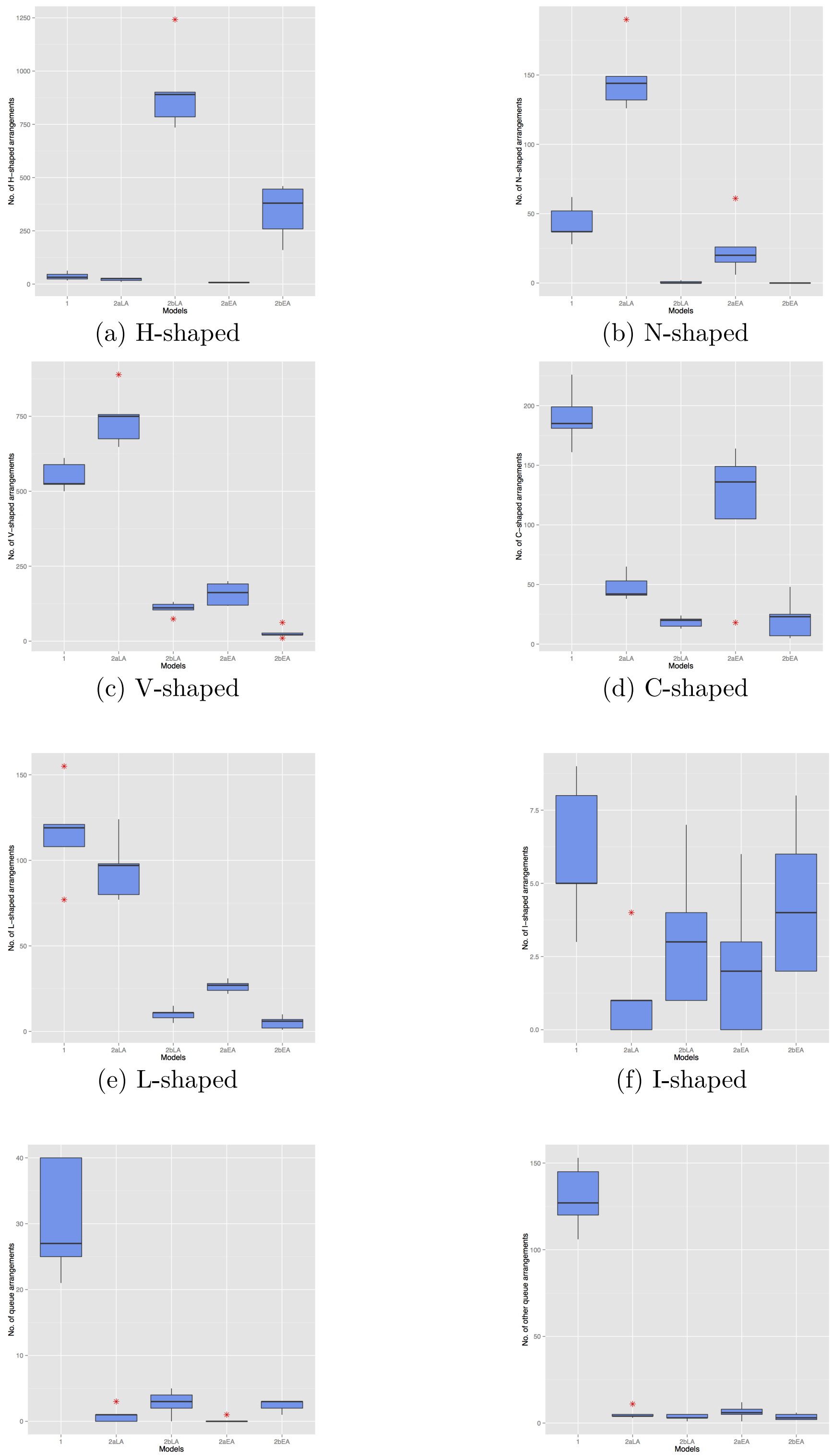
Figures 6.3d, 6.3e and 6.3f show the median number of different types of open dyadic arrangements detected over five runs of the different models. We see that the field-of-view models generate more open dyadic arrangements than the centroid-based models. Model 2a LA generates a lot of L-shaped arrangements, followed by the C-shaped arrangements, and very few I-shaped arrangements. Model 2a EA mostly produces the C-shaped arrangement, followed by L-arrangements, and few I-shaped arrangements. The centroid-based models, Model 2b LA and Model 2b EA, produce comparatively fewer C, L and I-shapes.

Figures 6.3g and 6.3h show the median number of queue and other queue dyadic arrangements detected over five runs of the different models. We can see that the field-of-view models and centroid-based models produce very few queue and other queue arrangements.

The baseline version, Model 1, mostly produces V-shaped closed arrangements, followed by the N-shaped and H-shaped arrangements. Model 1 mostly produces the C-shaped open arrangements, followed by the L-arrangement, and very few I-arrangements. Compared to other models, Model 1 produces a lot of other queue and queue arrangements. Since agents do not readjust their position or orientation in Model 1, it seems reasonable that this model produces more queue arrangements than any other Party World model. Lastly, we found that none of the Party World models produce back-to-back arrangements.

6.2.4 Analysis of variations across models

We performed a one-way ANOVA, with Welch's correction to compensate for heteroscedascity, to examine whether the patterns of dyadic arrangements resulting from the Party World simulations vary significantly across models. Table 6.2 shows the F-statistics and p-values obtained by comparing the frequency of closed arrangements (the H, N, and V-shaped arrangements) generated by the different Party World models. The p-values indicate that the models produce significantly different



(a) H-shaped (b) N-shaped

(c) V-shaped (d) C-shaped

(e) L-shaped (f) I-shaped

(g) Queue (h) Other

Figure 6.3: Dyadic arrangements resulting from the Party World models

numbers of H, N, and V-shaped arrangements.

We performed the Tukey Kramer test to do pairwise comparisons. The results confirmed that Model 2b LA and Model 2b EA are significantly different from one another and the other models in terms of generating H-shaped arrangements. Model 2b LA produces significantly more H-shaped F-formations, while Model 2b EA produces significantly fewer H-shapes. Considering this result together with the size of F-formations resulting from the different models (cf. section 5.3.2), we found that almost all dyadic F-formations resulting from Model 2b LA are H-shaped, and the few dyadic F-formations resulting from Model 2b EA are mostly H-shaped.

For the N-shaped arrangements, we only compared Model 1, Model 2a LA and Model 2a EA, because Model 2b LA and Model 2b EA did not produce N-shaped arrangements. The results showed that Model 2a LA produces significantly more N-shapes than Model 1 and Model 2a EA. Lastly, the Tukey Kramer test for V-shaped arrangements also showed significant variations between models. Model 1 and Model 2a LA produce significantly more V-shaped dyadic patterns than the other models. Results also showed that Model 2a EA and Model 2b EA vary from one another. The full pairwise comparisons are shown in figure A.2 in appendix chapter A.

H-shaped		N-shaped		V-shaped	
$F_{4,20}$	p value	$F_{2,12}$	p value	$F_{4,20}$	p value
32.8701	$p<0.001$ (4.08e-05)	37.733	$p<0.001$ (1.262e-04)	152.9306	$p<0.001$ (1.165e-08)

Table 6.2: Summary of one-way ANOVA to assess variations in closed dyadic arrangements

Table 6.3 shows the F-statistics and p-values obtained by comparing the frequency of open dyadic arrangements resulting from the Party World models. The differences in the occurrence of C-shaped arrangements in the different Party World models was highly significant ($p<0.001$). The follow-up Tukey Kramer test showed that Model 1 and Model 2a EA vary significantly from one another and from all the other models. The p-value was again highly significant (< 0.001) for comparisons on L-shaped arrangements. The follow-up Tukey Kramer test showed that Model 1 and Model 2a LA vary significantly from all other models. Lastly, with respect to the I-shaped arrangements, the models behaved similarly in that they all generated very few I-shaped patterns. There were no significant differences between models in terms of generating I-shaped dyadic patterns (p-value = 0.06156). The full pairwise comparisons are shown in figure A.2 in the appendix chapter A.

C-shaped		L-shaped		I-shaped	
$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value
56.2636	$p<0.001$ (2.843e-06)	52.2724	$p<0.001$ (1.651e-06)	3.2176	**$p>0.05$** (6.156e-02)

Table 6.3: Summary of one-way ANOVA to assess variations in open dyadic arrangements

Table 6.4 shows the F-statistics and p-values obtained by comparing the frequency of queue and other queue arrangements resulting from the different Party World models. The p-values indicate that there are variations between the Party World models, while the follow-up Tukey Kramer test confirmed that Model 1 produces the highest number of queue and other queue arrangements than any other model (see figures A.2g and A.2h in appendix chapter A).

Based on these results, we conclude that hypothesis H4a is supported for certain types of dyadic spatial arrangements, i.e. there are significant variations between the Party World models in terms of their ability to generate certain types of dyadic arrangements. There are variations in the closed arrangements such as the H, N and V-shaped arrangements. There are also variations in the C and L-shaped open arrangements and queue arrangements. However, there are no differences between the Party World models in terms of generating I-shaped and back-to-back arrangements – the models produce very few I-shaped arrangements and no back-to-back arrangements.

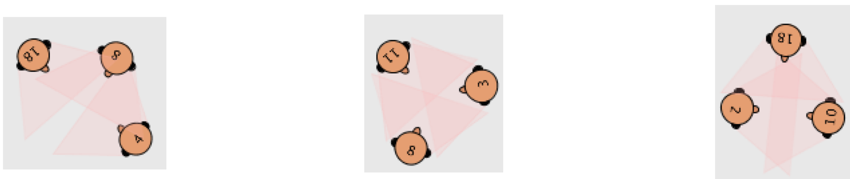
queue		other queue	
$F_{4,20}$	p value	$F_{4,20}$	p value
17.8757	$p<0.001$ (2.487e-04)	46.4098	$p<0.001$ (3.013e-06)

Table 6.4: Summary of one-way ANOVA to assess variations in queue arrangements

6.3 Analysis of the Spatial Patterns of Multi-party Agent Clusters

In this section, we compare the roundness of the spatial arrangements of multi-party agent clusters resulting from the different Party World models. In chapter 2, we found that the spatial patterns of multi-party conversational groups have not been analysed in a detailed manner like dyadic groups. The most common assumption has been that multi-party groups are circular. Other assumptions include triangular, square, parallelogram, oval, and semi-circular arrangements.

We could have classified the spatial patterns of multi-party agent clusters resulting from the Party World models using the regular shape categories. However, except for the multi-party agent clusters that resulted from Model 2b EA, none of the other models produced perfectly regular shapes. Hence, regular shapes were an inappropriate classification scheme to identify the spatial patterns of multi-party agent clusters; at least not as straightforward a way as the classification of dyadic arrangements based on θ_{dyad} (cf. section 6.2.1). We did not manually annotate the shapes due to issues of ambiguity, e.g. figures 6.4a, 6.4b, 6.4c can all be classified as triangles, or only figure 6.4b can be classified as a properly triangular arrangement.



(a) Triangle arrangement 1 (b) Triangle arrangement 2 (c) Triangle arrangement 3

Figure 6.4: Examples of triangular arrangements

To overcome these potential limitations, we developed a strategy to detect the roundness of shapes of multi-party agent clusters, and to classify the shapes based on the degree of roundness. We adapted the convex hull gift wrapping algorithm to detect the roundness of shapes. The roundness detection module is built into the Party World simulation platform. The module provides estimates of the roundness of shapes in the range of [0,1], where values approaching 1 means very circular, and values approaching 0 means not very circular. By using the degree of roundness as a measure, we attempt to achieve an objective classification of the spatial patterns of multi-party agent clusters produced by the Party World models.

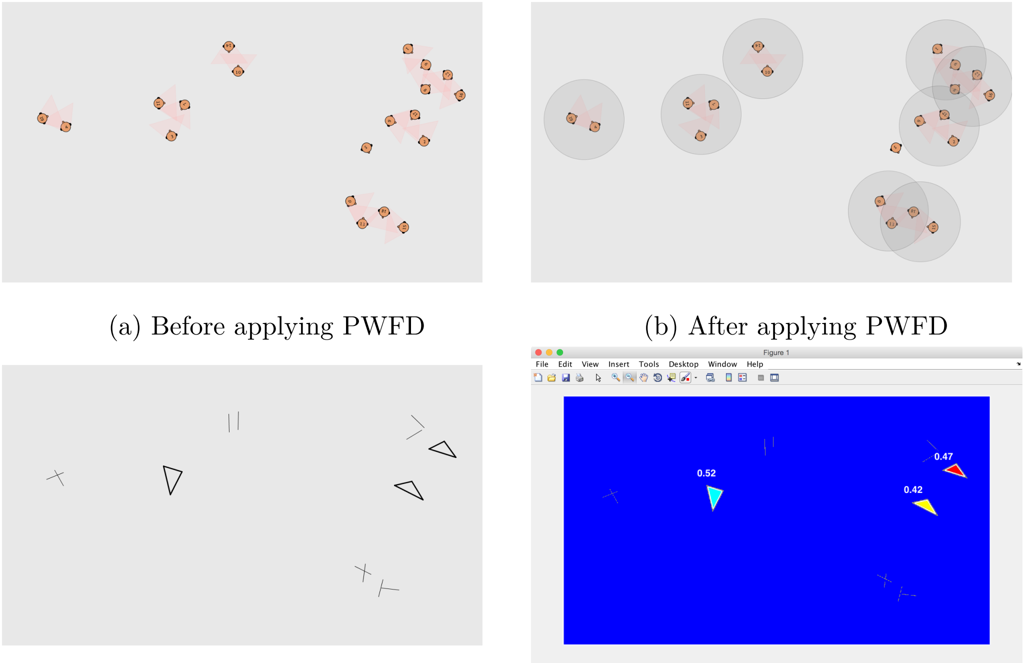
6.3.1 The Convex Hull Gift Wrapping module

We developed a module, based on the convex hull gift wrapping algorithm or Jarvis algorithm (Jarvis, 1973), to compute the convex hull of a multi-party agent cluster, i.e. the pattern formed by connecting the body positions of the outermost agents in a cluster. We have called this the Jarvis module and it runs at the same time as the Party World simulations. The module takes the multi-party F-formations (i.e. 3 or more participants) identified by the PWFD method as input, and computes the convex hull of each of these F-formations. Figure 6.5a is a screenshot from Model 2a LA, figure 6.5b shows the F-formations detected by PWFD in the screenshot, figure 6.5c shows the outcome of the Jarvis module after processing the convex hull of each F-formation detected in the screenshot.

The convex hull shapes detected by the Jarvis module are fed into another module, built using the image processing tools in MATLAB, for detecting the roundness of shapes. Figure 6.5d is the result obtained by processing the convex hull shapes seen in figure 6.5c. We classified the results obtained from the roundness estimation module under ten categories – multi-party agent clusters whose roundness is in the range of [0, 0.1] is one category, range [0.11, 0.2] is another category, and so on, with roundness in the range of [0.91, 1.0] as the tenth category.

6.3.2 Data for analysis

The multi-party arrangements resulting from the Party World simulations were obtained over five 30-minute runs of each model. There were 360 images per run of a model and 1800 images gathered over five runs of each Party World model. When the screenshots were taken, the multi-party F-formations were detected (using PWFD), and transformed to convex hulls (using Jarvis). After the simulations were complete, the images (with the convex hulls) were input into the roundness estimation module. It took around 40 hours on an iMac PC (Intel Core 2 Duo processor with



(a) Before applying PWFD

(b) After applying PWFD

(c) After applying Jarvis

(d) After roundness estimation

Figure 6.5: Demo of Jarvis and roundness estimation modules to compute the roundness of multi-party agent clusters

speed 2.66 GHz and 4GB RAM) for the roundness estimation module to process the images from all the five Party World models.

6.3.3 The different types of multi-party arrangements resulting from the Party World models

It can be seen in figure 6.6a that very few agent clusters with roundness in the range of [0, 0.1] are generated by the Party World models. Models 1, 2a LA and 2a EA produce more clusters with roundness in the range of [0.11, 0.5] than Models 2b LA and 2b EA. Figure 6.6b shows that Models 2b LA and 2b EA generate more shapes with roundness in the range of [0.51, 0.6] than the other three models. We also found that Model 2b LA produces a lot of clusters with roundness in the range of [0.61, 0.7].

Model 2b EA generates more clusters with roundness in the range of [0.71, 0.8] than Model 2b LA, as well as the other models. Model 2b EA also produces a lot of clusters with roundness in the range of [0.81, 0.9] and some in the range of [0.91, 1.0]. Model 2a EA results in quite a few clusters with roundness in the range of

[0.81, 0.9]. Models 1, 2a LA and 2b LA produce very few clusters with roundness in the range of [0.81, 0.9] and none with roundness in the range of [0.91,1.0].

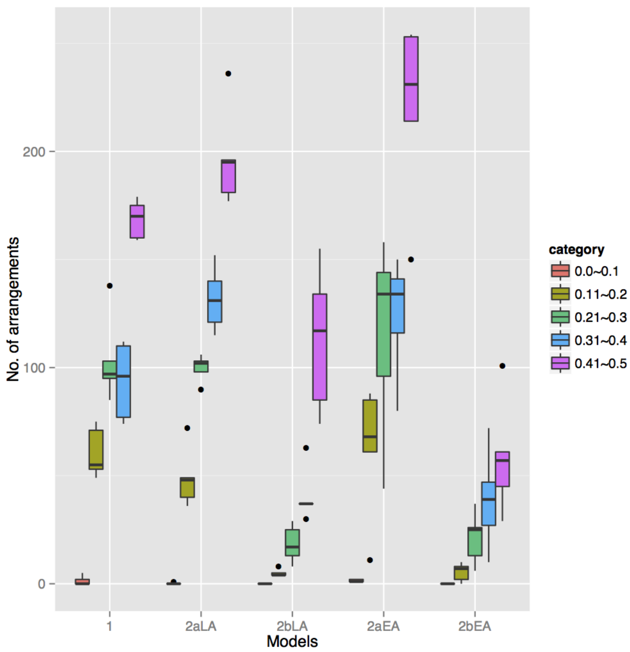
6.3.4 Analysis of variations across models

Table 6.5 shows the F-statistics and p-values obtained by comparing F-formations, with roundness in the range of [0.11, 0.5], resulting from the different Party World models. The p-values indicate significant differences between the Party World models in terms of the multi-party F-formations they produce with roundness in the range of [0.11, 0.5]. The follow-up Tukey Kramer test to obtain pair-wise comparisons confirmed that Models 1, 2a LA and 2a EA vary significantly from Models 2b LA and 2b EA in terms of producing F-formations with roundness in the range of [0.11, 0.2]. Likewise, for F-formations with roundness in the range of [0.21, 0.3] and [0.41, 0.5], the Tukey Kramer test confirmed that Models 1, 2a LA and 2a EA vary significantly from Models 2b LA and 2b EA. For shapes with roundness in the range of [0.31, 0.4], we found that Model 2a LA varies significantly from Model 1, as well as from Models 2b LA and 2b EA. Furthermore, Models 1 and 2a EA vary significantly from Models 2b LA and 2b EA. The full pairwise comparisons are shown in figure A.3 in the appendix chapter A.

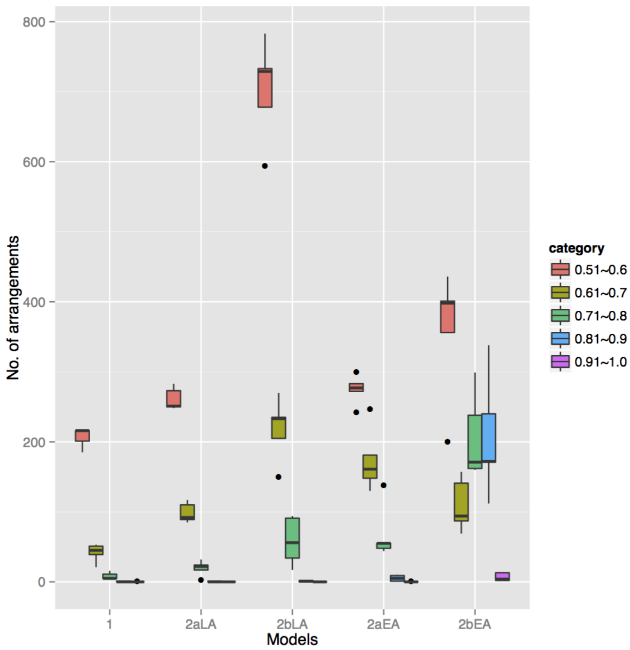
[0.11, 0.2]		[0.21, 0.3]		[0.31, 0.4]		[0.41, 0.5]	
$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value
35.5639	$p<0.001$ (2.183e-05)	81.2662	$p<0.001$ (2.597e-07)	29.0352	$p<0.001$ (1.982e-05)	22.1915	$p<0.001$ (1.106e-04)

Table 6.5: Summary of one-way ANOVA to assess variations in multi-party F-formations (Roundness: 0.1 to 0.5)

Table 6.6 shows the F-statistics and p-values obtained by comparing F-formations, with roundness in the range of [0.51, 0.9], resulting from the different Party World models. For F-formations with roundness in the range of [0.51, 0.6], the Tukey Kramer test showed significant differences between Model 1, Model 2b LA, and Model 2b EA. For F-formations with roundness in the range of [0.61, 0.7], there were significant differences between Model 1 and Model 2b LA. There were also variations between Models 2a LA, 2a EA and 2b EA. Lastly, the Tukey Kramer test



(a) Roundness 0.0 to 0.5



(b) Roundness 0.5 to 1.0

Figure 6.6: Roundness of multi-party arrangements resulting from the Party World models

showed that Model 2b EA differs significantly from all the other models with respect to generating F-formations with roundness in the range of [0.71, 0.9]. The full pair-wise comparisons are shown in figures A.3e, A.3f, A.3g and A.3h in the appendix chapter A.

[0.51, 0.6]		[0.61, 0.7]		[0.71, 0.8]		[0.81, 0.9]	
$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value	$F_{4,20}$	p value
54.5579	$p<0.001$ (1.459e-06)	25.8756	$p<0.001$ (4.313e-05)	15.2093	$p<0.001$ (5.337e-04)	7.4873	$p<0.05$ (5.546e-03)

Table 6.6: Summary of one-way ANOVA to assess variations in multi-party F-formations (Roundness: 0.5 to 0.9)

Because only few or no shapes with roundness in the ranges of [0, 0.1] and [0.9, 1.0] were produced by the models, we did not include these two categories in our analysis. However, apart from these two categories, there are considerable differences between the different Party World models in terms of the roundness of F-formations they produce. Models 1, 2a LA and 2a EA produce F-formations with roundness mostly in the range of [0.11, 0.5]. Among these, Model 2a EA produces more F-formations with roundness in the range of [0.11, 0.5], followed by Model 2a LA, and lastly Model 1. The centroid-based Model 2b LA produces a lot of F-formations with roundness in the range of [0.51, 0.7]. Model 2b EA tops the list in generating F-formations with roundness in the range [0.71, 0.9]. Therefore, as expected based on observations of the qualitative time series data in section 5.4, the centroid-based model with an extended neighbourhood influence produces more round multi-party F-formations than the field-of-view models. These results support Hypothesis H4b – the degree of roundness of multi-party F-formations resulting from the Party World simulations does vary depending on the model used.

6.4 Temporal organisation of agent clusters

An F-formation is a temporally regulated system and it lasts notwithstanding changes in participants (Kendon, 1990). The timeline analysis in the previous chapter, which showed the unique clustering patterns observed in some of the Party World models, confirmed qualitatively that the temporal dynamics of agent clusters varies between models. To further confirm this finding, we investigate whether the duration for which agents participate in clusters differs between the different Party World models, despite the stop time being fixed at 30 seconds in the models. This is an exploratory analysis for testing hypothesis H5 – the duration for which agents participate in clusters will vary across the different Party World models. The data used for analysis was obtained from one 30 minute run of each of the five Party World models.

6.4.1 Data for analysis

We implemented a timekeeping module in the Party World simulation platform to keep track of the duration for which agent clusters last. Several agent clusters are formed during each (30 minute) run of the simulations. On the other hand, each one of the twenty agents in the system participates in one or more clusters in each run of the simulation. The timekeeping module records the duration of every cluster in which each agent participates in. The module creates a dataset consisting of N columns, where N = 20, corresponding to the twenty agents used in the simulations. Therefore, for every run of the simulation, column C0 keeps track of the duration of all the clusters that qualify as F-formations in which agent A0 participates. As before, the agent clusters that qualify as F-formations are identified using the PWFD method.

The timekeeping module records a screenshot of the simulation once every five seconds, as well as recording the F-formations (if any) in which each of the twenty agents are participating in, and the time for which the F-formations have lasted. The duration of F-formations is calculated based on the membership of agents. Between screenshots, if the F-formation in which a particular agent participates has retained at least one member more than just the agent, then the duration of the F-formation is increased by 5 seconds. On the other hand, if an F-formation in which an agent participates has ended, either because the agent itself resumed movement, or because there was a complete turnover of other agents in the group, the duration counter is reset to zero. After each run of the simulation, we identified the longest lasting F-formation corresponding to each agent in the system. Our analysis of the temporal trends of participation in agent clusters in the different Party World models is based on this data.

6.4.2 Analysing the duration of agent clusters

Figure 6.7 shows the maximum duration for which each agent participated in the same F-formation in Model 1. The red line in the graph at the 30-second mark denotes the predetermined stop time used in the simulations. Each row on the Y-axis denotes the duration of the longest lasting F-formation in which the agent (whose id is the same as the row number) participated. The duration for which F-formations lasted (i.e. beyond the red line) is a result of the dynamics of the positional-orientational behaviour of agents brought about by Model 1. As seen in figure 6.7, the maximum duration for which an agent participated in the same F-formation in Model 1 is just under 200 seconds.

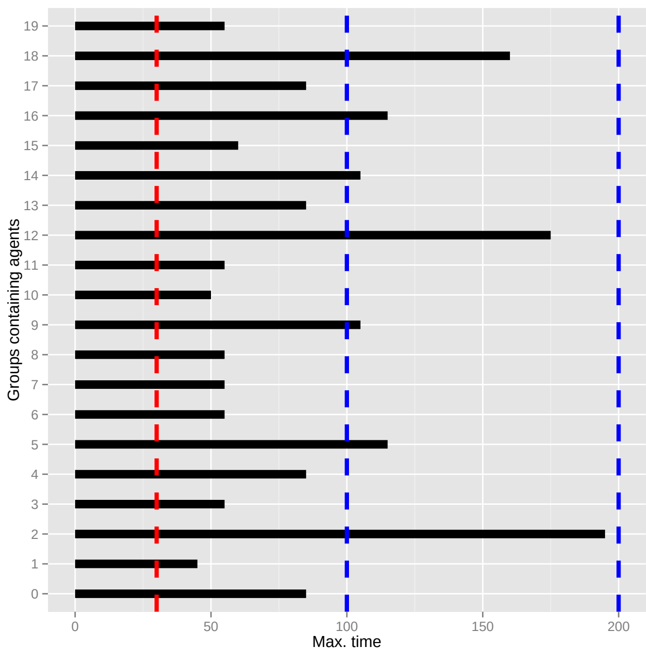
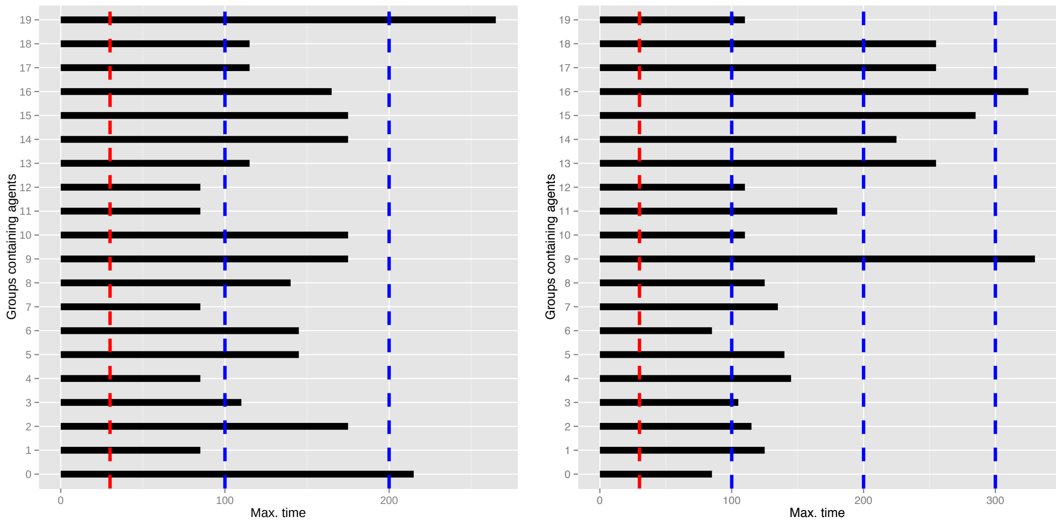


Figure 6.7: The maximum duration for which each agent participants in an F-formation in Model 1

Figure 6.8a shows the maximum duration of F-formations corresponding to each agent in the system using Model 2A LA. As seen in the figure, the maximum duration for which an agent participated in the same F-formation in Model 2a LA is over the 200-second mark. Furthermore, unlike Model 1, several agents remained in the same F-formation for longer a time in Model 2a LA – 15 out of the 20 agents remained in the same F-formation for at least 100 seconds. The results of Model 2a EA in figure 6.8b show that 18 out of the 20 agents participated in the same F-formation for more than 100 seconds, while 7 of them participated in the same F-formation for more than 200 seconds. A couple of them remained for longer than 300 seconds in the same F-formation.

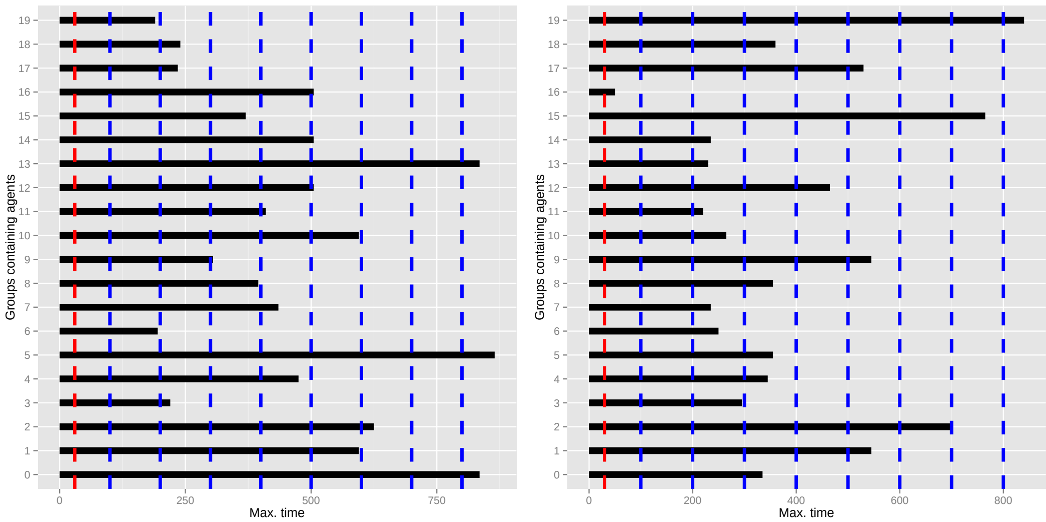


(a) Model 2a LA (b) Model 2a EA

Figure 6.8: The maximum duration for which each agent participants in an F-formation in the field-of-view models

Figure 6.9 shows the maximum duration of F-formations corresponding to each agent in the system using Models 2b LA and 2b EA. The maximum duration for which an agent participated in the same F-formation in Model 2b LA is over the 750-second mark, while in Model 2b EA it's over the 800-second mark. We see that most agents remained in the same F-formation for more than 200 seconds in both the centroid-based models.

Based on the results, we conclude that Hypothesis H5 is true when considering F-formations as the unit of analysis, i.e. the duration for which agents participate in F-formations differs between the different Party World models. The centroid-based models cause agents to remain in the same F-formation for longer than the field-of-



(a) Model 2b LA (b) Model 2b EA

Figure 6.9: The maximum duration for which each agent participants in an F-formation in the centroid-based models

view and baseline models. On the other hand, there are no appreciable differences between the local and extend neighbourhood influences within models in terms of affecting the time for which agents participate in the same F-formation.

We draw three inferences by comparing these results with the timeline analysis in the previous chapter. Firstly, in the previous chapter, we reported the observation that Model 2a EA results in big clusters that almost never ends, notwithstanding changes in participants. Our earlier observation that agents move in and out of the big cluster in Model 2a EA correlates with the latest finding that the maximum time for which agents participate in the same F-formation in Model 2a EA is just over 300 seconds. After this time, agents withdraw from the cluster, but bump into it again after wandering around only for a few seconds.

Secondly, in the previous chapter, we reported the observation that Model 2b EA often results in one big cluster and two or three smaller clusters. Considering this observation in light of the temporal pattern in figure 6.9b suggests that, not only do agent clusters last long in Model 2b EA, but agents also remain in the same cluster for a longer time than the field-of-view models.

Thirdly, the temporal patterns in figure 6.9a suggest that in Model 2b LA agents participate in the same F-formation almost as long as they do in case of Model 2b

EA, despite not showing any unique clustering pattern as observed in Model 2b EA. This happens as a consequence of two effects. The first reason is the splitting of clusters observed in Model 2b LA described in the previous chapter, i.e. clusters split when agents readjust their position and orientation based on the local centroid point. The second reason is our design of the time keeping module, where we consider that an F-formation lasts even if one member other than the agent is the same between screenshots. Since this condition will be met when the clusters split, as it often occurs in Model 2b LA, we find that Model 2b LA produces a temporal pattern similar to Model 2b EA in terms of the longest duration for which agents participate in the same F-formation. In summary, not only do clusters last for different durations across models, but the duration for which agents remain in the same cluster also varies between the different Party World models.

6.5 Summary of Findings

Our objective in this chapter was to explore and understand the influence of the positional-orientational behavioural rules, used in each of the different Party World models, on the spatial patterns and temporal characteristics of the resulting agent clusters. We found that:

1. Dyadic arrangements: We used nine different categories of dyadic arrangements – Ciolek and Kendon's (1980) classification of H, C, I, N, V and L-shaped arrangements together with the queue, back-to-back and other queue arrangements – as the reference to classify the shapes of dyadic agent clusters resulting from the Party World simulations. We found that:

1a. The centroid-based approach versus the field-of-view approach: We found that Models 2b LA and 2b EA mostly produce only H-shaped arrangements. On the other hand, Models 2a LA and 2a EA generate almost all types of dyadic arrangements. Based on this finding, and that H-shaped arrangements are normally preferred for intimate and confrontational interactions (cf. Scheflen (1975)), it can be said that the field-of-view approach is more effective than the centroid-based approach in simulating realistic dyadic shapes.

1b. The baseline approach: Despite having no rules to enable positional-orientational readjustments, we found that Model 1 still generates a lot of V, C and L-shaped arrangements. Model 1 produces comparatively fewer H and N-shaped arrangements. The findings can be attributed to the model assumption that movement is in the forward direction. Because of this, it's likely that when agents stop moving, they end up facing one another, even if not in face-to-face arrangements. Nevertheless, compared to all the other models, Model 1 produces considerably more queue and other queue arrangements. For this reason, Model 1 may not be better than the field-of-view approach in simulating realistic dyadic F-formations.

1c. Side-by-side arrangements : None of the Party World models generate a considerable number of I-shaped arrangements. I-shaped arrangements are normally preferred when there is an external focus of attention (cf. section 2.7.1). In future studies, we could model an external event/object in the Party World environment, to analyse its effects on the resulting spatial patterns of agent clusters.

2. Multi-party arrangements: The spatial patterns delineated by multi-party conversational groups have not been studied in a detailed manner like dyadic arrangements. Therefore, instead of classifying multi-party arrangements based on categories such as triangle, circle, square, etc., we chose the roundness of shapes as a measure. We implemented the Jarvis module and the roundness detection module to estimate the roundness of multi-party agent clusters resulting from the Party World simulations. Roundness was measured in the range of [0, 1.0] with 0 meaning least circular and 1 meaning most circular. We then used the measures of roundness to compare the differences between models.

2a. Comparison between models: The roundness estimates of F-formations resulting from Model 2b EA were mostly greater than 0.5, and compared to all the other models, Model 2b EA produces the most circular F-formations. The other centroid-based model, i.e. Model 2b LA, produced many F-formations with roundness in the range [0.51, 0.6]. Generally, the field-of-view approach resulted in F-formations that are not as circular as the ones produced by the centroid-based approach. The roundness of F-formations resulting from Model 1 was also on the lower end of the range.

3. Duration of agent clusters: The effect of the centroid-based model on the duration of agent clusters and the time for which agents remain in the same cluster have never been examined before. We implemented a time keeping module in the Party World simulation platform to record the maximum duration for which agents remain in the same F-formation. We combined the results with the insights from the timeline analysis in the previous chapter and inferred that:

3a. Difference between approaches: The centroid-based approach causes agents to remain in the same cluster for a longer time than the field-of-view and baseline models. Within models, the local and extended neighbourhood influences do not cause appreciable differences in the maximum time for which agents remain in the same cluster. On the other hand, unlike the local neighbourhood influence, the extended neighbourhood influence results in unique clustering patterns in Model 2a EA and Model 2b EA. In Model 2a EA the clusters themselves last long notwithstanding changes in participants, while in Model 2b EA the clusters last long as well as retaining the same participants for a long time.

3b. Correspondence to a real-world observation: Considering the temporal patterns of agent clusters resulting from the different Party World models against their spatial patterns allows us to reason as follows. Agent clusters resulting from the Party World models which produce F-formations with roundness < 0.5 last for a shorter time than agent clusters resulting from the models which produce F-formations with roundness > 0.5. Similarly, closed dyadic F-formations last longer than open dyadic F-formations. This inference is consistent with the real-world observation that conversational groups last longer when their spatial arrangement is closed instead of being open (cf. section 2.7.1 in chapter 2). In real-world situations this happens because closed arrangements are preferred in closed spaces with fewer intrusions, whereas open arrangements are preferred in open spaces with more intrusions, and thus tend to end soon. We did not model or consider the differences between closed and open spaces in the Party World models, even so, the spatio-temporal characteristics of agent clusters resulting from the Party World models is consistent with this real-world observation.

Based on our findings in the previous chapter and this chapter, it is evident that there are differences between the Party World models, in terms of the social, spatial and temporal characteristics of the agent clusters they produce. Nevertheless, we do not yet know how exactly the characteristics of agent clusters resulting from the models correspond to the characteristics of real-world conversational groups. Validation is the scientific way of determining whether a model produces results that resemble a real-world target phenomenon (cf. section 4.5.1 in chapter 4). The real-world phenomenon we are interested in is the organisation of spatial behaviour in face-to-face conversations and the spatial patterns of conversational groups. There is not enough (accessible) data on this target phenomenon for us to validate the Party World models. We therefore recorded videos of naturally occurring conversational groups in a real-world social situation and used the data to validate the models. In the next chapter, we will describe the Drinks Reception (DR) party dataset we have created for the purpose of this thesis, to validate the Party World models. In chapter 8, we will present the analysis and results of the validation study. Finally, based on the comparison of the results from the different Party World models (in the previous and current chapters) and the validation (in chapter 8) of the models, we will discuss which of the Party World models simulates the spatial patterns of conversational groups in a realistic way.

Chapter 7

Sampling the Spatial Patterns of Real-world Conversational Groups: The Drinks Reception (DR) Dataset

---

In this chapter, we describe the Drinks Reception (DR) dataset, which we created for the purpose of validating the Party World models. The dataset contains images of naturally occurring conversational groups at a drinks reception party, captured using five overhead cameras. The dataset also contains the foot position and head orientation of the people in the images. It also has the camera calibration parameters used to map the location of a person in an image to real-world coordinates. There are 944 images per camera angle and 4720 images in total. The dataset contains images from the start to the end of the party, which in turn enabled us to recreate the positional-orientational presence of participants for the entire duration of the event.

The rest of this chapter is organised as follows. In section 7.1, we begin by describing the process of filming a drinks reception party using multiple overhead cameras. In section 7.2, we describe the procedure used to convert the camera images into a format suitable for determining people's foot position and head orientation. In section 7.3, we describe our approach to annotate these. We summarise the contents of the DR dataset in section 7.4. We conclude the chapter with a discussion about using the DR dataset for validating our Party World models. We also list the advantages and limitations of the DR dataset.

7.1 Filming a Drinks Reception Party

We filmed a drinks reception party held after the distinguished seminar series conducted by the School of Electronic Engineering and Computer Science (EECS) at Queen Mary University of London (QMUL). Research, teaching and administrative staff and students from the University and external guests participated in the party.

7.1.1 The venue

The party was held at the EECS Hub (also known as the Hub) at QMUL. It's an indoor space that is roughly rectangular in shape and has an open plan kitchen. Figures 7.1 and 7.2, captured using overhead cameras, show the Hub, where apart from the kitchen area at the front of the room (with all its equipment and a table), there is some furniture at the back of the room, and a side table to one side of the room.

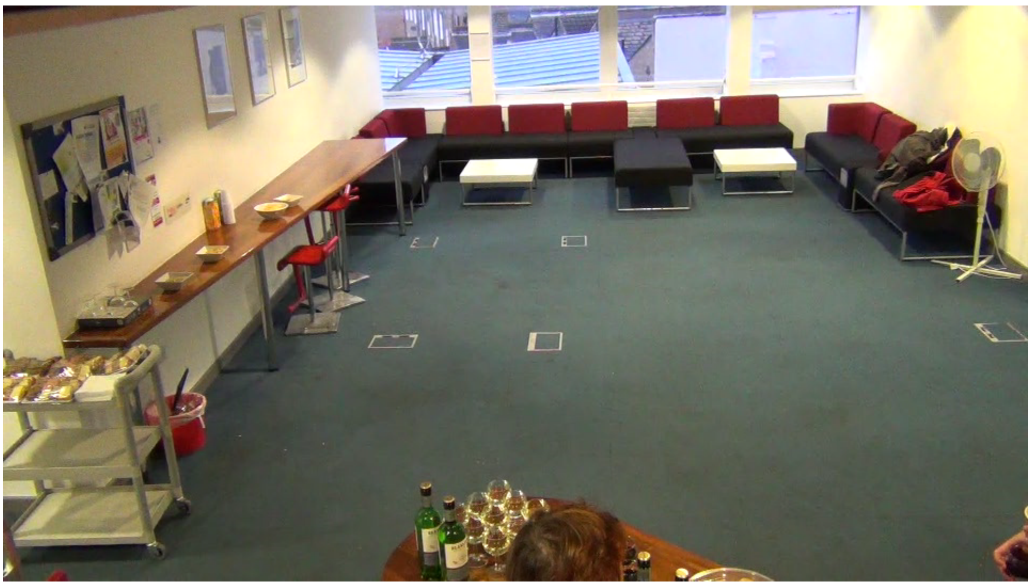


Figure 7.1: The party venue: The Hub

(a) Corridor leading to the Hub (b) Corridor leading away from the Hub

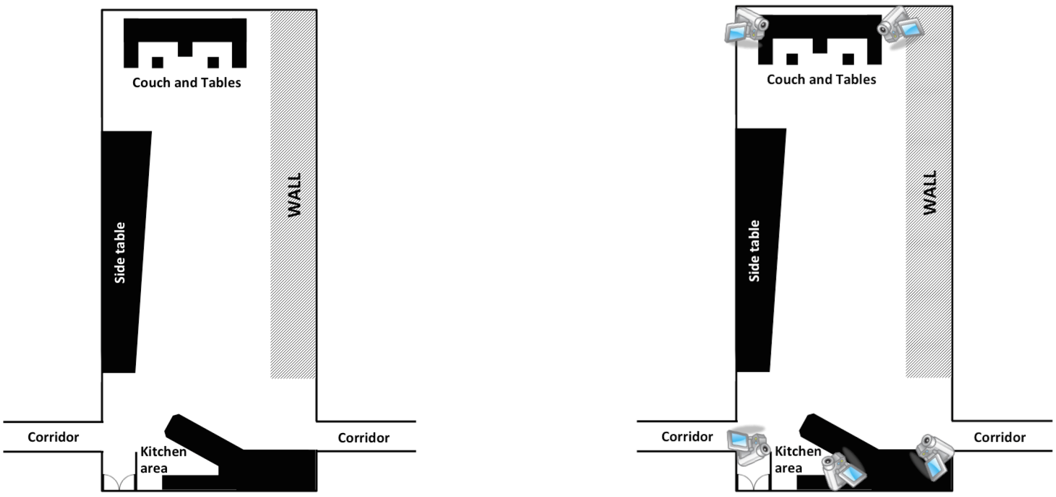
Figure 7.2: Corridors leading to and away from the Hub

7.1.2 Research ethics approval for filming

We filmed the drinks reception party after obtaining the approval of the QMUL research ethics committee (QMREC1274 - Filming Free-standing Conversations). A few days ahead of the event, we sent an email to the residents of EECS, informing them about the filming. On the day, we put up notices of "Filming in Progress" throughout the school building, along with the name and contact details of the investigator. We noted that the investigator is looking at how people use space during social gatherings, especially during face-to-face conversations. We offered to recognise and address objections to filming, but none were received. Furthermore, it appeared that people were not aware or conscious of being filmed, because the cameras were placed in an unobtrusive overhead position.

7.1.3 Equipment used for filming

The Hub is rectangular in shape and about 11.89 meters long and 7.07 metres wide. Figure 7.3a shows the floor plan of the Hub. We used five cameras to cover the entire area of the Hub. Figure 7.3b shows the placement of cameras in the Hub. We used SONY® hand-held cameras for filming (figure 7.4a). We used the clamps and mounts shown in figures 7.4b and 7.4c to fix the cameras to girders supporting the false ceiling in the Hub. The ceiling in the Hub was not high enough to obtain a proper bird's eye view of the room. Hence we mounted the cameras at a 45° downward inclined angle as shown in the figures. We borrowed all the filming equipment from the School of EECS at QMUL.



(a) Floor plan of the Hub

(b) Camera placement in the Hub

Figure 7.3: Floor plan of the Hub and camera placement

(a) Camera used for filming

(b) Clamp for holding cameras

(c) Mount used to secure cameras to ceiling

Figure 7.4: Equipment used for filming

7.1.4 Filming using multiple cameras

We fixed the cameras to the ceiling in marked positions (figure 7.3b), and turned them on manually, before the lecture ended and the participants started arriving in the Hub. We used the clapper board technique to synchronize the videos obtained from the five different cameras. The investigator stood at a spot, which was in the field of view of all five cameras, and clapped once loudly after recording started in all the cameras. The videos acquired from the individual cameras were then imported into Apple® Final Cut Pro[1] and synchronized using the built-in *sync using clapboard* feature.

The synchronization process resulted in 5 time synchronized videos of equal duration. The videos offer a partially overhead view of the Hub, because the cameras were not mounted exactly perpendicular to the floor. After setting up the equipment, i.e. within 6-7 minutes after filming started, the investigator left the Hub and returned only after the party ended to turn off the cameras and secure the videos. A couple of participants continued talking after the party ended. However, owing to camera battery and memory card restrictions, filming had to be stopped at the close of office hours at 17:00 p.m.

7.2 Post-processing of the Videos

We adapted the camera calibration algorithm (Bradski, 2000), the bird's eye view transformation algorithm (Bradski and Kaehler, 2008), and the image rectification algorithm (MATLAB and Toolbox, 2014) to convert the camera images into a format suitable for determining the foot position and head orientation of participants. In this section, we describe the step-by-step procedure used for converting the images.

7.2.1 Step 1: Still frame generation

The first step was to convert the videos into image sequences. We used an open source multimedia framework called ffmpeg[2] to convert the synchronized videos into sequences of synchronized images. We retrieved one image per camera angle once

[1] https://www.apple.com/uk/final-cut-pro/

[2] https://www.ffmpeg.org/

every five seconds of the videos, which resulted in 944 images per camera angle, and 4720 images in total. Figure 7.5 is an example of a time synchronized multi-camera image from the DR dataset.

(a) Camera 1 (b) Camera 2 (c) Camera 3

(d) Camera 4 (e) Camera 5

Figure 7.5: Example of a time synchronized multi-camera image in the DR dataset

7.2.2 Computer vision processing

We performed bird's eye view transformation on the images from the DR dataset, to render them appropriate for determining the foot position and head orientation of participants. It is a requirement to obtain a homography matrix, by performing the camera calibration process, to be able to perform bird's eye view transformation.

The camera calibration process helps in removing lens distortions effects from the images, and in relating camera measurements (in pixels) to real world measurements (e.g. measured in meters). The bird's eye view transformation helps in creating bird's eye view images of the Hub (and the people in it), corresponding to the partially overhead views achieved using the overhead camera placement (cf. section 7.1.3). The third stage is the image rectification process, for registering the bird's eye view images of the Hub onto a floor plan of the Hub. We elaborate each step in the following sub-sections.

7.2.2.1 Camera calibration

Our aim was to determine the foot position and head orientation of participants in the images in the DR dataset. To do this, we had to first achieve an exact mapping between the three-dimensional real world coordinate system, and the two-dimensional image coordinate system (rendered by the camera). Camera Calibration is a process that makes this mapping possible, by mathematically correcting for the distortions introduced by the camera lens (Bradski and Kaehler, 2008, p. 370).

To begin with, we captured a series of images of a regular pattern, i.e. a chessboard, from various angles and distances, using the same camera that was used for filming (the SONY® hand-held camera in figure 7.4). Figure 7.6 provides an overview of this step; notice the chessboard being held at different angles and distances. These images were then fed to a collection of readily available functions in OpenCV[3]. The outcomes were a set of values: the *camera intrinsics matrix*, the *lens distortion coefficients* and a *homography matrix*. The first two values help in correcting the still frames for distortions introduced by the lens used in the camera. The homography matrix relates positions of the points on an object plane (in the real world) to corresponding points on an image plane (rendered by the camera). We used the OpenCV guide to perform the camera calibration (Bradski, 2000).

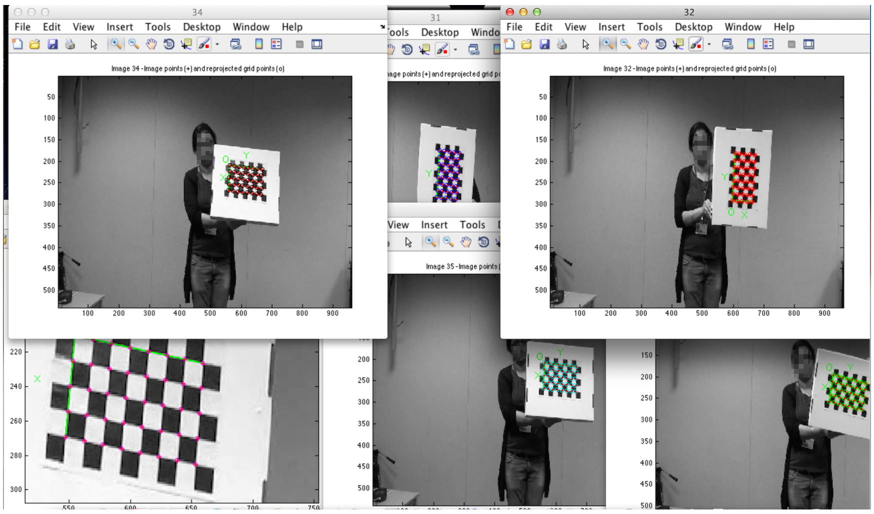
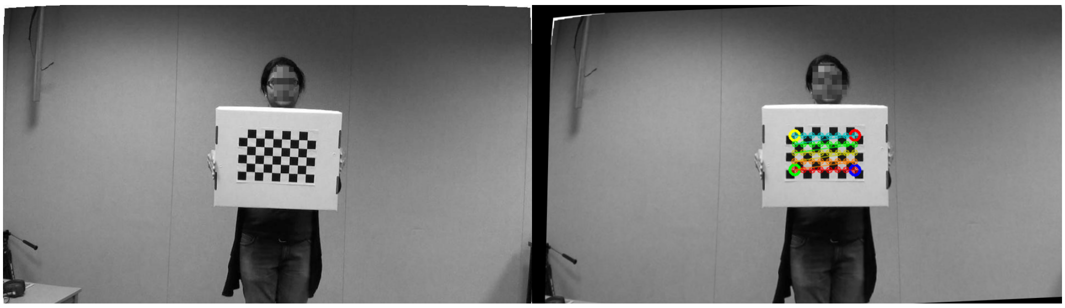


Figure 7.6: Overview of the camera calibration process

[3]OpenCV is an open source computer vision library. See www.opencv.org

7.2.2.2 Bird's eye view projection

After the camera calibration, we performed the bird's eye view transformation, which is a process used to obtain the top view perspective of an image. We adapted Bradski and Kaehler's (2008, p. 409) model for performing the bird's eye view transformation. The model takes the homography matrix resulting from the camera calibration process, and an undistorted camera image obtained by correcting for lens distortions, using the camera intrinsics matrix and lens distortion parameters, also obtained from the camera calibration process (see figure 7.7a). The output of the bird's eye transformation of the undistorted image is shown in figure 7.7b. Because filming was done from an almost overhead view, there are no significant changes in the image after undergoing bird's eye view transformation. We generated the bird's eye view of each image in the DR dataset, e.g. figure 7.8a is an image from the DR dataset and 7.8b is its bird's eye view.



(a) An undistorted chessboard image (b) Bird's-eye view transformation of 7.7a

Figure 7.7: Example of an undistorted image and bird's-eye view transformation

(a) An image from the DR dataset (b) Bird's-eye view transformation of 7.8a

Figure 7.8: Demo of transformation of an image from the DR dataset

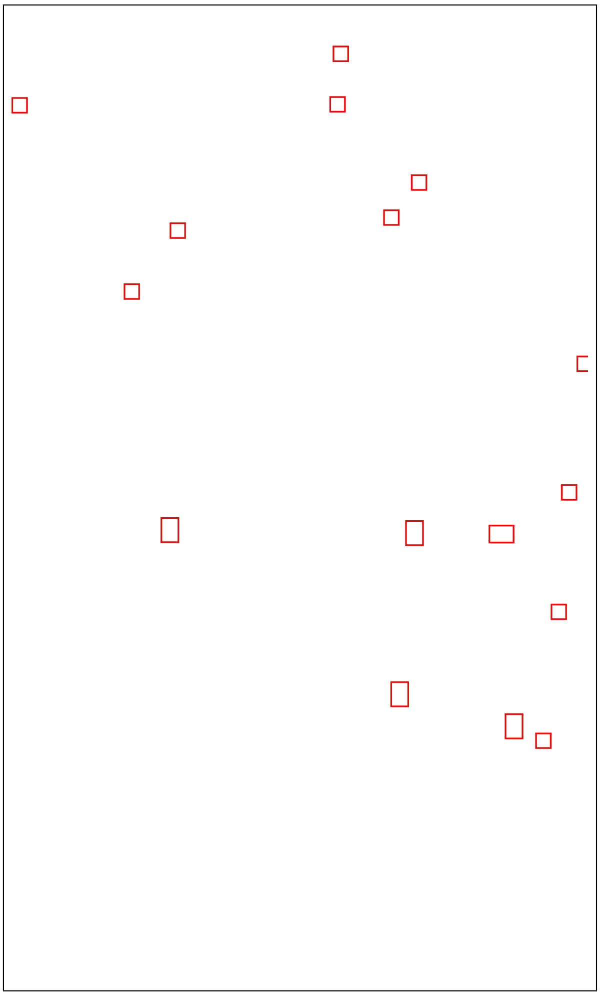
7.2.2.3 Image registration

Image registration is the final stage of computer vision processing we did to transform the images in the DR dataset. The process registers bird's eye view images (obtained from the previous step) onto a floor plan of the Hub. Doing this transforms one image (i.e. bird's eye view) into the coordinate system of another image (i.e. the floor plan of the Hub). In other words, image rectification allows a projection between two planes: the floor in each image and the floor plan of the Hub.

We followed the following steps for the image registration process. Firstly, we created a floor plan of the Hub to an exact $1/10^{\text{th}}$ scale of the width x height dimensions of the Hub. Secondly, we replicated the obvious corners, edges, and geometrically regular shaped structures visible on the floor of the Hub (e.g. see figure 7.9a), also to a $1/10^{\text{th}}$ scale on the floor plan of the Hub. These are the red-coloured squares and rectangles in figure 7.9b. Next, the bird's eye view images from the DR dataset and the floor plan of the Hub are both imported into MATLAB, using a script that allows selecting corresponding pairs of control points in the two images. Figure 7.10 shows the control points selection tool. With the chosen set of control points, the script then performs a projective transformation, to bring the two images into alignment. Figure 7.11 is a bird's eye view image from the DR dataset, and figure 7.12 shows the bird's eye view image registered onto a floor plan of the Hub.

The image registration process leads to an outcome where selecting any (x,y) position in the image, corresponds to an exact physical location (X,Y) in the actual floor space of the Hub, such that X= kx and Y = ky, where k = 10, i.e. the scale adopted in drawing the floor plan of the Hub. Therefore, at this stage, clicking on an individual's foot position (i.e. an x,y position) in an aligned image yields a scaled version of the exact physical location (i.e. the X,Y position), where the individual was standing in the Hub during the party. In other words, the process enables obtaining an individual's foot position with respect to the ground plane, in terms of the real world coordinate system.

(a) Spatial structures in the floor of the Hub



(b) Spatial structures replicated in the floor plan of the Hub

Figure 7.9: Recreating the spatial structures in the floor of the Hub

Figure 7.10: Image registration: Control point selection tool

Figure 7.11: A transformed image from the DR dataset

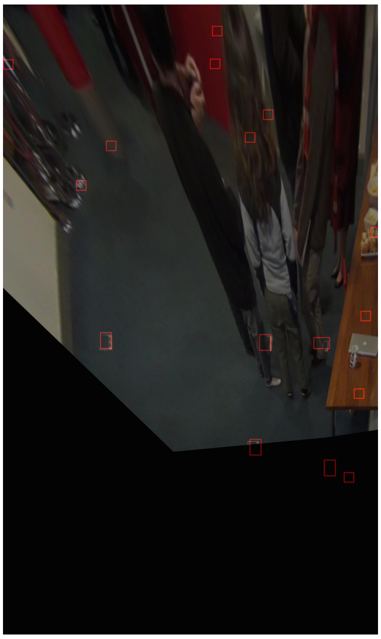


Figure 7.12: Bird's-eye view image aligned with the floor plan of the Hub

7.3 Manual Annotation of the Images

In this section, we describe the process used to determine the foot position and head orientation of participants, in the rectified images in the DR dataset. Annotation was done manually by the investigator for 21 participants. The open plan kitchen in the Hub occasionally attracted people other than participants during the party. These people walked in and out of the Hub to use the kitchen. We did not annotate the position and orientation of people who only used the kitchen, or walked through the Hub, without participating in the party. We also did not consider people who were in the Hub for less than 5 minutes interacting with one or more of the participants. Annotating the foot position and head orientation of the 21 participants in the 4720 (i.e. 944 images per camera angle x 5) images took about 100 hours. It must be noted that not all 21 participants featured in all the images. Some of them were in the Hub for a shorter time than others, and none of them were there for the entire duration (i.e. approximately 78 minutes).

7.3.1 Annotation of foot position

We built an interface in MATLAB, which allows an annotator to click on the foot of participants in the rectified images (from the DR dataset), and record the (x,y) coordinates of the mouse click. The images aligned to the floor plan of the Hub from the DR dataset are displayed in the left pane of the interface. Clicking on the foot of a person in the images records the (x,y) coordinates of the mouse click. As noted before, since the images are aligned to the floor plan of the Hub drawn to an exact scale, the (x,y) coordinates of the mouse click correspond to where participants were actually standing on the floor of the Hub during the party. The rectified images are not as clear as the camera images, and so, the camera image corresponding to a rectified image was displayed in the right pane of the interface, as a guide to the annotator, i.e. to improve the annotator's confidence while clicking on the foot of participants. Figure 7.13 shows the interface.

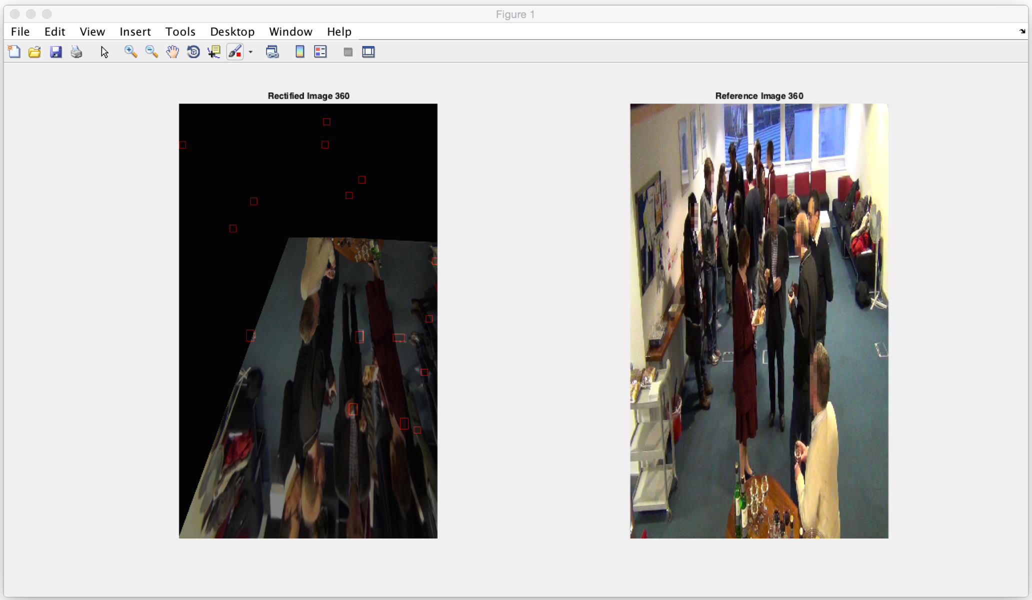


Figure 7.13: Interface for recording the foot position of participants

The annotation of foot position was performed in the following way. For every camera angle, the entire sequence of aligned and original image pairs were loaded in order, one pair at a time, allowing the annotator to record the foot position of every individual participant whose feet were visible in the images. Every participant was assigned a unique ID – we used participant's initials or pseudonyms to create the unique ID for each participant. A mouse click on the foot of a particular participant in the image triggered a prompt that allowed the annotator to register the unique ID of the participant. The (x,y) location was automatically populated based on registering the mouse click in the image. The image number was also populated automatically while transitioning between images. Clicking *OK* in the prompt recorded the data in a spreadsheet, whereas clicking *Cancel* allowed the annotator to redo the selection. Figure 7.14 shows the prompt.

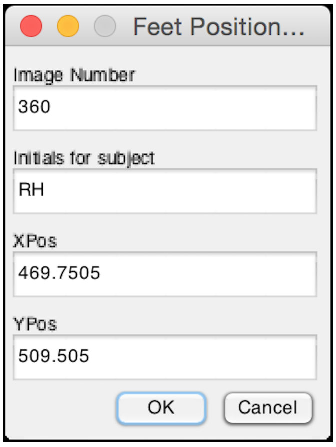


Figure 7.14: Prompt for registering data about foot position

We calculated the weighted average of the (x,y) values for an individual who was captured in more than one camera at the same time. We found that not all participants were captured in all cameras at every time step. Even if they were, their feet may be cut off or occluded from an image. However, any particular participant was likely to have been captured by more than one camera at any one time. Therefore, for every image, we calculated the weighted average of the (x,y) coordinates of the mouse clicks corresponding to a particular participant, across cameras. Table 7.1 shows an example of the weighted average calculation for the foot position of a participant identified as RH in image 286. The value (0,0) in the table denotes that RH was not captured by Cam4 (or Camera 4) or that RH's feet were not seen in image 286 obtained from Cam4.

Image	Participant	Value	Cam1	Cam2	Cam3	Cam4	Cam5	Average
286	RH	x	599.08	577.16	561.12	0	588.12	581.37
		y	393.31	388.93	394.41	0	426.20	400.71

Table 7.1: Example: Determining the weighted average of a participant's foot position

The weighted average calculation ensured recording the best estimate of an individual's foot position, by reducing the uncertainty related to: (1) clicking the mouse on slightly different locations in the images from the different cameras, (2) not being sure where to click on the feet. For instance, figure 7.15 shows examples of feet placement in the images from the DR dataset. For the placement seen in figure 7.15a, we clicked roughly at the centre of the left foot and the right foot. For other feet positions, such as the ones in figures 7.15b and 7.15c, we clicked on one of the spots indicated in the figures. Calculating the weighted average helped minimize the discrepancies across different camera angles and patterns of feet placement.

(a) Feet placement 1 (b) Feet placement 2 (c) Feet placement 3

Figure 7.15: Examples of mouse clicks corresponding to feet placement

7.3.2 Annotation of head orientation

We developed a MATLAB interface, similar to the one used for the annotation of foot position, for the classification of participants' head orientation under one of the following eight categories: (i) North, (ii) South, (iii) East, (iv) West, (v) North East, (vi) North West, (vii) South East and (viii) South West. This multi-class manual labelling convention is inspired by existing work on training automatic head orientation detectors (Tosato et al., 2013).

The MATLAB interface used is shown in figure 7.16. Unlike the processed images used for the annotation of foot position, the images originally acquired from the cameras were used for the annotation of head orientation. A red-coloured square, encompassing a participant's head, appears when a mouse click is made roughly at the center of the head. The square aims to single out a participant's head from the rest of their body, and is intended only as a visual reference, for making better decisions about head orientation classification. A prompt also appears along with the bounding box, which allows recording the participant ID and an appropriate value for their head orientation, under one of the above eight categories (see prompt in figure 7.16).

Figure 7.16: MATLAB interface for manual annotation of head orientation

The classification of head orientation under the eight categories is based on assuming a particular wall of the Hub as a frame of reference (see figure 7.17). More specifically, facing towards the frame of reference is assumed as North, and the right

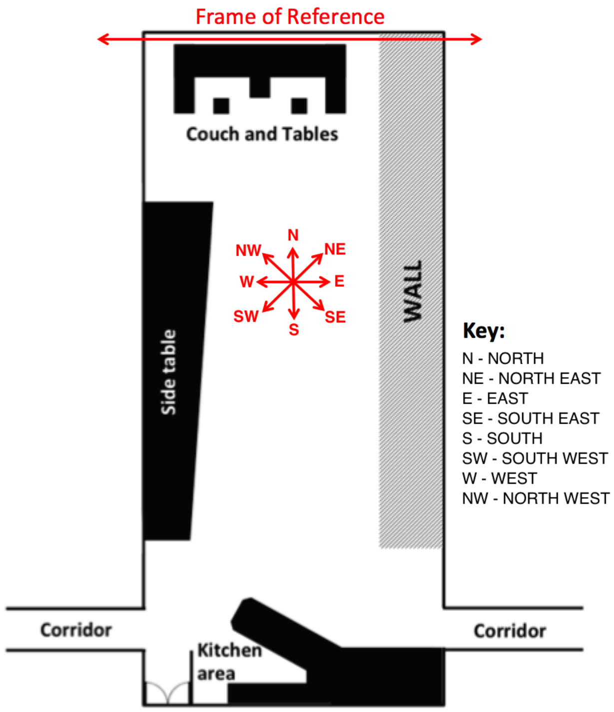


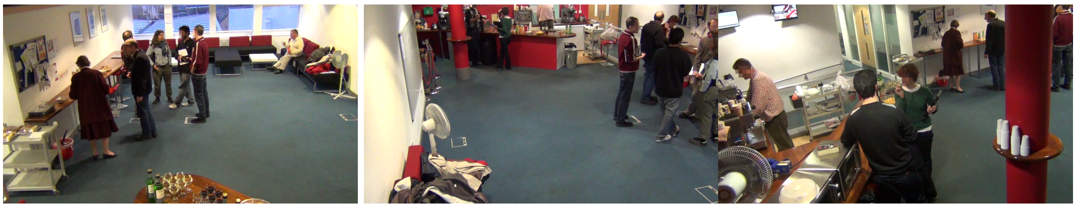
Figure 7.17: A fixed frame of reference in the Hub

opposite is assumed as South. Starting from North, North West, West, South West, South, South East, East and North East orientations are spaced at 45° intervals along the anti-clockwise direction.

7.4 The DR Dataset

At the end of the image processing and annotation tasks, the DR dataset contained images showing naturally occurring conversational groups at a party, as well as a spreadsheet containing the foot position and head orientation of the people in the images. More specifically, the dataset contained: (i) Original camera images – one folder for each camera angle containing a sequence of 944 images captured when filming; (ii) Computer-processed images – one folder for each camera angle containing a sequence of 944 floor aligned images, (iii) a spreadsheet summarising the foot position (with respect to the ground plane) and head orientation (with respect to the chosen frame of reference) of 21 participants across the 944 images, and lastly, (iv) the camera calibration parameters.

Figure 7.18 shows a set of time synchronized multi-camera images from the DR



(a) Image from camera 1 (b) Image from camera 2 (c) Image from camera 3

(d) Image from camera 4 (e) Image from camera 5

Figure 7.18: A set of time synchronized multi-camera images from the DR dataset

dataset. Using the position and orientation of each participant in these images, we synthesised a Party World type of representation of the Hub, see figure 7.19. When the images shown in figure 7.18 were captured, there were 9 participants in the Hub, each of whom are identified with a unique participant ID in the DR dataset. The reconstituted Party World type replication of figure 7.18 shows the nine participants, represented like agents in the Party World simulations, bearing their unique IDs on the head. The actual methods used for comparing the spatial, temporal and social characteristics of human conversational groups captured in the DR dataset, and the agent clusters resulting from the Party World simulations, will be discussed in the next chapter.

7.5 Discussion

In this chapter, we introduced the DR dataset, containing images of conversational groups at a drinks reception party, and information regarding the position and orientation of the people at the party. We described the filming process undertaken to collect the images for the dataset, the techniques used for post-processing the images, and lastly, the methods used to annotate the foot position and head orientation of participants.

The advantage of the DR dataset is three-fold. First is the creation of a dataset

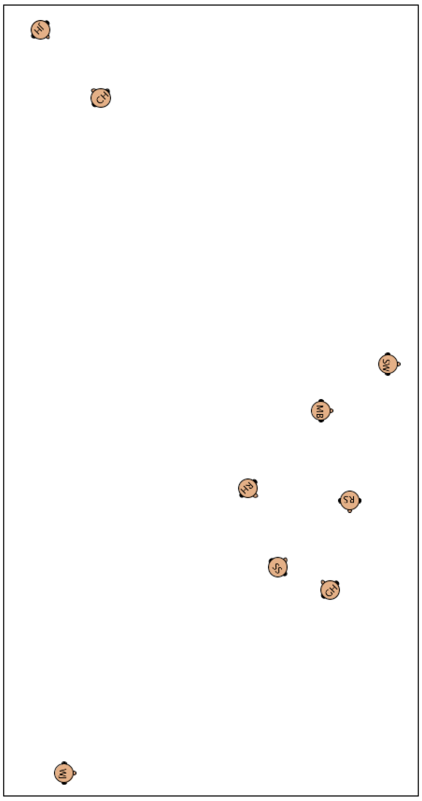


Figure 7.19: A reconstituted Party World type image of the Party at the Hub

capturing the occurrence of naturally occurring conversational groups. We could not find a dataset that had as many images and as many participants as the DR dataset. More importantly, we could not find a dataset that captured the positional-orientational presence of participants for the entire duration of a party. For example, the Coffee Break dataset[4], which was accessible to us, had only 75 or fewer images. Furthermore, capturing conversational groups at a party, resulted in an outcome comparable to the Party World simulations, which are also based on the party context.

The second advantage is the multi-camera filming technique used to create the DR dataset. The images from the five cameras, placed at different spots in the Hub, ensured greater coverage of the space, thereby minimising instances of participants being cut off or occluded. The multi-camera images also enabled determining the weighted average of the foot position of participants. The third advantage of the

[4]https://profs.sci.univr.it/ cristanm/datasets/CoffeeBreak/index.html

DR dataset is the detailed specification of the step-by-step procedure used to create the dataset.

On the other hand, it may be argued that using sensors like motion capture systems (MoCap) or accelerometers, instead of traditional video cameras, would have resulted in more accurate measurements of foot position and head orientation in the DR dataset. We could have also avoided the detailed computer vision-based processing and manual annotation of images, which was expensive in terms of researcher time. However, we faced some practical difficulties when we tried to use MoCap equipment for the process. (1) Setting up the MoCap equipment, especially to cover the space used by more than twenty participants at a time, was difficult. (2) Requesting 20 to 30 participants to suit up with markers for the MoCap process proved challenging. Not to mention the suit and markers constrained participants' movements. Furthermore, the annotator had to make up for any missing marker data, e.g. when the markers fall off when participants transition between capture zones. Using the Microsoft Kinect sensor[5] is a less intrusive alternative to MoCap. Nonetheless, capturing the position of about 20 people simultaneously, required a sophisticated network of Kinect sensors, which in turn posed other constraints.

The quality of our manual annotation, performed only by the investigator, may also be questioned. The weighted average strategy we used for determining the foot position of participants offers a certain level of consistency in this regard, however, it is possible to question our approach to label head orientations. As noted in section 7.3.2, the strategy we used for labelling head orientations is inspired by existing methods, used for developing automatic head orientation detectors. In these works, the head orientation of participants were labelled manually in a fraction of the dataset, as a way of training models to perform automatic multi-class head orientation classification, using machine learning methods. Most of the existing detectors classify head orientations under four (front, back, left and right), or six (front, back, left, right, front back and front left) categories, due to the inherent difficulties in deducing back-facing head orientations from images. We could have

[5]http://www.microsoft.com/en-us/kinectforwindows/

used these automatic detectors in our dataset to ensure consistency, but using our multi-camera approach and the static frame of reference, we were able to classify head orientations under eight different categories, resulting in more fine-grained information.

In the next chapter, we will use the foot position and head orientation of participants in the DR dataset, to analyse the social, spatial and temporal characteristics of the conversational groups at the drinks reception party. We will then compare our findings with the results of the Party World simulations to identify if, and which, of the Party World models result in agent clusters that correspond more closely to reality.

Chapter 8

Validating Party World Models using the DR dataset

The process of comparing the results of a simulation with observations of a real-world phenomenon, which is the target of the simulation, is referred to as validation in agent-based modelling (Gilbert, 2008). Validation confirms if a model correctly simulates its target. For Party World simulations the target is conversational groups formed in the real world. In this chapter, we validate the Party World models by comparing the social, spatial and temporal characteristics of agent clusters resulting from the Party World simulations with the characteristics of conversational groups captured in the DR dataset. Based on our validation study, we identify if, and which, of the Party World models generate agent clusters that most closely resemble real-world conversational groups.

The rest of this chapter is organised as follows. In section 8.1, we begin by analysing the number and size of F-formations in the DR dataset – we compare the measures with the ones resulting from the Party World simulations, to identify the similarities and differences. In section 8.2, we analyse the different types of dyadic and multi-party arrangements of conversational groups in the DR dataset and compare it with the outcomes of the Party World models, to assess the similarities and differences. In section 8.3, we present a qualitative analysis of the coexistence and evolution of the spatial-orientational arrangements of conversational groups in the DR dataset, comparing it with the social, spatial and temporal dynamics of agent clusters resulting from the different Party World models. We end the chapter by providing a summary of the key findings.

8.1 Analysis of the Number and Size of F-formations in the DR dataset

In this section, we analyse the number and size of F-formations detected in the DR dataset, and then compare the results with those obtained from the Party World simulations. While performing the analysis, we take into consideration the limitations of the simulated Party World, when compared to the drinks reception party filmed at the Hub (cf. chapter 7).

8.1.1 Adapting the F-formation detection methods for the DR dataset

We used the HVFF, GCFF, and PWFD methods to detect the F-formations in the DR dataset. The foot position and head orientation of each of the 21 participants in the images were given as input to detect the F-formations. The HVFF and GCFF methods also required two other input parameters to identify the F-formations in the images – one of them was used as a reference for screening interpersonal distances and the other was used for screening the relative orientation of participants.

Determining the values for the reference parameters was straightforward when dealing with the Party World dataset. We used the predetermined separation distance used in the Party World simulations, i.e. 45 pixels, as the reference value for screening interpersonal distances. Based on the advice of the developers of the HVFF and GCFF methods, we used 60° as the reference value for screening the relative orientation of participants. The HVFF and GCFF methods worked with values around the reference range to identify the F-formations in the Party World dataset.

For the DR dataset, we ran each of the automatic F-formation detection methods using different values for the screening parameters. The values were selected on a trial and error basis. We selected a few images from the DR dataset to do a trial run of the HVFF and GCFF methods with different values for the reference parameters. Values that did not allow a reliable detection of F-formations were not used, e.g. some lower values did not detect any valid F-formations, while some higher values detected two separate groups as one big F-formation. Consequently, reference values that did not produce a reasonable estimation of F-formations were not used for the actual detection.

For HVFF, the value of *radius* is a key input parameter, and it is used to screen the interpersonal distances between individuals. We ran HVFF with two different radius values: 45 HVFF(1) and 90 HVFF(2). The parameter *stride* in the GCFF model is the equivalent of radius in the HVFF model. We ran GCFF with two different stride values: 20 GCFF(1) and 40 GCFF(2). For PWFD, the interpersonal distance (d) and the angle of relative orientation ($\theta_{screening}$), are the key screening parameters. We ran PWFD with three different value combinations – PWFD(1) (d=45, $\theta_{screening}$= 60), PWFD(2) (d=45, $\theta_{screening}$= 90), and PWFD(3) (d=60, $\theta_{screening}$= 90).

8.1.2 Analysis based on the number of F-formations

The number of F-formations detected by the HVFF, GCFF and PWFD methods using different values for the reference parameters are shown in figure 8.1. As seen in the figure, all three methods detected more than 2000 F-formations in the 944 images of the DR dataset. In chapter 5, we found that most Party World models produced around 2000 F-formations in 360 images (cf. figure 5.6), whereas in the DR dataset, the total number of F-formations detected (in all 944 images in the dataset) is around the 2000 mark. This may appear to suggest that the Party World simulations generate three times as many F-formations compared to real life. However, there is an important difference between the Party World models and the DR dataset, which may have caused this difference.

The Party World simulations were run with 20 agents, and coincidentally, the DR dataset includes 21 participants. Therefore, the number of agents in the Party World

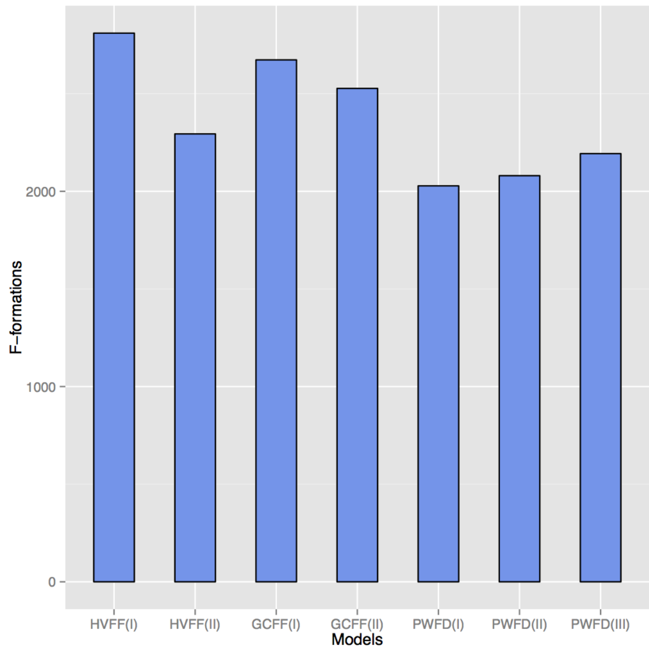


Figure 8.1: F-formations detected in the DR dataset

simulations and the number of participants in the DR dataset is almost the same. However, the 20 agents in the Party World simulations were always included in the simulations, whereas in the drinks reception party, we found that participants were not present in the Hub for the entire duration of the party. Participants arrived and left at different times in the drinks reception party. Due to this difference between the Party World simulations and the drinks reception party, it is not possible to conclude that the simulations produce three times as many F-formations compared to real life.

We found that 7 people per image was the average in the DR dataset, i.e. sum of number of people in each image (6543) divided by the number of images in the DR dataset (944). Since this is approximately three times less than the number of agents in the Party World dataset, we might argue that the estimates of F-formations in the DR dataset and the Party World dataset are roughly equivalent. However, in future extensions to the Party World models, we propose adding rules that would allow agents to arrive and depart at different times as in real life. The idea is developed further in the rest of the chapter.

Furthermore, in chapter 5, we found that the HVFF and GCFF methods detected more F-formations than the PWFD method in the Party World dataset (cf. figure 5.6), and a similar trend is observed here in figure 8.1: HVFF (1 and 2) and GCFF (1 and 2) detected more F-formations than PWFD (1, 2 and 3). We found that HVFF and GCFF detected more F-formations when the radius and stride values were lower (45 and 20 respectively), while they detected fewer F-formations when the radius and stride values were higher (90 and 40 respectively). On the other hand, PWFD detected more F-formations when d and $\theta_{screening}$ values were higher, and fewer F-formations when the values were lower. In section 8.1.4, we explain how the screening parameters affect the performance of the automatic F-formation detection methods, and argue why it is important to choose appropriate values for input parameters in models.

8.1.3 Analysis based on the size of F-formations

Figure 8.2 shows the size of F-formations detected in the DR dataset. PWFD (3) detected the most 7-member F-formations, followed by PWFD(2) and GCFF(40). All the other methods detected F-formations consisting of between two and six members.

Figure 8.2 also shows that there are lot more dyadic and triadic F-formations in the DR dataset than F-formations of 4 or more people. In chapter 5, we found that the size of F-formations resulting from the Party World simulations showed a similar trend, i.e. more dyadic and triadic F-formations than bigger groups (cf. figures 5.11a, 5.11b and 5.11c).

In chapter 5, we also found that Models 1, 2a LA and 2b LA generate many F-formations with 2 to 6 or 7 members, Model 2a EA generates F-formations with 2 to 8 or 9 members, and Model 2b EA generates F-formations of even up to 13 or 14 members (cf. section 5.3.2). However, again, since the entry and exit characteristics of people in the DR dataset did not match with the assumptions of the Party World simulations, we cannot draw a one-to-one comparison between the size of F-formations in the DR dataset and the Party World dataset.

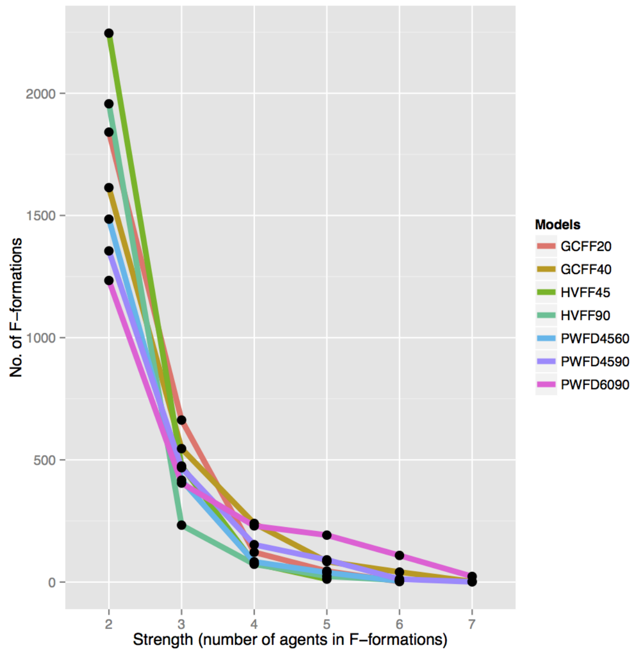
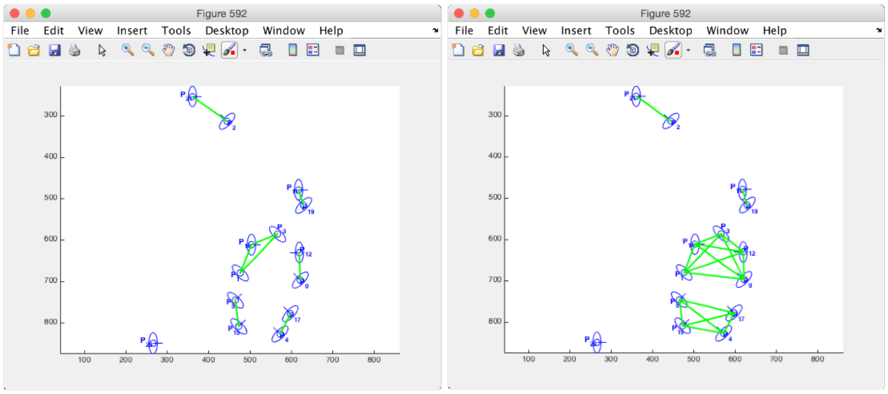


Figure 8.2: The size of F-formations in the DR dataset

8.1.4 Comparing the results of the different F-formation detection methods

In this section, we compare the performance of the HVFF, GCFF and PWFD methods. As evident in figure 8.1, HVFF(1) detected more F-formations than HVFF(2). This is because the former detected many groups consisting of fewer agents, while the latter detected fewer groups comprised of many agents. Similarly, GCFF(1) detected more F-formations with fewer members, while GCFF(2) detected fewer F-formations, but they were bigger groups with more members. Generally, when using the HVFF and GCFF methods, increasing the value of the key input parameters (i.e. radius and stride) leads to detecting bigger F-formations. Increasing the value of the key input parameters in PWFD, d and $\theta_{screening}$, also leads to detecting bigger F-formations – PWFD(3) detected bigger F-formations than the ones detected by PWFD(1) and PWFD(2).

The differences in the performance of the different versions of the F-formation detection methods can be illustrated with the following example. In the DR dataset, there was one big conversational group, which was established at the start of the party (i.e. when guests started arriving in the Hub after the seminar), and lasted almost until the end of filming. Figure 8.3 shows the F-formations and their members detected in image 598 of the DR dataset by GCFF model I (stride=20) and GCFF model II (stride=40) respectively. As seen in the figure, GCFF(1) detected more F-formations that are smaller in size, whereas GCFF(2) detected comparatively fewer but bigger F-formations. Similar differences were evident in the performance of the HVFF and PWFD methods.



Stride = 20 **Stride = 40**

Figure 8.3: F-formations in image 598 in the DR dataset detected by the GCFF(1) and GCFF(2) methods

The differences in the performance of the different F-formation detection methods leads to the following reasoning: although the methods identify potential F-formations, their performance is not definitive. There are variations in the assessments of F-formations, especially when detecting groups of four or more people, but F-formations of 2 to 3 people are correctly detected by all methods. The failure to correctly detect bigger F-formations could be attributable to two reasons. Firstly, the influence of the input parameters (radius, stride, d and $\theta_{screening}$) in identifying the F-formation centres (in case of HVFF and GCFF) or the overlap of transactional segments (in PWFD). Evidently, providing higher values for the input parameters allows the methods to identify bigger F-formations. At the same time, we also found that if values are increased beyond a certain threshold, then two or three nearby dyadic F-formations were detected as one big F-formation.

We consider the second reason for the failure to detect bigger F-formations as a more fundamental one. We found, in Chapter 2, that Hall's (1966) conceptualisation of interpersonal distances and Kendon's (1990) concept of overlapping transactional segments were based on observations of smaller conversational groups. The understanding has been extended by the HVFF, GCFF and PWFD methods to detect bigger groups. Hence, the question arises if the bigger groups operate differently, due to the practical limitations associated with several people occupying a relatively small area, and coordinating their spatial-orientational behaviour within the available space? Perhaps then the personal-social distances and the 45° transactional segment overlap have to be compromised in view of accommodating all members within a reasonable visual-auditory range. For example, in bigger groups, some people will be closer to each other than to others. Likewise, if arrangements are not properly circular, some people will be closer to the group centre than the others. However, when based on strictly defined parameters, such as the values of radius, stride, d and $\theta_{screening}$, the F-formation detection methods cannot consider such practical limitations and, hence, may fail to detect bigger F-formations.

Since computational modelling, video-based ethnography, and computer vision are three different disciplines, there are no prior examples of integrated approaches to link models of F-formation simulation with models of F-formation detection, and compare them with real-life conversational groups. But here we have shown that the F-formations identified in the DR dataset varies depending upon the method used for detecting F-formations, and the values used for the screening parameters.

Even if the values for screening parameters are based on empirical evidence, e.g. Hall (1966) distance zones, it has to be recognised that real-world measures cannot be easily translated to virtual equivalents. That is to say, the meaningful distances and relative orientations for face-to-face conversations could vary between conversational groups occurring in the real world, the ones analysed computationally (as we have done here), and the ones being simulated (as in the Party World models). One way of estimating reasonable values for the input parameters is to run the simulation and/or detection methods several times, with different value combinations (i.e. to experiment with wider ranges of values than we have done here), to identify the parameter combination(s) that produce the most realistic outcomes. The Party World simulation platform and models can be used as test beds for this purpose.

Thus, our observations are that: (1) the simulation and detection of F-formations are both highly influenced by the values chosen for the spacing and orientation of interactants. (2) The values of interpersonal spacing and mutual orientation applicable to real-life conversational groups (as established in existing theories of human interaction) may not be translatable to virtual environments. The latter may be true particularly when dealing with groups of four or more members. Our observation does not agree with the existing evidence that the interpersonal distances preferred for face-to-face conversations are the same in real and virtual worlds (cf. discussion in section 3.6 of chapter 3). The disagreement may be due to the fact that most existing models dealt with only groups of four to six members, and hence did not observe the lapses with bigger groups. In summary, based on our observations, we believe that systematic analysis, of the sort reported here, is required to ascertain what ranges of values provide the most realistic simulations and detections of F-formations under different conditions.

8.2 Analysis of the Spatial Patterns of Conversational Groups

In this section, we present the results obtained by running the dyadic pattern detection module and the Jarvis algorithm-based roundness detection module on the F-formations detected in the DR dataset, to identify and classify the shapes of the F-formations in the dataset.

8.2.1 Analysis of Dyadic spatial patterns

In chapter 6, we found that the centroid-based models (Model 2b LA and Model 2b EA) mostly generate H-shaped dyadic arrangements. The baseline (Model 1) generates a lot of queue arrangements. The field-of-view models (Model 2a LA and Model 2a EA) generate a combination of H, V, C, N and L-shaped arrangements. We also found that none of the Party World models generate more than a few I-shaped arrangements.

In the DR dataset, we expected to find all of Ciolek and Kendon's (1980) six categories of dyadic arrangements, but fewer or no queue arrangements. Figure 8.4 shows the percentage of F-formations under the different categories of dyadic arrangements, detected using the PWFD method in the DR dataset. We found that a majority of dyadic F-formations were V-shaped, followed by the L, H, C, N and I-shaped arrangements. There were less than 2% of queue arrangements. For the purpose of comparison, we have replotted the dyadic shapes resulting from the Party World models (cf. figure 6.3) here in figures 8.5, 8.6 and 8.7, by combining queue and other queue as queue arrangements.

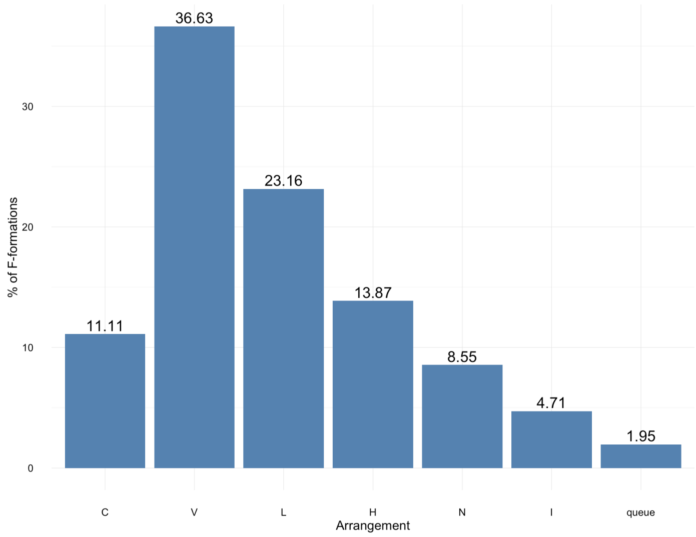


Figure 8.4: Classification of dyadic arrangements in the DR dataset

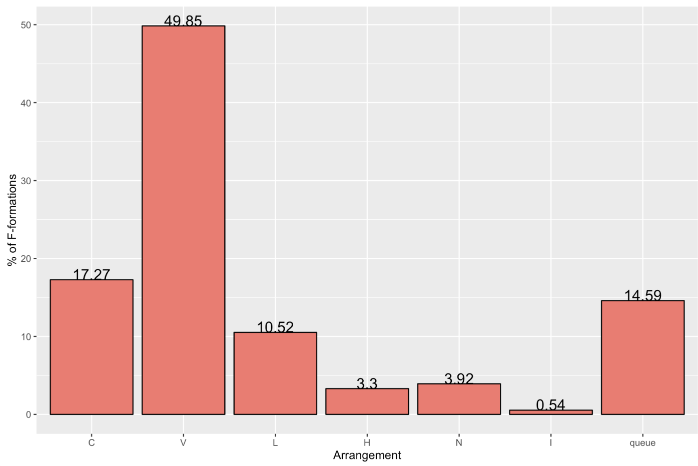


Figure 8.5: Classification of dyadic arrangements in Model 1

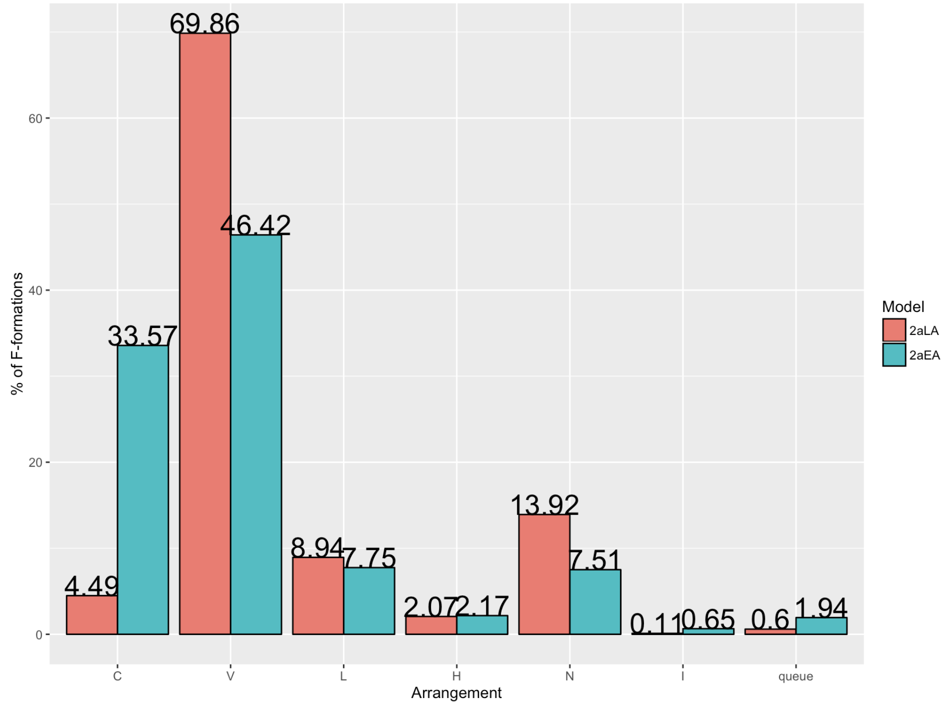


Figure 8.6: Classification of dyadic arrangements in Model 2a LA and Model 2a EA

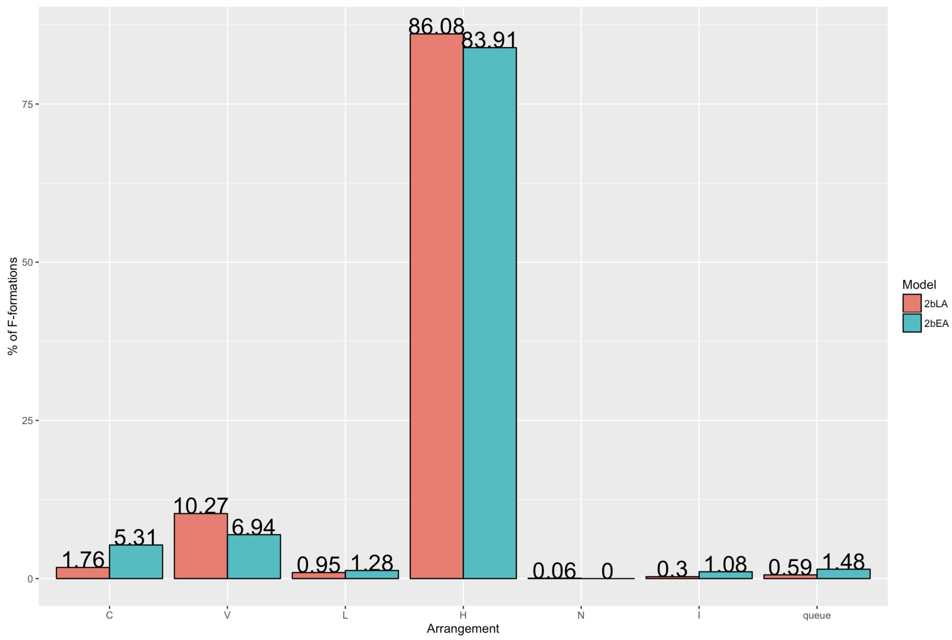


Figure 8.7: Classification of dyadic arrangements in Model 2b LA and Model 2b EA

We performed a chi-square test of independence to examine the relation between the Party World dataset and the DR dataset and the shapes of dyadic F-formations in the datasets. The relation between these variables was significant ($p < 0.001$). A z-test of column proportions further determined that Model 1 produced a significantly higher percentage of queue formations than the DR dataset, and the centroid-based models produced a significantly higher percentage of H-arrangements than the DR dataset.

The z-test also showed that DR dataset had a significantly higher percentage of I-shaped arrangements than the Party World models. As pointed out in section 6.5, in future Party World simulations, we could model an external event/object with an attractive force to check if it improves the emergence of I-shaped arrangements. The proportion of L, C, N and V-shaped arrangements in the field-of-view models were statistically different to those of the DR dataset. Nonetheless, compared to the baseline and centroid-based models, the field-of-view models had a lot more V, C, L and N-shaped arrangements and fewer queue arrangements like the DR dataset. Based on this reasoning, we argue that the field-of-view models (Model 2a LA and Model 2a EA) are better than the baseline and centroid-based models to simulate dyadic F-formations. The results of the z-test analysis are shown in figure A.4 in appendix chapter A.

8.2.2 Analysis of Multi-party spatial patterns

In chapter 6, we found that the extended neighbourhood influence in the centroid-based model (Model 2b EA) produces a considerable number of F-formations with roundness > 0.7, i.e. very round F-formations. There were at least 40% of F-formations with roundness > 0.7 in the Model 2b EA dataset. On the other hand, at least 25% of F-formations in Model 1 had roundness < 0.3 (least circular shapes). The common link between the models was that the highest percentage of F-formations they produced had roundness in the region of [0.5-0.6].

In the DR dataset, we expected to find multi-party F-formations with different degrees of roundness, but fewer F-formations with roundness < 0.3 (least round), or > 0.7 (very round). Figure 8.8 shows the frequency of the different types of multi-party F-formations (classified based on roundness) in the DR dataset. The F-formations were detected using the roundness detection module on the outcomes of PWFD(1), where d = 45 and $\theta_{screening}$ = 60 (see section 8.1). Figure 8.9 shows the frequency of the different types of multi-party F-formations detected using PWFD(2), where d = 45 and $\theta_{screening}$ = 90. Lastly, figure 8.10 shows the frequency of the different types of multi-party F-formations detected using PWFD(3), where d = 60 and $\theta_{screening}$ = 90.

The overall trend observed is that PWFD(3) detects bigger groups, and hence, the roundness of F-formations are relatively higher than the roundness of F-formations detected by the PWFD(1) and PWFD(2) methods. As seen in figures 8.8, 8.9 and 8.10, most multi-party F-formations detected in the DR dataset have roundness in the range of 0.51 to 0.6. We found that very few multi-party F-formations in the DR dataset have roundness less than 0.3 or greater than 0.8. For the purpose of comparison, we have replotted the roundness of multi-party F-formations resulting from the Party World models (cf. figure 6.6) here in figures 8.11, 8.12 and 8.13.

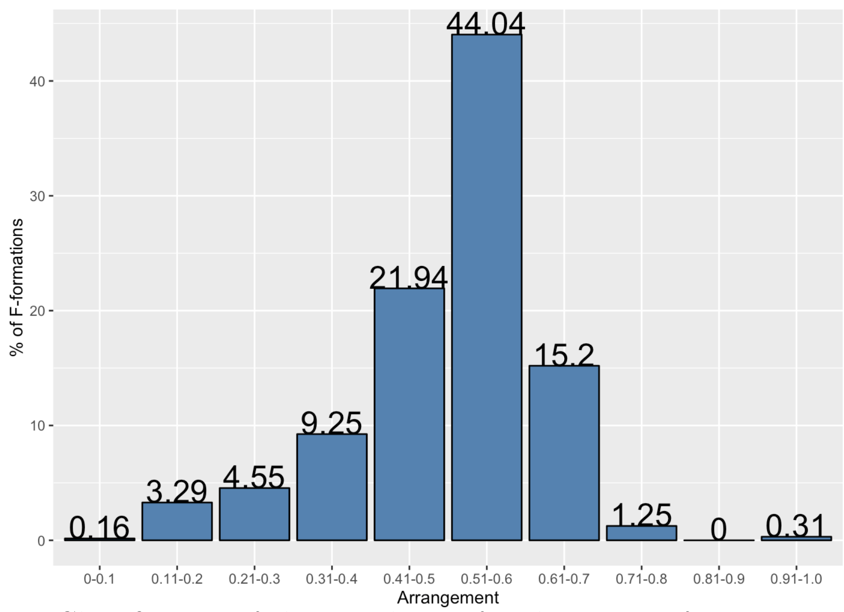


Figure 8.8: Classification of the roundness of multi-party F-formations in the DR dataset (based on the outcomes of PWFD(1))

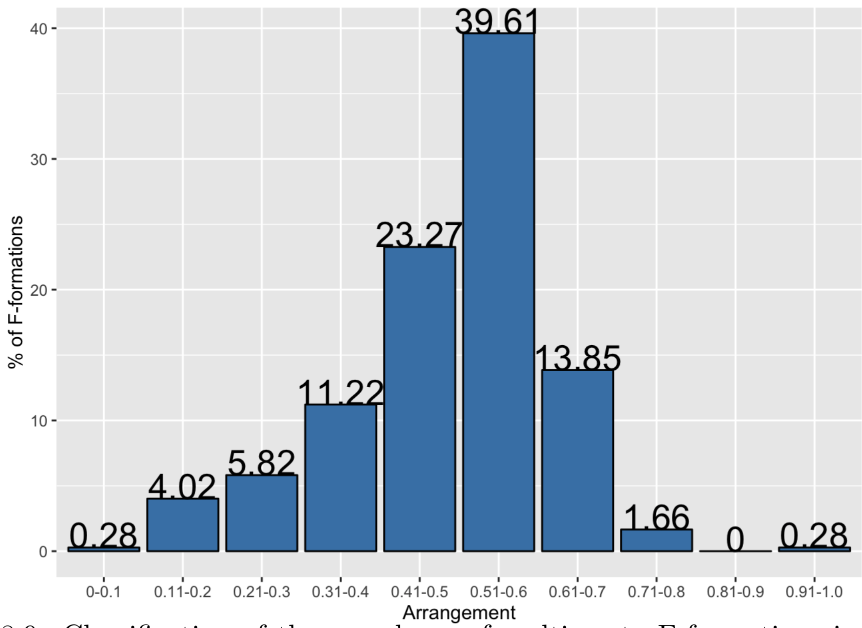


Figure 8.9: Classification of the roundness of multi-party F-formations in the DR dataset (based on the outcomes of PWFD(2))

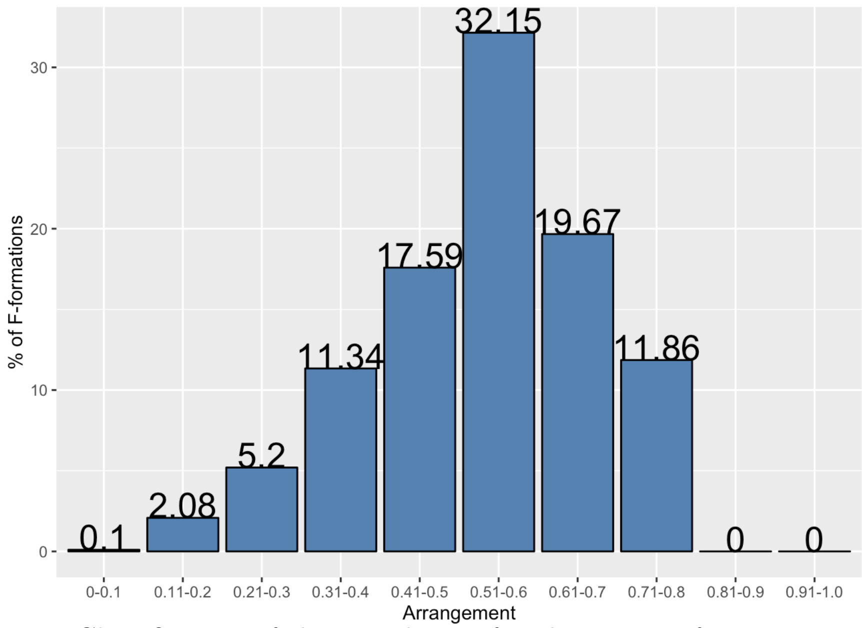


Figure 8.10: Classification of the roundness of multi-party F-formations in the DR dataset (based on the outcomes of PWFD(3))

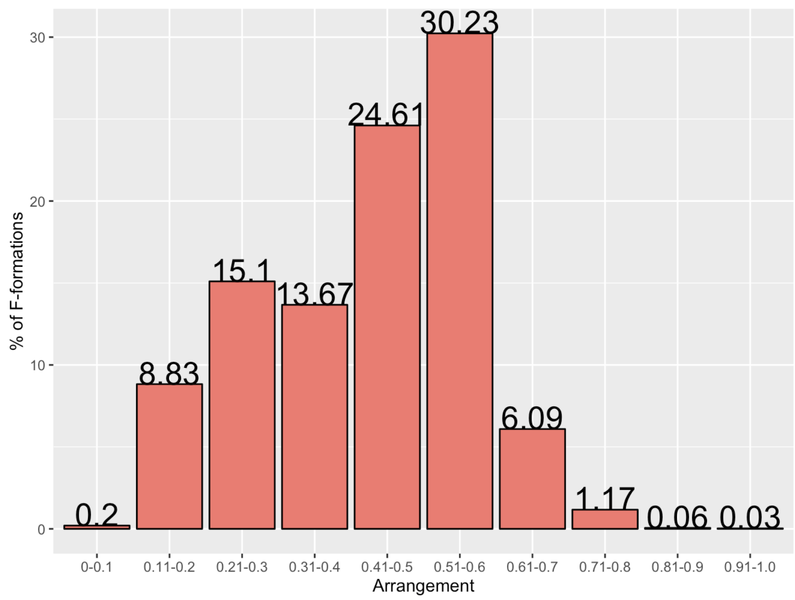


Figure 8.11: Classification of the roundness of multi-party F-formations in Model 1

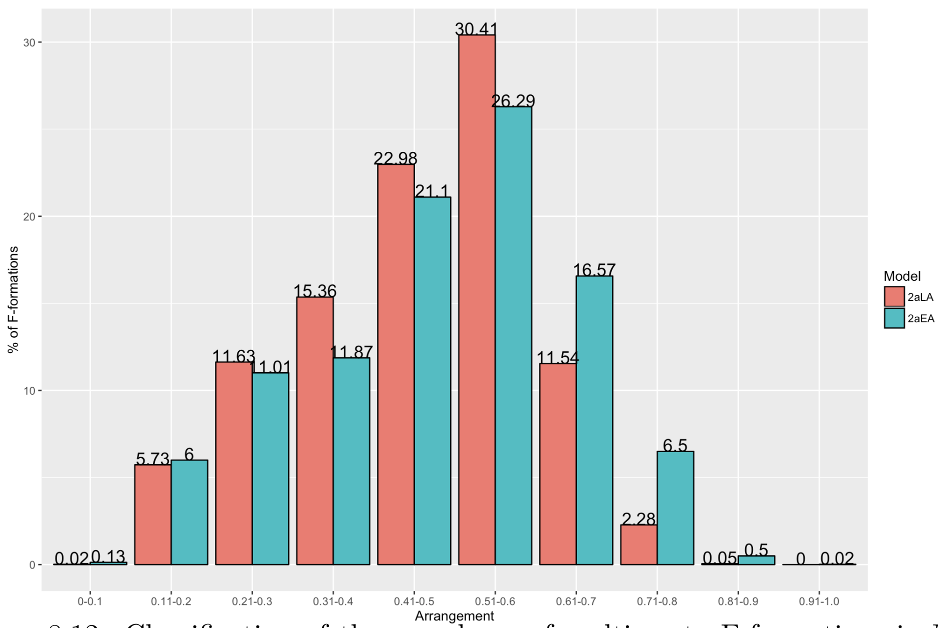


Figure 8.12: Classification of the roundness of multi-party F-formations in Model 2a

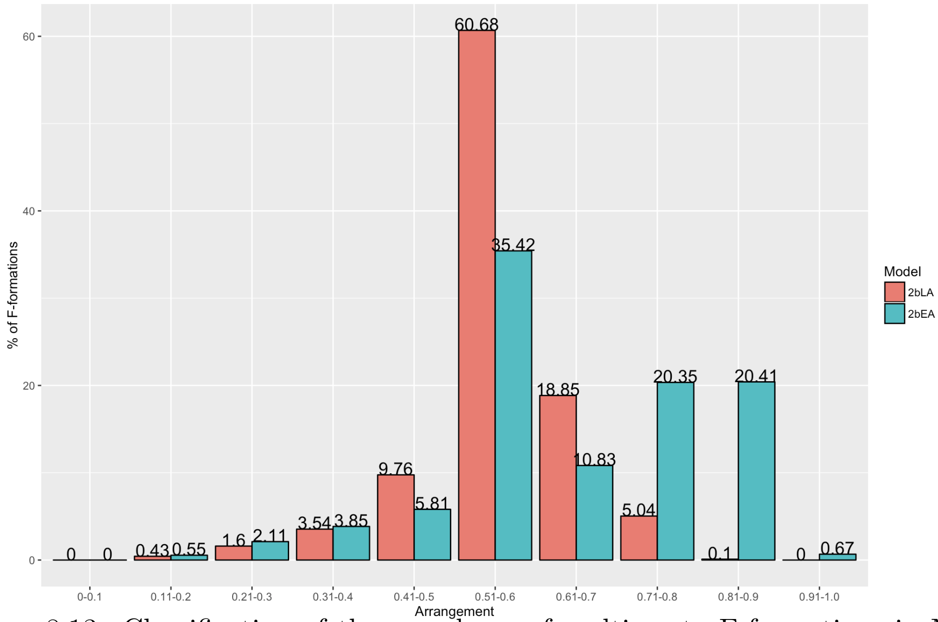


Figure 8.13: Classification of the roundness of multi-party F-formations in Model 2b

As seen in the figures, the roundness of most multi-party F-formations in the DR dataset, as well as the Party World dataset, are in the range of [0.51-0.6]. Furthermore, the roundness of F-formations resulting from the field-of-view models look like a normal distribution, and appear to have similarities with the roundness of F-formations in the DR dataset. We performed independent-samples Kruskal-Wallis test, which showed significant differences between the distributions of different categories of multi-party F-formations in the Party World models and the DR dataset ($p<0.001$). Pairwise comparisons further determined that there is no significant difference between Model 2a EA and the DR dataset (when using PWFD(1) and PWFD(2)), while all other pairwise comparisons showed significant differences. Hence, the roundness of multi-party F-formations resulting from Model 2a EA are not significantly different to the roundness of multi-party F-formations found in the DR dataset (detected using PWFD (1) and PWFD(2)).

The figures also show that, compared to all other models, Model 1 produces a lot more F-formations with roundness less than 0.3, and Model 2b EA produces a lot of F-formations with roundness greater than 0.7, neither of which resembles the trend in the DR dataset. This result shows that the notion of perfectly circular arrangements for multi-party F-formations is inadequate when describing real world conversational groups. It also challenges existing models, which use the centroid-based approach, to simulate perfectly circular F-formations. In summary, based on our results, we argue that field-of-view models work better to simulate multi-party F-formations whose roundness most closely resembles those of real-world conversational groups.

8.2.3 Spatial distribution of conversational groups

Figure 8.14 shows the spatial distribution of participants in the Hub during the drinks reception party. Each distinctly coloured circle in figure 8.14 represents one of the 21 participants. The figure was created by plotting the foot positions of each the 21 participants for the entire duration of the party.

As seen in the figure, participants were spread across the Hub, but often gathered around the kitchen counter and the side table, i.e. places where food and drinks were available. Since we did not model the influence of food and drinks in our Party World simulations, we cannot do a one-to-one comparison between the spatial distribution of agent clusters in the Party World simulations and the spatial distribution of conversational groups in the DR dataset. Existing models of coexisting conversational groups have also not considered the influence of food and drinks on the spatial distribution of conversational groups. Kendon (1990) provides an example of a triadic conversational group that occurs next to a food table. Nevertheless, a detailed study investigating the influence of food and drinks on the functioning of coexisting conversational groups has not been performed.

Although a one-to-one comparison is difficult, a subtle comparison, between the spatial distribution of agents in the Party World simulations and participants in the DR dataset, is still possible. In chapter 5, we found that agent clusters resulting from Models 2b LA and 2b EA use only a small fraction of the available space. Model 1 and Model 2a LA resulted in a much greater spatial distribution of agent clusters. None of these patterns resemble the spatial distribution seen in figure 8.14, but the one from Model 2a EA has a similarity – a big conversational group, where participants keep altering their position and orientation over time, but the group

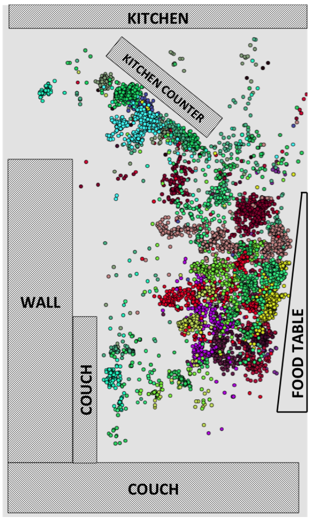


Figure 8.14: Spatial distribution of conversational groups in the Hub

itself does not relocate, and hence, participants cluster in a particular area (cf. figure 5.20d).

We chose the party context for our simulations to model coexisting conversational groups, where participants can move between groups. Such coexisting conversational groups are a common occurrence at parties, and less likely in places like the beach, corridors, hallways, etc. The spatial distribution of conversational groups in the DR dataset suggests that the influence of the location of food/drinks should be considered in a model of coexisting conversational groups. In future Party World simulations, we could model food tables in the Party World environment with an attractive force, to check if it leads to a spatial distribution similar to the one in the DR dataset.

8.2.4 The number of F-formations per image

The number of F-formations per image detected by the HVFF, GCFF and PWFD methods in the DR dataset is shown in figure 8.15. As seen in the figure, there are few F-formations before image number 270. This is because the party did not properly begin until after the first 20-25 minutes, i.e. until guests started arriving from the seminar, which preceded the drinks reception party. But since filming happened in the Hub, which is a kitchen and refreshment area, people kept coming in and out of the Hub, sometimes interacting in groups of two or three. Most of these conversational groups ended before or around the time participants started arriving in the Hub, except for a couple of individuals who stayed back, and continued interacting with the party guests.

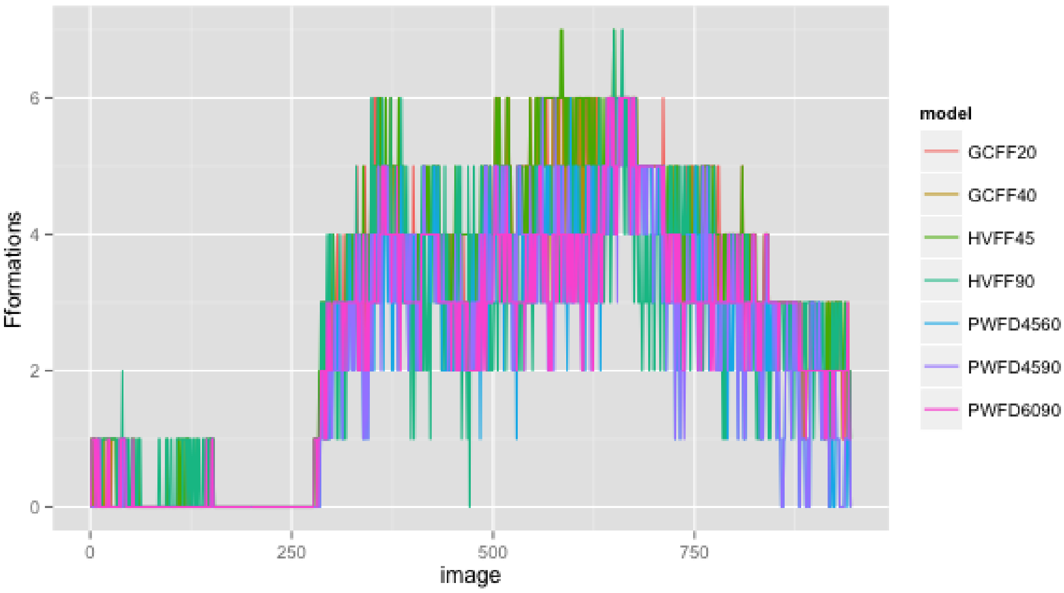


Figure 8.15: The density of F-formations in the DR dataset

8.3 Qualitative Analysis of the Spatio-temporal Patterns of Conversational Groups

The analysis and comparisons in the previous sections did not reveal the characteristics of the spatio-temporal dynamics, or membership dynamics, of conversational groups in the DR dataset, and its similarities to the Party World dataset. That is to say, we performed a quantitative analysis of the similarities and differences in the number, size and shapes of agent clusters and human conversational groups, by considering F-formations as the unit of analysis. However, we have not explored the dynamics of people forming, sustaining, ending, and moving between conversational groups. To this end, we present a qualitative analysis of changes in the spatio-temporal and membership characteristics of conversational groups in the DR dataset.

The analysis is similar to the timeline analysis of agent clusters resulting from the Party World simulations (cf. section 5.4 of chapter 5). We examine the spatial and temporal characteristics and membership dynamics of conversational groups in the DR dataset, and compare them with the characteristics of agent clusters resulting from the different Party World models. Our analysis is based on the reconstructed Party World type of images in the DR dataset (we used the original camera images for additional reference). The instances described here were chosen by visually inspecting the images in the DR dataset, to provide a concise overview of the social, spatial and temporal characteristics of coexisting conversational groups.

8.3.1 Spatial arrangements of agent clusters before the start of the party

In conversational groups, which existed before the start of the party, interactants were standing quite far apart, and sometimes, not even facing each other. Figure 8.16 shows examples of such arrangements. Since there were not many people in the Hub before the party, it may have been that interactants did not worry about intrusions/intruders, and so, did not stand close to one another. Furthermore, Ciolek and Kendon (1980) and Kendon (1990) note that the spatial arrangements of conversational groups may not be well-defined if the encounters last only for a short time. In section 8.2.4, we noted that conversational groups lasted only for a short time before the start of the party, so this could have also been a reason.

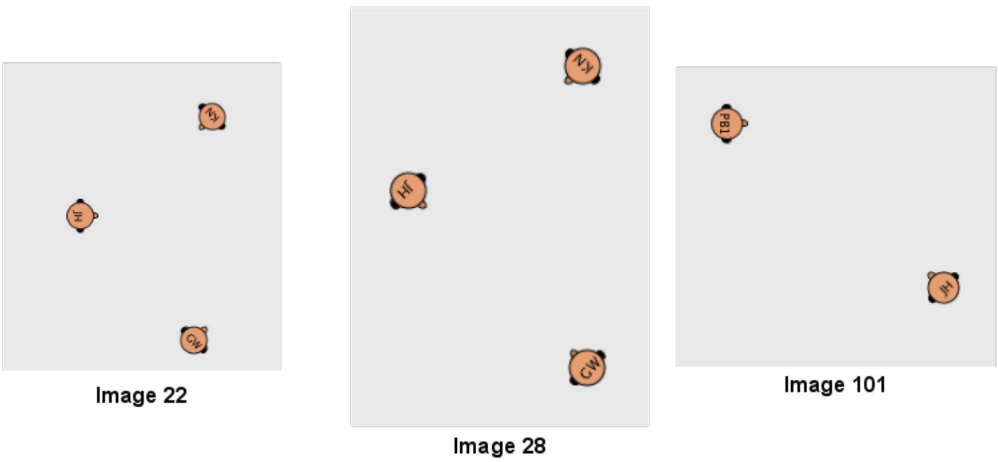


Figure 8.16: Examples of conversational groups before the start of the party

8.3.2 Spatial arrangements at the beginning of the party

Participants started arriving in the Hub from image 277 onwards, and this can be corroborated with figure 8.15, where the number of F-formations increases from image 280 onwards. A notable aspect about the arrival of participants was that those who arrived together, or around the same time, started forming conversational groups almost instantly. We noticed a rough clustering of people, but no clearly defined spatial arrangements until image 287. However, by image 308, i.e. roughly within 4 minutes after participants started arriving at the Hub, there were clearly defined and tightly organised conversational groups, where participants were standing next to and facing each other. In image 335, two new members joined one of the conversational groups that already had four participants, and the resulting arrangement is shown in figure 8.17. As seen in the figure, the arrangement of the conversational group is not a perfect circle.

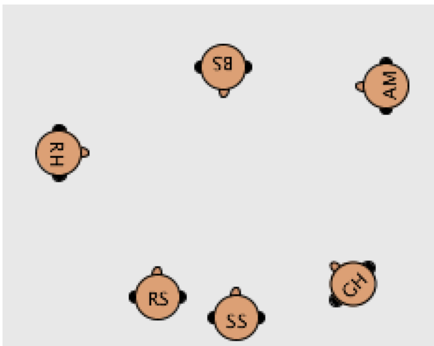


Figure 8.17: Spatial pattern of a 6-member conversational group in image 335 of the DR dataset

As of image 343, there were three distinct conversational groups, each well separated from one another, in the Hub. The stability of these conversational groups was momentarily disrupted in image 346, when more people started arriving in the Hub, but participants reorganised their position and orientation to compensate for the changes. By image 369, there were four distinct conversational groups in the hub, as shown in figure 8.18.

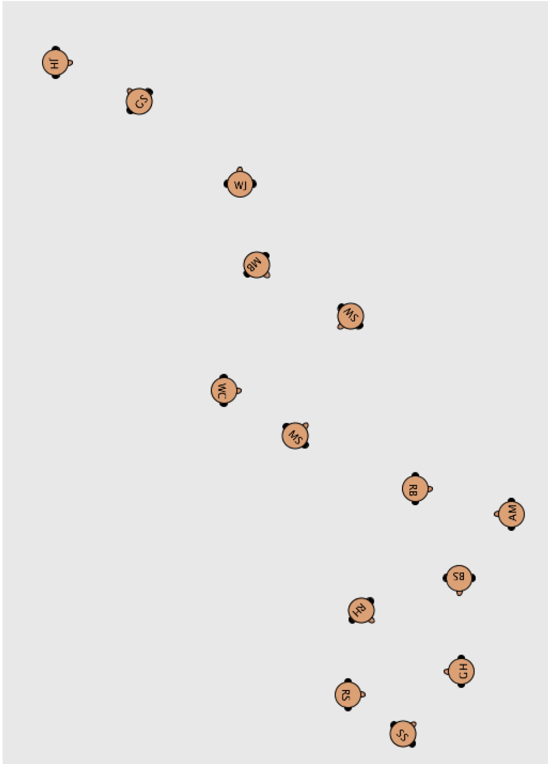


Figure 8.18: Four distinct conversational groups in image 369 of the DR dataset

8.3.3 Changes in participants and the spatial patterns of groups

As of image 424, SW, WJ and EC[1] were in a triadic arrangement; the V-shaped arrangement of AM and RB underwent several changes since image 369; MB, WC and MS were in an (almost) equilateral triangle; the big group comprising of BS, RS, RH, GH and SS were in a closed, but not in a perfectly circular arrangement (see figure 8.19).

As of image 450, AM stopped interacting with RB, and joined the big conversational group from an outsider position (see figure 8.20). The outsider position of AM was clearer in the reconstructed bird's eye view Party World type image of the Hub (i.e. in figure 8.20) than in the original camera image. In several other cases, too, the reconstructed images enabled capturing such subtle yet important aspects of the spatial organisation of conversational groups.

There were several changes over the next few minutes, both in the spatial patterns of conversational groups, as well as in their membership. As of image 495, there were

[1]EC is not seen in the image due to missing annotation, i.e. because the foot position of EC was occluded in the camera image.

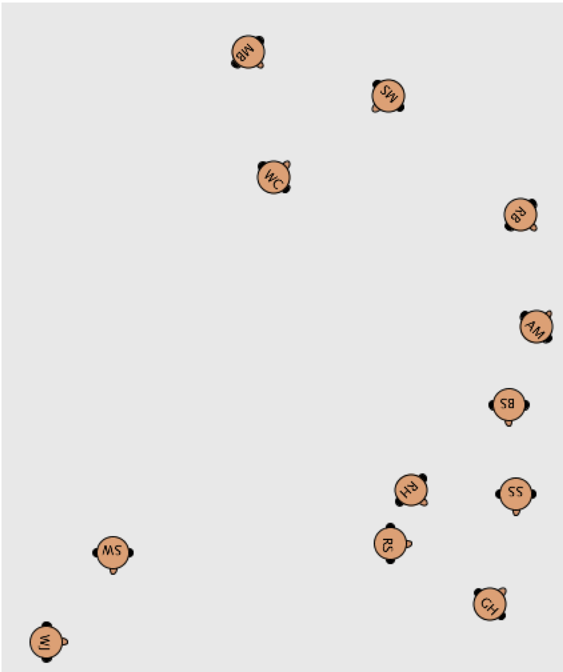


Figure 8.19: Changes in the four conversational groups as of image 424

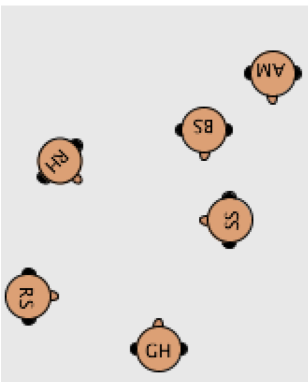


Figure 8.20: The outsider position of AM near the big conversational group

three conversational groups in the Hub, two dyadic arrangements (involving EC and RB and MB and WC, respectively) and one big conversational group involving MS, BS, RH, RS, GH and SS. As seen in figure 8.21, the big group was closed but not circular. The dyadic arrangement of MB and WC was an N-shaped closed pattern, while the arrangement of EC and RB was a V-shaped closed arrangement.

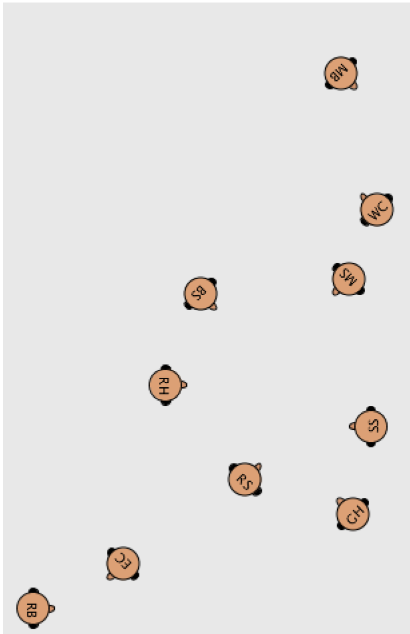


Figure 8.21: Three distinct conversational groups in image 495 of the DR dataset

By image 505, there were three dyadic groups and one 7-member group (figure 8.22). WJ and SW were interacting at the front of the hub (top left of the image) in a V-shaped arrangement; EC and RB continued in a V-shaped arrangement at the back of the hub (bottom centre of the image); MB and WC transitioned from an N-shaped arrangement to a H-shaped arrangement. Lastly, the 7-member group is not a proper circle, but a semi-closed (a.k.a. semi-open) arrangement with an open spot between BS and MS.

8.3.4 Conversational roles determining the spatial organisation of groups

The spatial organisation of the big group in image 8.23 is an example of conversational roles determining the positional-orientational arrangement of participants. AM and MS are interacting with each other, while others in the group are look-

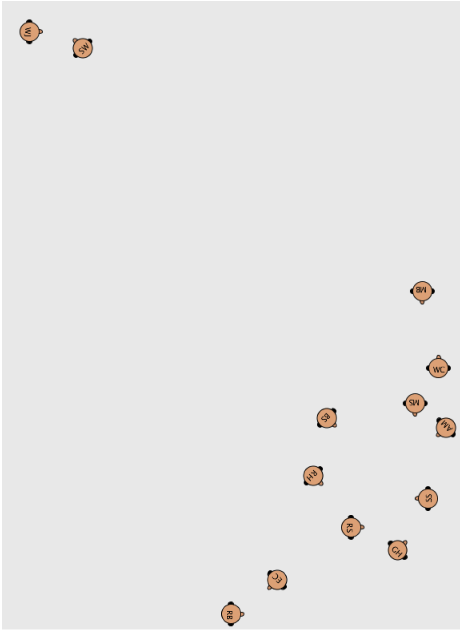


Figure 8.22: Four distinct conversational groups in image 505 of the DR dataset

ing at the duo. AM and MS are facing each other, almost directly, when talking, while others (BS, RH, RS, GH and SS) are standing in a semi-circular arrangement, looking at the duo.

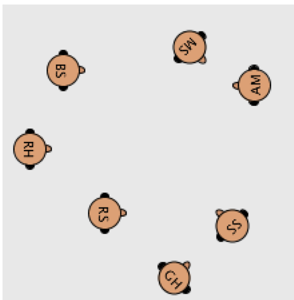


Figure 8.23: Example: Conversational roles determining the spatial layout of conversational groups (image 531)

8.3.5 The absence of perfectly circular arrangements

As of image 569, more participants joined the big group, and it had 10 members, but the layout was not a proper circle (see figure 8.24). Normally, in indoor environments, bigger groups may not have the space needed to establish and maintain properly circular arrangements. But in this case, even when there was empty space available in the Hub, the big group did not spread around to form/maintain a regular shaped pattern. One reason could be that interactants preferred continuing their interactions in the same spot, where they first formed the group, i.e. irrespective of gaining or losing members. We found this to be the case with all conversational groups (small or big) in the DR dataset. We found that interactants alter their position and orientation several times, which in turn causes the group to shift slightly, but relocation never happens.

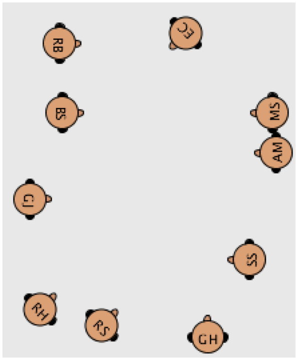


Figure 8.24: The 10-member conversational group

Another motivation to remain in the same spot, particularly for the big group, could have been the presence of food and/or furniture. We noticed that participants tried to access food and drinks, or lean on the side table, while interacting. This observation confirms that the environment influences the spatial patterns of conversational groups (Ciolek and Kendon, 1980; Kendon, 2010).

Notwithstanding the potential reasons, the non-circular arrangements of conversational groups in the DR dataset suggest that the perfectly circular (or regular shaped) patterns resulting from the centroid-based approach with an extended neighbourhood influence, does not correspond to reality. The (non-regular) spatial patterns resulting from Models 2a LA, 2a EA and 2b LA more closely resemble the spatial patterns of conversational groups in the DR dataset.

8.3.6 The splitting of bigger groups

The splitting of bigger conversational groups into sub-groups is a common phenomenon (Kendon, 1990), and it occurred in the DR dataset, too. The big conversational group is fragmented in image 625 (see figure 8.25). In chapter 5, we showed that the splitting phenomenon seldom occurs in models with an extended neighbourhood influence, whereas the LA models allow it. This is because, Model 2a LA and Model 2b LA result in semi-open or open spatial arrangements, which allow agents to leave the group more easily. In this regard, it appears that the LA Party World models produce more realistic outcomes than the EA models.

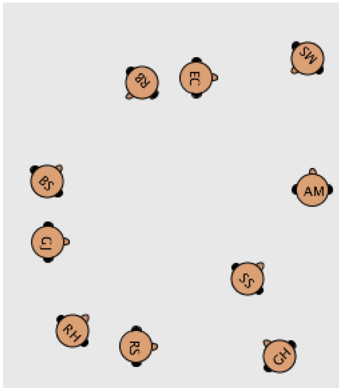


Figure 8.25: The splitting of a bigger conversational group into sub-groups

8.3.7 The onlooker behaviour

In image 632, some people are actively conversing – RB, SS, RS and GH and EC, MS and AM, while others are overlooking – BS, GJ and RH (see figure 8.26). Such onlooker type of spatial patterns result from Models 1 and 2a LA, and sometimes also from Model 2b LA, but seldom from Models 2a EA and 2b EA. This is because the EA models cause every agent in a group to readjust with respect to every other agent in the group. Consequently, agents end up facing one another more or less directly (2a EA), or facing the centroid point (2b EA), and these make agents appear as active participants and not just onlookers.

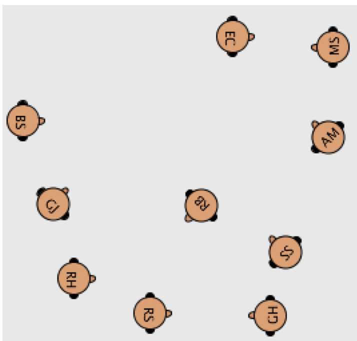


Figure 8.26: An example of onlooker behaviour

8.3.8 The aggregation of smaller groups

Another phenomenon observed in the DR dataset is that the smaller groups, which split from a bigger group, aggregate again. However, the aggregated groups do not last long. The phenomenon of bigger groups splitting into sub-groups, and then, grouping again happens only in the LA Party World models. However, in the DR dataset, it is the same set of people who break off and join together, whereas in Models 2a LA and 2b LA, there is a change of agents. Furthermore, we found that the splitting and regrouping phenomenon takes longer in the DR dataset, approximately 8.5 minutes in one prominent case. However, in Models 2a LA and 2b LA, the splitting and regrouping occurs within a few seconds.

Kendon (1990) notes that F-formations continue to exist even when there is a complete turnover of participants. This proposition potentially justifies the change of agents in the clusters resulting from Models 2a LA and 2b LA. Nevertheless, we propose that future versions of Party World models could include an *interaction affinity* parameter, to allow agents to stay in or return to the same groups.

8.3.9 The proximity of sub-groups

Another interesting phenomenon in the DR dataset is that the sub-groups remain close to one another despite there being a lot of empty space in the Hub, e.g. see figure 8.27. In chapter 5, we showed that the splitting of groups in Model 2b LA, causes sub-groups to remain close to one another.

8.3.10 The entry and exit of participants

From image 750 onwards, i.e. roughly 35 to 40 minutes after the party began, participants started leaving. As of image 823, there were four groups, each having only two or three members. On the other hand, the Party World models had 20 agents from the start to the end of the simulations. Since the arrival and exit of participants in the DR dataset contradicts with the models in terms of agents being always present, in future versions, we propose to make the environment unbounded, and assign different entry, exit and re-entry times for agents.

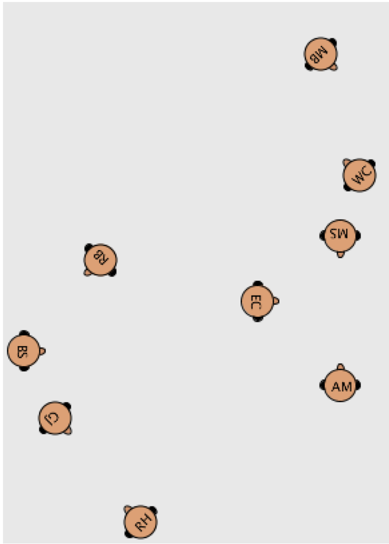


Figure 8.27: The proximity of sub-groups

8.3.11 Spatial arrangements at the end of the party

Towards the end of the party, we found that interactants were standing quite far apart, as they were in conversational groups which existed before the party began. This again confirms that, in less crowded environments, participants are relaxed about their spacing and orientation behaviour. Therefore, in future versions of the Party World models, where the agent population varies over time, we propose to implement positional-orientational behavioural rules relative to the crowdedness of the environment.

8.4 Summary of Key Findings

There is a lack of evidence that existing models to simulate the spatial patterns of human conversational groups have been validated systematically as we have done in this chapter. To the best of our knowledge, the analyses presented in this chapter are the first quantitative and qualitative comparative studies of the spatial patterns of human versus simulated conversational groups. Following is a summary of findings:

Lack of perfectly circular arrangements. The conversational groups in the DR dataset did not have a perfectly circular, or any other regular-shaped, pattern. The currently popular Model 2b EA always produces regular shapes. Hence we argue that Model 2b EA is not suitable for simulating the spatial patterns of real world multi-party conversational groups. Models 2a LA, 2b LA and 2a EA produce more realistic spatial patterns of multi-party groups than Model 2b EA. We do not recommend Model 1 because it generates a lot of queue arrangements which were not found in the DR dataset.

Models 2a LA and 2b LA result in smaller, variably-shaped clusters, which occur in nearby locations. We observed this trend in almost all the images between 277 to 850 in the DR dataset. Model 2a EA causes smaller groups to aggregate into a bigger group. This phenomenon was also observed in the DR dataset, however, unlike Model 2a EA, where almost all the agents aggregate into a group, in the DR dataset, only a fraction of participants form a big group. Based on this, we argue that a local neighbourhood influence is better than an extended neighbourhood influence to simulate conversational groups. Another option would be to activate Model 2a EA to form big groups on specific occasions. For example, a specific event, such as an announcement or cake-cutting at a party, could be modelled as a triggering event.

The dyadic arrangements assume H, N, V, C, L and I-shapes. The dyadic arrangements in the DR dataset assume all the six types of arrangements identified by Ciolek and Kendon (1980) – H, N, V, L, C and I-shaped arrangements. This finding leads us to reject Model 2b LA and Model 2b EA for simulating dyadic groups, because they mostly produce only H-shaped arrangements. We also reject Model 1, because it produces a lot of queue arrangements, whereas the DR dataset had only 2% of queue arrangements. Although the proportions of the dyadic shapes resulting from the field-of-view models were not statistically similar to those of the DR dataset, they produce a lot more V, C, L and N-shaped arrangements and fewer queue arrangements than the other Party World models. Hence, we argue that the field-of-view models work better for simulating realistic dyadic patterns. Lastly, since we found that Party World models produce significantly less I-shaped arrangements than the DR dataset, we proposed modelling an attractor event in future simulations, to see if that leads to the emergence of more I-shaped arrangements.

The roundness of multi-party F-formations. We found that the roundness of multi-party F-formations generated by Model 2a EA most closely resembles the roundness of multi-party F-formations detected in the DR dataset. Although not statistically significant, the distribution of roundness of multi-party F-formations resulting from Model 2a LA has some similarities to the roundness of F-formations in the DR dataset. Model 2b LA also has some similarities to the DR dataset in respect of this measure, but it produces a lot more F-formations with roundness in the range of [0.5-0.6]. Model 1 produces a lot of F-formations with roundness < 0.3 and Model 2b EA produces several F-formations with roundness > 0.7. Neither of these correspond to the roundness of F-formations detected in the DR dataset. Hence, we argue that the field-of-view models and the centroid-based model with a local neighbourhood influence are more suitable for simulating multi-party conversational groups in respect of the roundness measure.

Need to model the influence of food/drinks on conversational groups. In the DR dataset, we found that conversational groups clustered around the table with food and drinks. Although we checked the robustness of models by including obstacles in the environment, we did not model the influence of furniture or food on conversational groups. Consequently, we were unable to provide a one-to-one comparison of the spatial distribution of agent clusters resulting from the models, and the spatial distribution of conversational groups in the DR dataset. Nevertheless, we were able to provide a simpler comparison. The spatial distribution of agent clusters resulting from Models 2b LA and 2b EA are restricted to a small fraction of the available space, whereas the ones resulting from Model 1 and Model 2a LA are more dispersed. Model 2a EA produces a unique outcome, where the spatial distribution of agents is highly concentrated at any one particular spot. Considering these results, we recommend that a combination of Models 2a LA and 2a EA may be ideal to achieve the right amount of spatial distribution and concentration of agent clusters, with future models also including some attractor areas, such as the food and drinks tables in the DR dataset.

People arrive and leave at different times. In the DR dataset, we found that participants arrive and leave at different times, whereas agents are always present in the Party World simulations. This is a limitation of the Party World simulations, where the environment is bounded, and so agents are always present. We propose making the Party World environment unbounded, thereby allowing agents to arrive, leave, and re-enter at different times. We hope this will make the simulations more realistic.

Furthermore, we propose adding rules that will cause agents to adapt spacing and orientation patterns, depending upon the crowdedness of the environment. We hope this could lead to observing well-spaced out conversational groups at the start and the end of simulations, similar to the DR dataset. We also hope that allowing agents to arrive and depart at different times would control the number of F-formations resulting from the simulations. It was shown in section 8.1 that 360 images (captured at 5 second intervals) from the Party World simulations had three times as many F-formations as the ones detected in the 944 images (also captured at 5 second intervals) of the DR dataset.

The splitting and rejoining of groups. In the DR dataset, we found that the same set of people form, split, and re-establish conversational groups. Although the Party World models replicate the process of splitting and rejoining, there are usually changes in the membership. Model 2a LA and Model 2b LA cause bigger groups to split into sub-groups or smaller groups to aggregate into a big group. However, it is usually not the same set of agents that are involved in the process. The change in membership may be justified (cf. Kendon (1990)), but we propose including an *interaction affinity* parameter in the models, which would allow agents to remember conversational encounters and partners, and thereby, enable them to to stay in or return to the same groups. We believe the interaction affinity parameter would account for the effect of familiarity evident in the DR dataset, where participants already knew each other, and hence may have formed F-formations with their more familiar counterparts.

Transitioning between groups. Participants in the DR dataset did not shift between groups very often, but agents frequently shift between groups. We hope that adding the *interaction affinity* parameter would address this issue as well, especially in Models 2a LA and 2b LA, where agents transition a lot. In Model 2a EA, agents remain in the same group for relatively longer, but the group seldom splits into sub-groups. Model 2b EA results in one or two big groups and few smaller groups with relatively stable membership.

Taking into account all of the investigated parameters, we argue that among the Party World models, the field-of-view models are the most suitable for simulating the spatial patterns of dyadic and multi-party conversational groups. We also argue that local neighbourhood influence is better than extended neighbourhood influence, because it reproduces most of the characteristics of real-world conversational groups, and is cognitively more plausible and computationally efficient. It's a cognitively plausible assumption that individuals would readjust with respect to their immediate neighbours, rather than remembering and relating their positional-orientational behaviour with respect to every other member in the group. We also proposed activating the EA influence selectively in Model 2a to form big groups on specific occasions, e.g. we could model neighbourhood influence as a dynamic parameter which alternates between LA and EA depending on triggering events. We also discussed the need for modelling the influence of environmental factors (e.g. food table), and different entry and exit times for agents, when modelling coexisting conversational groups in a party context. However, most importantly, we have shown that the field-of-view models, with their simple rules of movement, positioning and orientation behaviour, lead to the emergence of realistic spatial patterns of conversational groups.

Chapter 9

Findings from a User Evaluation Study

In this chapter, we present the findings from a pilot user evaluation study, which was conducted before the development of the final versions of the Party World models. The study is inspired by Braitenberg's (1986) formulation of simple vehicles eliciting perceptions of complex emotions, such as the display of love, fear, aggression, etc. During the study, we asked participants to evaluate images and videos generated from earlier versions of the Party World models, to identify agent clusters that resembled human conversational groups. The findings of the study led to the development of the final versions of the Party World models, i.e. Models 1, 2a (LA and EA) and 2b (LA and EA). The study also resulted in identifying issues that warrant further research, especially with regard to intimate conversational encounters.

The rest of the chapter is organised as follows. In section 9.1, we begin by reviewing the earlier Party World models. In section 9.2, we review the image and video analysis tasks in the user evaluation study. In section 9.3, we describe the procedure of the study. In section 9.4, we present the results of comparing participants' assessments of appropriate and inappropriate arrangements of conversational groups, in the images and videos from the Party World simulations. In section 9.5, we review participants' qualitative descriptions of the characteristics of agent clusters resulting from the simulations. We end the chapter with a summary of major findings and by identifying potentially important areas for further research.

9.1 Earlier Party World Models: Model 0 and Model 1

The earlier Party World models used in the evaluation study were the Level 0 and Level 1 models. In Level 0 simulations, agents wander randomly, while being oblivious to the presence of one another. Agents also stop moving randomly, and they remain stationary for a few seconds, before moving again. Since agents don't recognise other agents in the system, they tend to collide with one another when moving, or overlap with one another in the stationary state.

The Level 0 model is evidently unrealistic for simulating the spatial behaviour of humans. It's like Braitenberg's (1986) formulation of the most basic vehicle. Nevertheless, our objective was to assess participants' evaluations of the modest Level 0, to see if they rejected the agent clusters resulting from the model as invalid representations of human conversational groups.

In Level 1 agents move around randomly, however, unlike Level 0, agents recognise one another, and hence, they either stop moving when the separation distance condition is violated (i.e. similar to Party World Model 1), or steer away from one another. The latter makes it look as though agents avoid one another. Due to the separation, or avoidance behaviour, in Level 1 agents never collide with one another as they do in Level 0.

For the user evaluation study, we used the shape shown in figure 9.1 to denote agents, i.e. instead of the blob shape used in the actual Party World simulations. Figure 9.1 is a scaled-down version of a three-dimensional bird's eye view of the head and shoulders of a person. Just like the actual Party World models, we ran the Level 0 and Level 1 simulations with twenty agents, each identified by a unique number on their head.

Figure 9.1: The shape of agents in Level 0 and Level 1 simulations

9.2 Image and Video Analysis Tasks

The user evaluation study involved two tasks: an image analysis task and a video analysis task.

The image analysis task: We ran the Level 0 and Level 1 simulations for 5 minutes, capturing a screenshot of the simulation window, whenever any agent stopped moving. The images generated from the Level 0 simulation were collated as one dataset, and the ones generated from Level 1 were collated as another dataset. The images were printed on a A4 paper with a marking on the top left corner to identify the level to which the images belonged. The markings were only for the investigator's reference – participants were not made aware of the concept of levels, or what the markings meant. By doing so, our intention was to investigate if participants' evaluations of the agent clusters in Level 0 images were any different from the ones in Level 1 images.

In no fixed order, we retrieved 10 images each from the Level 0 dataset and the Level 1 dataset, to create a test set for each participant. We gave each participant a set of test images, a green ink pen, and a red ink pen. We asked them to imagine that they were looking at a group of agents from a bird's eye view. We allocated 30 minutes for the image analysis task. During this time, participants had to:

- Draw a circle around the different agent clusters, which appeared like conversational groups, in each image using a green ink pen.
- Describe in writing, right next to the circles drawn in step 1, the nature of conversations within the groups.
- Draw a circle around agent clusters, which appeared inappropriate, in each image using a red ink pen.
- Describe in writing, right next to the circles drawn in step 3, the reason for inappropriateness.

Figure 9.2 shows an image annotated by one of the participants. We found that some participants took longer than the prescribed 30 minutes to annotate the images.

Some participants finished the task around 30 minutes but did not annotate all the 20 images. At the end of the study, we found that each participant had annotated anywhere between 15 and 20 images. It was interesting to observe that participants took time to study the images and make annotations.

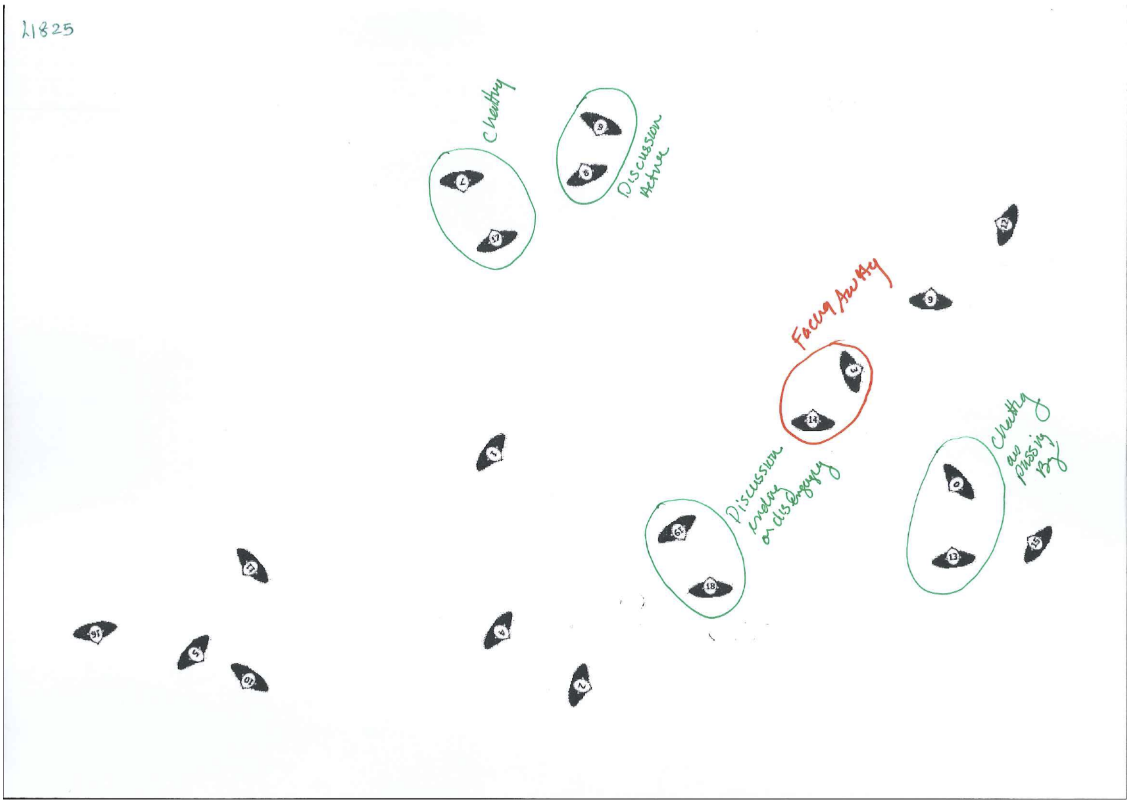


Figure 9.2: Example of an annotated image from the user evaluation study

The video analysis task: We ran the Level 0 and Level 1 simulations for 5 minutes each, making a movie of the simulation, using a screen recording software[1]. We imported the videos in OpenCV, and built a new functionality, to allow participants to place markers on the videos. The marker functionality allowed participants to place a green-coloured square marker on the video by clicking the left mouse button and a red-coloured square marker by clicking the middle mouse button. We set the delay between mouse clicks and the markers appearing on the video to zero in order to minimize latency errors.

We showed the same two videos, one from Level 0 and another from Level 1, to all

[1]Apple QuickTime Player. http://www.apple.com/quicktime/download/

the participants. Once again, participants were not made aware of the concept of levels, or what the two different videos denoted. At the start of the task, we asked participants to imagine that they were going to look at a group of agents moving, from a bird's eye view. We asked each participant to choose and monitor one agent (from 0 to 19) throughout the two videos. While monitoring, we asked participants to place a green marker next to the chosen agent, if they thought the agent started a conversation, or joined an ongoing conversation involving other agents. We asked them to place a red marker next to the chosen agent, if they thought the agent stopped interacting, and left a conversational group. Before the actual video analysis task, we showed a sample video to participants, to allow them to become familiar with the mouse clicks and markers that appeared in the video. We recorded all annotated videos using the same screen recording software.

9.3 The Evaluation Study

We conducted the evaluation study after obtaining the approval of the QMUL research ethics committee (QMREC 0716 – K-space analysis). We recruited seven participants for the study, who were either postgraduate research students, or post-doctorate research fellows at QMUL. There were 3 male and 4 female participants, all fluent speakers of English – some were native speakers and others spoke English as their second language. Participants performed the tasks separately at different times and on different days. They took around 60 to 75 minutes to complete both the tasks.

Before the study, participants were given time to read the instructions for the image analysis and video analysis tasks, and clarify any doubts. We also asked participants to sign consent forms before the study began. After this, participants were left alone to do the image analysis task. At the end of the task, participants handed the images to the experimenter, and were given a five minute break, during which time the experimenter set up the computer for the video analysis task. Once again, participants were left alone to do the task. At the end of the video analysis task, participants completed a questionnaire, which asked "Could you describe what the agents were doing in the videos, as it seemed to you?". Participants volunteered for the study and were not compensated for their participation.

9.3.1 Data obtained from the study

The image analysis task resulted in images, where participants had drawn green- and red-coloured circles around agent clusters, with a brief description identifying the nature of conversations, or the reason for the perceived inappropriateness of arrangements. In some cases, participants had only drawn circles, without providing any description. We were able to collect 117 images in the image analysis task – 59 images in Level 0 and 58 images in Level 1. We recorded three videos from the video analysis task, of which the first one was a training video (although participants were unaware of it) and so excluded from our analysis. The remaining two were videos of the Level 0 and Level 1 simulations respectively. We also collected participants' responses to the post-study questionnaire.

9.4 Comparison of Assessments of Agent Clusters

The evaluation study was done before finalising the actual Party World models. The intention then was to investigate if two-dimensional bird's eye view simulations of agent movement, positioning, and orientation, elicited perceptions of real-world conversational groups. We performed the study to establish the feasibility of our line of inquiry. The hypothesis we wanted to test was that:

Hypothesis I: Modelling the movement, spacing and orientation of agents in a virtual environment will elicit perceptions of human conversational groups.

Hypothesis II: Level 0 will elicit more inappropriate clusters and fewer appropriate clusters than Level 1.

9.4.1 Number of conversational groups

We used the data acquired from the image analysis task to compare the number of conversational groups identified in Level 0 and Level 1 respectively. For Hypothesis I, we wanted to confirm if participants were able to identify any conversational groups at all, in the Level 0 and Level 1 images. For Hypothesis II, we wanted to confirm if the number of conversational groups identified in one level was significantly different from the other. Since in real life situations, it is uncommon for people to collide or overlap with one another, our expectation was that the number of conversational groups (if any) identified in Level 1 would be higher than the ones identified in Level 0.

The number of conversational groups that participants identified in Level 1 exceeded the number of conversational groups identified in Level 0. In Level 1, 183 conversational groups were identified in the 58 images, and in Level 0, 75 conversational groups were identified in the 59 images. A Wilcoxon signed ranks test indicated that participants identified significantly more conversational groups in the Level 1 model than the Level 0 model (p=0.018). The results support hypotheses I and II, participants did identify conversational groups in the simulations, and they identified more conversational groups in Level 1 than in Level 0.

9.4.2 Participation in conversational groups

We used the videos obtained from the video analysis task to compare the number of instances where agents were identified as participating in conversational groups in the Level 0 and Level 1 videos respectively. We expected that the inter-agent spacing rule in Level 1 will improve the likelihood of agents being seen as participating in conversational groups. The outcomes from the video analysis task confirmed our expectation – 28 instances of participation were recorded in the Level 1 video, whereas only 14 instances were identified in the Level 0 video.

9.4.3 Inappropriate arrangements of agents

We compared the number of inappropriate arrangements that participants identified in the Level 0 and Level 1 images respectively. Neither Level 0 nor Level 1 allows agents to face one another, but the latter allows agents to maintain reasonable interpersonal spacing. On the other hand, Level 0 results in overlapping configurations, such as the one shown in figure 9.3. Hence we expected that participants will identify more inappropriate arrangements in Level 0 than in Level 1.

As expected, the images obtained from the image analysis task showed that, participants identified more inappropriate arrangements in Level 0 than in Level 1. They identified 67 inappropriate arrangements in Level 0 and 25 inappropriate arrangements in Level 1. A Wilcoxon signed ranks test indicated that participants identified significantly more inappropriate arrangements in the Level 0 model than the Level 1 model (p=0.018). The result demonstrates that allowing agents to maintain reasonable inter-agent spacing reduces the perceived inappropriateness of the spatial arrangements of agent clusters.

Figure 9.3: The overlapping of agents in a Level 0 image

It was expected that participants would perceive more inappropriate arrangements in Level 0 than in Level 1. But it was unexpected that participants did not reject all the overlapping arrangements from Level 0 as inappropriate. For example, see participants' descriptions for some of the overlapping arrangements from Level 0 in figures 9.4 and 9.5.

It can be seen that some of the overlapping arrangements are seen as inappropriate, with comments such as *gang violence*, *kissing*, *about to get beats*, etc. On the other hand, there are also examples of overlapping arrangements seen as valid encounters, with comments like *embracing couple*, *couple with a child*, *boyfriend*

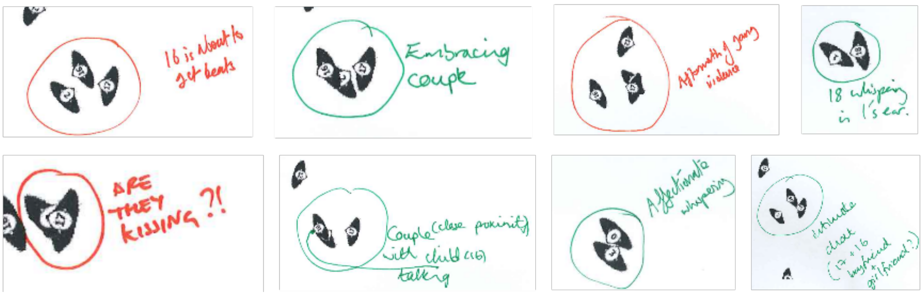


Figure 9.4: Participants' description of agent clusters

Figure 9.5: Different annotations for similar images

and girlfriend, etc. Therefore, while some participants indicated that overlapping arrangements were inappropriate, others felt that it indicated intimate personal relationships between agents. It was also surprising that participants attributed the inappropriateness to the type of encounter (i.e. gang violence, kissing, beating, etc.) but not to the spatial arrangement itself. That is to say, participants did not comment that overlapping configurations are not possible.

The fact that participants saw overlapping configurations, not as inappropriate spatial arrangements, but as valid encounters, demonstrating intimate personal relationships, agrees with Scheflen's (1975) observation that close interpersonal distances means intimate relationship between interactants. The remarks about fighting and gang violence agree with Hall's (1966) observation that, when angry or emphatic about the point being made, interactants move closer than normal.

Existing models have not explored the possibilities of modelling the spatial patterns of conversational groups involving intimates. The equally spaced patterns resulting from the centroid-based approach, evidently would not appear as intimate encounters, or encounters involving intimates. The closest related research is that of Bailenson et al. (2003) and Ennis and O'Sullivan (2012), who explored the effects of face-to-face conversations at the far phase of the intimate distance. However, modelling face-to-face interactions at the close phase of the intimate distance, has not received much attention, but our findings identify it as a potential avenue for further investigation. We could ask questions such as: (1) how couples or intimates are treated within a conversational group, i.e. in terms of modelling their positional-orientational behaviour with respect to one another, and the group, (2) how does their positioning and orientation change over time, etc.

9.4.4 Average duration of conversational groups

The video analysis task required participants to put markers on the videos to denote the start and stop times when agents participate in conversational groups. We define the time elapsed (in seconds) between the placement of a corresponding pair of start and stop markers as the *duration of engagement*. We implemented the simulations such that the time for which agents remain stationary is the same in both Level

0 and Level 1. However, the duration of engagement that participants identified, varied between the two models. The average duration of engagement identified in Level 0 was 12 seconds, while it was 10 seconds in Level 1. The actual values are not important here, but the difference observed in participants' evaluation of the temporal characteristics of agent clusters, is notable. We found there were differences in the temporal characteristics of agent clusters resulting from the final versions of the Party World modes, too (cf. section 6.4.2 in chapter 6).

9.5 Participants' Evaluation of the Characteristics of Agent Clusters

In this section, we analyse and compare participants' qualitative comments, about the characteristics of agent clusters in the Level 0 and Level 1 images and videos.

9.5.1 Participants' classification of conversations

The spatial layout of conversational groups depends on the nature of the conversation (Sommer, 1965; Cook, 1970; Kendon, 2010). We wanted to examine if participants' description about the nature/kind of conversations within agent clusters varied between Level 0 and Level 1. The types of conversations that participants commonly identified, both in Level 0 and Level 1, include: *talking*, *chatting*, *conversations*, *discussing*, and *paying visual attention*. Although there were no rules to coordinate the mutual orientation of agents, both in Level 0 and in Level 1, it is interesting that participants still remarked about paying visual attention. Perhaps because agents were moving in the forward direction, it made it look like agents were more or less facing each other, which in turn looked like paying visual attention.

The types of conversations more commonly identified in Level 0, but less commonly in Level 1, include: *whispering*, *arguing*, *conflict*, and *disagreement*. The categories identified only in Level 0 were: *embracing*, *advising*, *mediating argument*, and *debate*. This again correlates with Hall's (1966) suggestion that interactants stand closer than normal, perhaps within the intimate zone, to perform activities involving the use of body parts, e.g. embracing. Lastly, the categories identified only in Level 1 were: *greeting*, *briefing in passing*, *PhD supervision*, *flirtation*, *ranting*, *overhearing a conversation*, and *criticism.*

It was interesting that participants provided descriptions such as PhD supervision and criticism, to characterise the spatial arrangements of agent clusters, where one agent appeared to have a clearly more authoritative position than others. It agrees with Kendon's (1973) suggestion that, in conversational groups with unequal participation rights, one of the interactants assumes a head position. The differences in the descriptions of the spatial arrangement of agent clusters in Level 0 and Level 1 provide additional support for hypotheses I and II

9.5.2 Reasons attributed to inappropriateness of arrangements

We compared the reasons participants gave for inappropriate arrangements of agent clusters. We found that the reasons identified, both in Level 0 and in Level 1, were: *odd back-to-back arrangement*, *odd queue-like arrangement*, *odd back-turn*, and *odd orientation.* The reasons identified only in Level 0 were: *overlap*, *too close*, *odd group*, *all facing away and odd turning away.* The only other reason identified in Level 1 was that *(agents were) not facing each other.* It is notable that not facing each other was seen as inappropriate in Level 1, but not in Level 0, where a number of other reasons were given. Since agents maintain reasonable interpersonal distances in Level 1, there seemed to be an expectation for agents to face one another, too.

9.5.3 Comments about agents' activities

We compared participants' feedback on what they thought about the activities of agents in the Level 0 and Level 1 videos respectively. For Level 0, participants made comments like *"Milling around in an open space, like a college quad"*, *"Wandering aimlessly, sometimes stopping randomly, sometimes stopping to interact"*, *"They seemed to move a bit random"*, *"Walking around in an airline terminal in a confused and erratic way".* The comments confirm that participants attribute randomness to the spatial behaviour of agents in the Level 0 simulation.

On the other hand, for Level 1, participants made comments like *"Busy City*

Plaza? People rushing about. A station concourse?", *"The different speeds made it look more human."*, *''They seemed more intentional in their movements. The fact that they swing around and made decisive turns made it seem they had "intentions"...Their formation seemed more natural"*., and *"Agent 14 was avoiding talking to people at a public park or train station square."*[2]. There are two inferences here: firstly, the turning away (or avoidance) behaviour of agents seemed to have a noticeable effect. Secondly, participants attribute purposiveness to the spatial behaviour of agents, which is unlike the randomness attributed to Level 0. Once again, participants' comments about the activities of agents in Level 0 and Level 1 are consistent with hypotheses I and II.

9.6 Summary

Our study showed that the Level 0 and Level 1 models were both successful in eliciting perceptions of human conversational groups. Nevertheless, there were significant differences in the evaluations of agent clusters resulting from the two models. The lack of interpersonal spacing and overlapping were seen as inappropriate in Level 0, but at the same time, participants also provided plausible explanations to justify the lack of spacing between agents, e.g. kissing, embracing, holding a baby, etc. On the other hand, since agents maintain reasonable interpersonal distances in Level 1, there seemed to be an expectation for agents to face one another. This led us to consider that modelling rules for governing the orientation of agents is a gradual next step. Hence we came up with Models 2a and 2b.

Our study also revealed other issues, which led to the development of the final versions of the Party World models, as well as choosing a different method for validating the spatial patterns of agent clusters.

The irrelevance of avoidance behaviour. Level 1 had two different rules to prevent agents from colliding with each other. One is to stop moving and the other is to steer in a different direction. We found that participants perceive the latter as avoidance behaviour, which was seen as an issue, particularly in the video analysis task.

[2]This comment was made by a participant who tracked agent no. 14

If the agent that participants had to monitor was moving away more often instead of stopping, it caused participants to believe that the agent preferred avoiding conversational engagements. Furthermore, the avoidance behaviour did not contribute to our research questions, which aimed to establish the effects of simulating spacing and orientation behaviour in leading to the emergence of conversational groups. Hence, for the actual Party World simulations, we developed Model 1 similar to Level 1, but without the rule that causes agents to avoid encounters.

The lack of circular arrangements was not seen as an issue. None of the arrangements resulting from the Level 0 or Level 1 simulations were circular, or of any other regular shape, and yet, there were no comments about non-circular (or non-regular) arrangements not being conversational groups. This provided another reason for us to develop Model 2a (LA and EA), where the behavioural rules did not cause agents to form regular shaped patterns alone.

Evaluating images can be misleading. Just by looking at the images, participants were unable to differentiate stationary agents from those that were moving at the time the screenshots were made. Consequently, participants considered all 20 agents while evaluating the images, and this caused some uncertainty in their evaluations. Including moving persons is also an issue in the automatic detection of F-formations (Gan et al., 2013). Therefore, while detecting the F-formations resulting from the actual Party World simulations (i.e. using the HVFF, GCFF and PWFD methods), we only included stationary agents in the assessments.

Level 0 cannot be the baseline. The Level 0 model described here differs from the Level 1 model in two aspects. (1) In Level 0, agents cannot perceive one another, whereas in Level 1 agents can perceive one another. (2) In Level 0, agents stop moving randomly, whereas in Level 1, agents stop moving only when the minimum separation distance condition is violated. Therefore, it's a step change from Level 0 to Level 1. Furthermore, the complete lack of perception and the random stopping behaviour in Level 0, which causes agents to overlap, are modest assumptions which do not pertain to our research questions. Hence, for the actual Party World models, we did not implement the Level 0 model. We considered Model 1 (equivalent to Level 1) as the baseline, followed by Models 2a (LA and EA) and 2b (LA and EA).

Alternative to questioning participants. The comments participants provided during the tasks suggested a tendency to overinterpret the images used in the tasks. Furthermore, in section 2.10, unobtrusive observation of naturally occurring behaviour was considered better than questioning participants, when trying to develop a systematic understanding of the behaviour being studied. Therefore, for validating the actual Party World models, we decided to compare the spatial patterns of agent clusters resulting from the models against the spatial patterns of naturally occurring conversational groups.

Modelling intimate encounters / encounters involving intimates. Our user evaluation study highlighted the relevance of modelling face-to-face encounters at the close phase of the intimate distance. Potential questions include: (1) how to model the positional-orientational relationship of intimates within conversational groups, (2) if and how does the spatial behaviour of intimates adapt over time during conversational encounters.

Part IV

Conclusions

Chapter 10

Conclusions and Future Work

10.1 Thesis summary

In this thesis, we have presented a simulation-based analysis of computational models for generating the spatial patterns of conversational groups. In face-to-face encounters, participants adjust their body position and orientation relative to one another in order to see and hear clearly. In our simulations, agents follow simple rules of movement, interpersonal spacing and mutual orientation, which in turn leads to the emergence of agent clusters.

The overarching goal of our research was to perform a systematic analysis of alternative models to simulating agent clusters, whose spatial, temporal, and membership dynamics resembles human conversational groups. The boid model has been extensively adapted to simulate the steering behaviour of agents, but Reynolds (1987) did not validate the flocking patterns resulting from the model with real bird flocks. Huth and Wissel (1992) proposed a fish schooling model similar to Reynolds's (1987) boid model, but the difference is that they compared the results of their model with real fish schools. The purpose of comparison was to identify which combination of rules is most likely to simulate fish schools with characteristics typical of real fish schools. Railsback and Grimm (2011) characterise this as doing "real science" with agent-based simulations. They also proposed the concept of *scenario contrast experiments* to compare alternative models intended to generate the same target phenomenon.

There is a huge body of work that focuses on the social, spatial and temporal aspects of human conversational groups. For example, the works of Hall (1966), Hall et al. (1968), Deutsch (1977, 1979), Ciolek and Kendon (1980) and Kendon (1990) paved the way for understanding the relationship between the positional-orientational behaviour of individuals and the emergence of F-formation systems. In chapter 2, we found that studies of human interaction were based on smaller conversational groups, i.e. dyadic, triadic, or groups of four to six people at the maximum.

Several computational models to simulate F-formations, i.e. as a means of simulating realistic arrangements of conversational groups, have been developed. In chapter 3, we found that a majority of these models use a centroid-based approach, wherein agents coordinate their position and orientation relative to the centroid point of a group. To the best of our knowledge, prior to our work, the agent clusters resulting from the centroid-based approach have not been systematically validated. Therefore, we had a two-point research agenda: (1) to devise an alternative to the centroid-based approach, and (2) to systematically evaluate and validate the spatial patterns of agent clusters resulting from the centroid-based approach, as well as our alternative approach.

In chapter 4, we presented our agent-based simulation platform called Party World, which allows simulating and collecting data pertaining to the spatial manifestation of conversational groups. We implemented Model 1 as the baseline, i.e. the modest version of the Party World simulations, where agents wander randomly, and stop moving when the separation distance is violated. We then implemented two different Party World models. In the first, agents readjust their position and orientation to cover as many neighbours as possible within a 180° frontal view, i.e. Model 2a. In the second, agents readjust their position and orientation with respect to the centroid point, i.e. Model 2b. We also varied the scope of neighbourhood influence on the positional-orientational readjustments of agents. If the scope is local, agents readjust their position and orientation only with respect to those immediately adjacent to themselves (Model 2a LA and Model 2b LA), whereas if the scope is extended, agents perform readjustments with respect to every other agent in the group (Model 2a EA and Model 2b EA).

We compared the working of Model 1, Model 2a LA, Model 2a EA, Model 2b LA and Model 2b EA using detailed qualitative and quantitative analysis. In chapter 5, we analysed the spatial patterns of agent clusters resulting from the different models, by considering an F-formation as the unit of analysis. We also performed a timeline analysis of agent clusters similar to Deutsch (1979) and Kendon (1990). In chapter 6, we identified the different types of dyadic and multi-party agent clusters resulting from the different models. Chapters 7 and 8 focused on validating the Party World models based on observations of conversational groups formed at a drinks reception party, using the DR dataset, which was created exclusively for the purpose of this thesis. Lastly, in chapter 9, we presented findings from a user evaluation study, conducted prior to the development of Models 1, 2a (LA and EA) and 2b (LA and EA).

Our findings showed that the spatial, temporal and social characteristics of agent clusters vary significantly between the different Party World models. Our findings also showed that the currently popular centroid-based approach, particularly the one with an extended neighbourhood influence (Model 2b EA), generates unrealistic spatial patterns and dynamics of conversational groups. On the other hand, the agent clusters resulting from our alternative to the centroid-based approach, i.e. Model 2a, corresponds more closely to real-world conversational groups. The main conclusions of the thesis are:

The simplistic baseline model generates a lot of queue arrangements. Although very basic, Model 1 generates plausible spatial arrangements for groups with fewer than four members. This happens because agents move in the forward direction, and thereby end up in relatively plausible arrangements when they stop moving. However, when there are more than three or four members in a group, the model becomes unrealistic because agents do not readjust their position and orientation, and end up in queue arrangements. The DR dataset had very few queue arrangements on rare occasions. Therefore, we argue that Model 1 is not suitable for generating realistic spatial patterns of conversational groups.

The centroid-based approach is unrealistic. Most existing computational models use the centroid-based approach to simulate conversational groups. Our findings show that this approach does not produce realistic spatial patterns or dynamics of conversational groups. The model always produces a vis-à-vis arrangement for dyads, an equilateral triangle for triads, a square/rhombus for four member groups, a pentagon for five, and a circle for groups of more than five members. The monotony of shapes resulting from the centroid-based model renders it inappropriate for simulating the target phenomenon, because while analysing the DR dataset, we found that conversational groups assume different shapes. Moreover, conversational groups in the DR dataset did not have regular shapes.

The swing-in and swing-out actions of the agents, i.e. when readjusting their position and orientation based on the centroid point, also does not correspond to reality. Furthermore, forming perfectly circular arrangements like the ones resulting from Model 2b EA may not be possible in real life due to space constraints. While analysing the DR dataset, we found that even when the necessary space was available, a properly circular arrangement was never maintained.

Our alternative, the field-of-view model, produces more realistic shapes. Model 2a produces more realistic dyadic arrangements, i.e. ones that are similar to the shapes observed in the DR dataset, as well as to the categories identified by Ciolek and Kendon (1980), such as the H, N, V, C, L and I-shaped arrangements. Model 2a also produces multi-party arrangements whose shapes are similar to the conversational groups in the DR dataset in terms of the roundness of shapes. The spatial behavioural dynamics of agents in Model 2a, i.e. where each agent readjusts its position and orientation to cover as many neighbours as possible within a 180° frontal view, corresponds more closely to the spatial-orientational dynamics of conversational groups in the DR dataset.

Readjustments based on local and extended neighbourhoods. In this thesis, we investigated the effects of agents readjusting their position and orientation based on a local neighbourhood, i.e. immediate neighbours only, versus an extended neighbourhood, i.e. all the agents within a group. To the best of our knowledge, this kind of analysis has not been attempted so far – the centroid-based model has always been used in the context of what we refer to as an extended neighbourhood.

Our findings suggest that the spatial, temporal and social characteristics of agent clusters differ significantly depending on whether agents perform readjustments based on a local neighbourhood, or an extended neighbourhood. The former produces small to medium-sized groups (2 to 6 or 7 members), while the latter produces bigger groups (10 or more members). A related finding was that the former produces more groups with fewer members, while the latter produces fewer groups with more members in each group.

More importantly, the extended neighbourhood influence results in perfectly circular arrangements in the case of Model 2b EA, or one big group in the case of Model 2a EA – neither of these outcomes were found in the DR dataset. Firstly, the perfectly circular arrangements resulting from Model 2b EA, wherein the circle shrinks and expands as people move in and out relative to the centroid point, does not happen in real-world interactions. Secondly, although big groups could be formed at a party, e.g. during an announcement or cake-cutting, the group would eventually disintegrate into smaller groups, but this never happens in Model 2a EA.

In summary, we found that Models 1 and 2b EA perform poorly in comparison to all the other Party World models. Model 2b LA performs better, but Models 2a LA and 2a EA perform the best when taking into consideration the different characteristics of conversational groups in the DR dataset. For example, Model 2a LA results in smaller, variably-shaped agent clusters similar to conversational groups in the DR dataset. Model 2a EA causes smaller groups to aggregate into a bigger group. This phenomenon was observed in the DR dataset, but unlike Model 2a EA, the participants did not all form one big group. Based on this, we argue that a local neighbourhood influence is better than an extended neighbourhood influence to simulate conversational groups. It's also cognitively more plausible that individuals would readjust their position and orientation relative to their immediate neighbours than with respect to every other member in the group. Nonetheless, Model 2a EA could be used to form big groups on specific occasions, e.g. by modelling a triggering event to activate extended neighbourhood influence for some time, before deactivating it again.

Differences between modes. We showed that the simulation and detection of F-formations are both highly influenced by the values chosen for the spacing and orientation of interactants. In this regard, our findings showed that real-world measures of interpersonal distances and orientation cannot be translated as it is to virtual equivalents, particularly when dealing with groups of four or more members. Our finding does not agree with previous reports suggesting that interpersonal distances preferred for face-to-face conversations are the same in real and virtual worlds. Perhaps the issue did not surface because existing studies were carried out with groups of five or fewer members. We argue that systematic assessments, of the sort described in this thesis, are required to identify suitable values for the input parameters, so that the models can produce realistic outcomes.

We have not modelled conversations in Party World models, because the scope of our research is about understanding the role of spatial-orientational behaviour in the functioning of conversational groups. However, Model 2a could be applied or exported to other software packages or frameworks, which have the functionality for dialogue systems. Since conversations can have a major influence on the positional-orientational arrangements of agents, we would need to add new rules to Model 2a, for considering the influence of conversational roles (i.e. speaker, addressee, listener, etc.) on the spatial behaviour of agents. In terms of extending the current Party World models, our findings suggest the following directions:

Modelling a non-empty Party World environment. Conversational groups in social situations such as parties rarely occur in an empty environment. There may be obstacles, furniture, food tables, etc., in the physical space where conversational groups occur. We modelled obstacles in one of our scenarios, but the spatial behavioural rules were the same for obstacles as with the edges of the environment. However, while analysing the spatial patterns of conversational groups in the DR dataset, we found that participants often clustered around the food table. Therefore, we suggest that future versions of Party World simulations could consider meaningful environmental features, and their influence on the spatial arrangements of agent clusters. For example, we could model a food table that attracts agents in the vicinity, i.e. as if to attract people who can see the food on the table. In this scenario, we would expect that as more agents arrive near the table, they would start forming groups or join existing groups.

Modelling the entry and exit of agents. In our Party World models, all twenty agents were present in the environment for the entire duration of the simulations. However, while analysing the DR dataset, we found that the party started with a few people, at one point there were around 15 participants in the Hub, and then, the party ended with only two people remaining in the Hub. That is to say, participants arrived and left at different times during the party. Therefore, future versions of the Party World models could incorporate a similar kind of dynamics for the entry and exit of agents, to model conversational groups in a party context.

Modelling intimates within conversational groups. Our user evaluation study showed that overlapping or very close spatial-orientational arrangements of agents are not seen as inappropriate, but as intimates, e.g. kissing, embracing, couple, etc. Existing models have not considered modelling the spatial-orientational arrangements of intimates within conversational groups, so we propose to investigate how intimates are treated within a conversational group, i.e. in terms of modelling their positional-orientational behaviour with respect to one another and the group.

10.2 Model 2a in a Semi-immersive Virtual Reality Environment

In addition to body position and head orientation, gaze, gesture and posture are other non-verbal behaviours that play a role in conversational groups. We did not consider these aspects in Party World models, as our research questions specifically focused on examining the effectiveness of alternative models, to control the movement, positioning and orientation behaviour of agents. However, it is recognised that several online environments require agents that are more expressive, and capable of making online interactions seem like real-world face-to-face conversations (Gillies et al., 2008). To this end, in addition to the work described in this thesis, we implemented a semi-immersive virtual reality demo, where non-player characters form small groups among themselves, as well as with a human participant. As noted in chapter 3, non-player characters are avatars that are autonomous, and not controlled by a human.

Our demo consists of a life-sized projection screen, a back projector, and a Kinect® motion capture sensor[1]. Whenever a participant arrives in front of the projection screen, the Kinect sensor placed right below the screen starts tracking the position of the participant's body. The skeleton coordinates are fed to a Python program[2] written on the SmartBody[3] character animation platform. The program controls the group formation behaviour of the non-player characters. Figure 10.1 is an image of the demo showing characters interacting (non-verbally) among themselves in small groups, figure 10.2 shows characters interacting among themselves and with a participant.

The part controlling the group formation behaviour of characters (in the Python program) is roughly based on Model 2a, i.e. characters adjust their body position and orientation to cover the participant and as many neighbouring characters as possible. The virtual world is populated with about 8 to 10 characters, each equipped with a movement engine to control navigation in three-dimensional space. At the start of the demo, the characters start walking in random directions, from random initial

[1]http://www.microsoft.com/en-us/kinectforwindows/purchase/
[2]https://www.python.org/
[3]http://smartbody.ict.usc.edu/

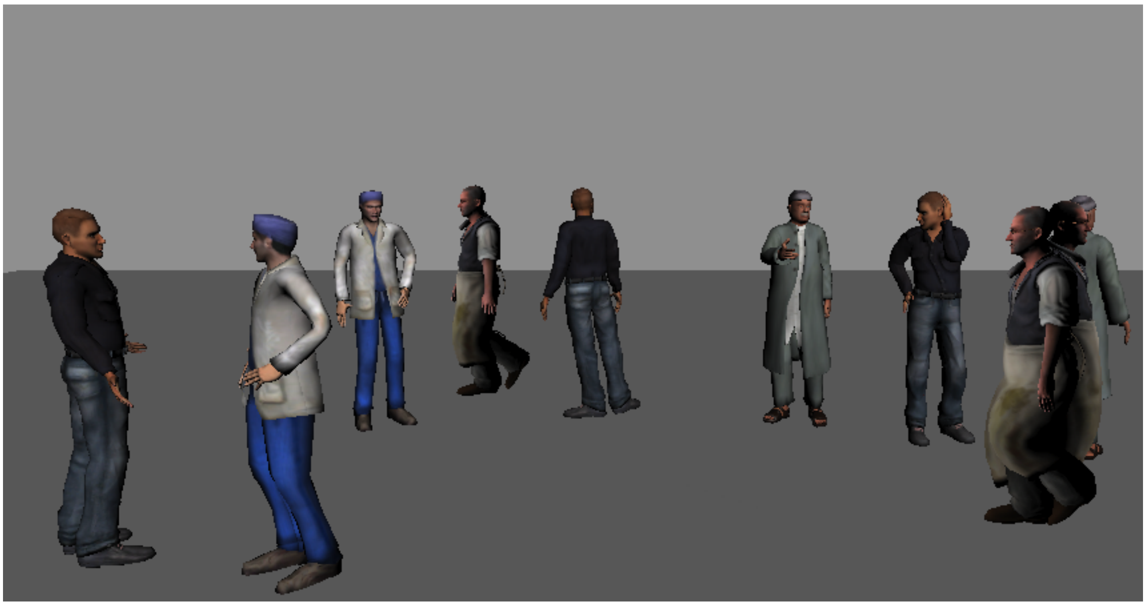


Figure 10.1: Characters interacting among themselves

Figure 10.2: Characters interacting among themselves and with a participant

positions and orientations. As the demo progresses, each character queries the scene manager (a software coordination module in the SmartBody platform), to retrieve the position and orientation of other characters in the system. The characters then calculate the distance between themselves and others. If the distance separating two or more characters is less than a predetermined threshold, they reorganise their position and orientation to cover as many immediate neighbours as possible within a 180° view.

Furthermore, by using a range of body motion scripts available readily in the SmartBody platform, the program controls the gaze and gesture of characters engaged in conversational groups. For example, characters direct their gaze, and wave and point at one another. They also place a hand on their chin (as if to think). The characters start leaving the group after a predetermined time.

The program allows characters to consider the participant as another avatar residing in a different space, i.e. the physical world. Whenever a participant arrives in front of the projection screen, a mapping schema included in the program translates the participant's real-world spatial coordinates (captured using the Kinect sensor) into corresponding virtual-world coordinates. Should a character encounter the participant, i.e. if the virtual-world coordinates of the participant falls within the predetermined separation distance of the character, it would then walk up to the screen to interact with the participant. If more than one character spots the participant, then all of them would walk up to the screen, and form a group with the participant (see image 10.3).

Participants experience the demo in a semi-immersive setup. The visuals on the screen change as the participant moves backwards and forwards and sideways of the projection screen. We did this by setting the field of view of the demo from the eye position of an imaginary avatar, whose position in the virtual-world coordinate system corresponds to the position of the participant in the real-world coordinate system.

Every time a participant steps in front of the projection screen, the first step

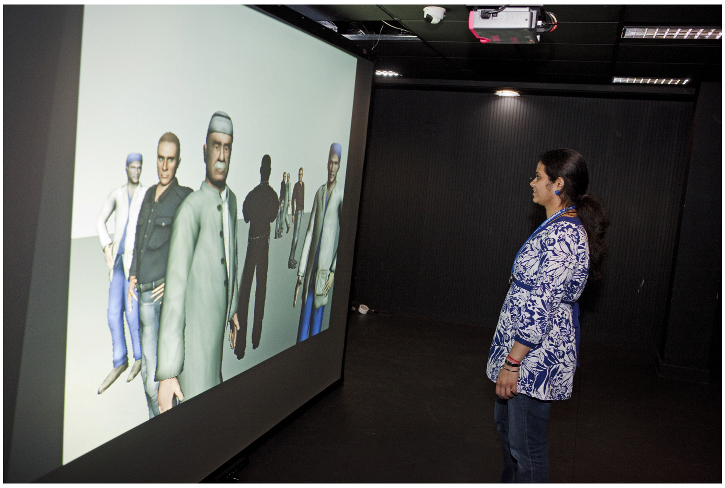


Figure 10.3: Characters engaging with a participant standing in front of the projection screen

is skeleton calibration, which the OpenNI software[4] requires to start tracking the movements of the participant by interfacing with the Kinect sensor. Figure 10.4 shows a participant standing in front of the projection screen, holding an initial calibration pose, for the Kinect sensor (below the screen) to start performing the skeleton tracking. Figure 10.5 shows the real-time skeleton tracking of a participant. Except for instructing participants to restrict their movements within a certain area in front of the projection screen, i.e. to ensure effective tracking, we did not restrict any other aspect of how they interacted with the system.

We exhibited the installation at various showcase events and workshops, such as the Barts and Queen Mary science festival, the Ends of Audience workshop, and the qMedia open studios[5]. During the events, participants reported that they found the interactive experience immersive and engaging. Participants also seemed to respond actively whenever the characters arrived at the screen and tried to engage by waving, pointing, gesturing, etc.

We also received suggestions for improvement. The first suggestion was that the characters could be more friendly. We did not design the look of the characters used in the demo; we used the ones freely available in the SmartBody package. We came

[4] http://structure.io/openni
[5] A video illustration of the demo can be found at https://vimeo.com/46091584

Figure 10.4: Participant's initial calibration pose

Figure 10.5: Real-time skeleton tracking of a participant

to know from the SmartBody development team[6] that the four characters, Brad, Doctor, Utah and Elder, were actually developed for the purpose of a language and culture training system for military use. This could have led to the characters to be regarded as unfriendly.

The second suggestion was to improve the ways in which characters engaged with participants. If participants readjusted their position in the real-world, the characters would also readjust their position and orientation in a corresponding manner, in order to continue facing the users. On the other hand, the waving, pointing and thinking gestures of characters were all implemented in an arbitrary fashion, i.e. characters would wave, point and think randomly, while interacting with the participant. However, participants expected that if they waved, the characters had to respond by waving or pointing. When characters did not respond, because they were not programmed with the ability to detect waving and respond to it, participants found it unusual/unfriendly.

We hope to address the concerns and improve the demo in the future. We hope to address comment 1 by including more friendly characters, especially some female characters[7]. To address comment 2, we hope to develop the software needed to detect and respond to the waving actions of participants. Although not directly linked to the goal of investigating the effect of simulating the spacing and orientation behaviour of characters, we think that addressing comments 1 and 2 would improve the interactive experience of the installation.

We also aim to use the immersive-interactive demo to capture and analyse how participants perceive and navigate among simulated conversational groups. Imagine that we ask participants to navigate across the room to a table where there is food. In the process of accomplishing the task, it will be interesting to see if, and how, participants respond to the conversational groups in the environment. Participants might avoid walking through the simulated groups similar to how they

[6] The experimenter is grateful to Dr. Ari Shapiro and the SmartBody team at the University of Southern California for their help through email exchanges, while developing and setting up the demo for exhibit.

[7] A couple of users had specifically noted about including female characters in the demo.

would avoid interrupting conversational groups in real life. Or, if they try to join the conversational groups, it will be interesting to see how they did it, i.e. if their actions are similar to what we would expect to see in real-world situations. Cafaro et al. (2016) have carried out a similar investigation, but they have used the centroid-based approach to simulate conversational groups, which, as we have seen from the work in this thesis, does not accurately model real conversational groups. Therefore, one avenue for future research could implement both Model 2a and Model 2b in the aforementioned context, to assess if, and how, the behaviour of participants varies with respect to the models.

10.3 Concluding remark

Railsback and Grimm (2011) remark that, the success of an agent-based (or individual-based) model is not so much in just formulating the model, but in analysing it. Reynolds (1987) proposed the seminal boid model for automating the steering behaviour of agents, whereas Huth and Wissel (1992) also validated their fish schooling model using data on real fish schools. In this thesis, we hope to have used scientifically robust methods to formulate, implement, evaluate and validate the different Party World models. There are no precedents of other agent-based models of human conversational groups that have done this. The contributions of this thesis have implications for models of face-to-face interaction used in intelligent virtual environments, human-agent interaction systems, intelligent surveillance systems, and artificial life systems. In addition, the Party World simulation platform and its models are potential test beds, for clarifying and supporting theories about the role of spatial behaviour in the functioning of conversational groups.

Appendix A

Full tables for Statistical Analysis

Model	HVFF	GCFF	PWFD
1	3.16%	3.12%	2.87%
2a LA	2.15%	2.14%	3.02%
2b LA	4.31%	2.75%	6.21%
2a EA	10.08%	6.54%	8.12%
2b EA	11.74%	11.39%	18.71%

Table A.1: Coefficient of variation in F-formations detected over five runs

Model	HVFF		GCFF		PWFD	
	Statistic (w)	p value	Statistic (w)	p value	Statistic (w)	p value
1	0.9486	0.7271	0.9681	0.8627	0.9106	0.4714
2a LA	0.7972	0.07694	0.8453	0.18	0.8033	0.08624
2b LA	0.9297	0.5945	0.9194	0.5258	0.8417	0.1698
2a EA	0.8374	0.1579	0.8003	0.08144	0.8615	0.2338
2b EA	0.9446	0.6987	0.87	0.2664	0.7527	0.03151**

** Value less than 0.05, denoting that sample deviates from normality.

Table A.2: Summary of Shapiro-Wilk's normality test for the number of F-formations detected across the different Party World models

Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	-38.0	-316.71266	240.7127	0.9937029*
Model 2b LA - Model 1	362.0	83.28734	640.7127	0.0072633
Model 2a EA - Model 1	-137.2	-415.91266	141.5127	0.5904882*
Model 2b EA - Model 1	-52.2	-330.91266	226.5127	0.9792998*
Model 2b LA - Model 2a LA	400.0	121.28734	678.7127	0.0029020
Model 2a EA - Model 2a LA	-99.2	-377.91266	179.5127	0.8218653*
Model 2b EA - Model 2a LA	-14.2	-292.91266	264.5127	0.9998671*
Model 2a EA - Model 2b LA	-499.2	-777.91266	-220.4873	0.0002635
Model 2b EA - Model 2b LA	-414.2	-692.91266	-135.4873	0.0020558
Model 2b EA - Model 2a EA	85.0	-193.71266	363.7127	0.8887120*

Table A.3: Tukey multiple comparisons of means (HVFF)

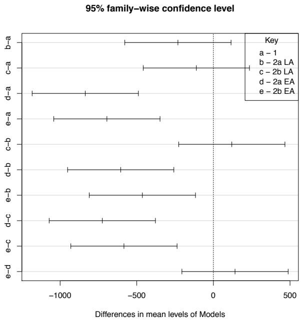
Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	-30.4	-270.6721	209.8721	0.9952705*
Model 2b LA - Model 1	382.8	142.5279	623.0721	0.0009960
Model 2a EA - Model 1	-63.2	-303.4721	177.0721	0.9313925*
Model 2b EA - Model 1	-22.4	-262.6721	217.8721	0.9985546*
Model 2b LA - Model 2a LA	413.2	172.9279	653.4721	0.0004247
Model 2a EA - Model 2a LA	-32.8	-273.0721	207.4721	0.9936727*
Model 2b EA - Model 2a LA	8.0	-232.2721	248.2721	0.9999756*
Model 2a EA - Model 2b LA	-446.0	-686.2721	-205.7279	0.0001710
Model 2b EA - Model 2b LA	-405.2	-645.4721	-164.9279	0.0005311
Model 2b EA - Model 2a EA	40.8	-199.4721	281.0721	0.9855853*

Table A.4: Tukey multiple comparisons of means (GCFF)

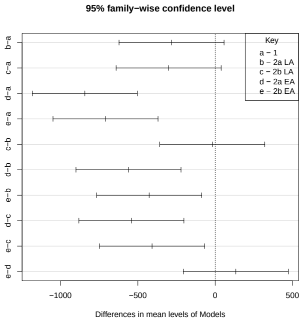
Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	59.2	-265.938371	384.33837	0.9813384*
Model 2b LA - Model 1	468.0	142.861629	793.13837	0.0028204
Model 2a EA - Model 1	-394.2	-719.338371	-69.06163	0.0128895
Model 2b EA - Model 1	-71.4	-396.538371	253.73837	0.9632492*
Model 2b LA - Model 2a LA	408.8	83.661629	733.93837	0.0095778
Model 2a EA - Model 2a LA	-453.4	-778.538371	-128.26163	0.0038198
Model 2b EA - Model 2a LA	-130.6	-455.738371	194.53837	0.7502646*
Model 2a EA - Model 2b LA	-862.2	-1187.338371	-537.06163	0.0000012
Model 2b EA - Model 2b LA	-539.4	-864.538371	-214.26163	0.0006388
Model 2b EA - Model 2a EA	322.8	-2.338371	647.93837	0.0522464***

Table A.5: Tukey multiple comparisons of means (PWFD)

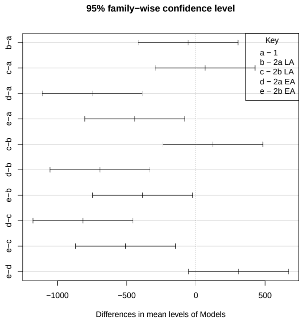
* p-value>0.05, hence not considered as significant. *** p-value marginally >0.05.



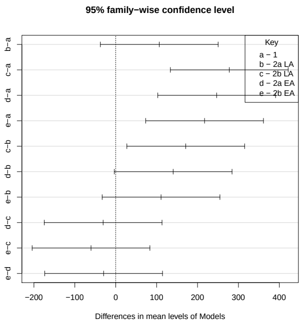
(a) 2 agent cluster GCFF



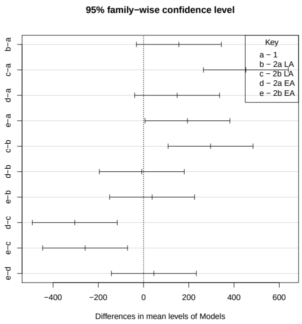
(b) 2 agent cluster HVFF



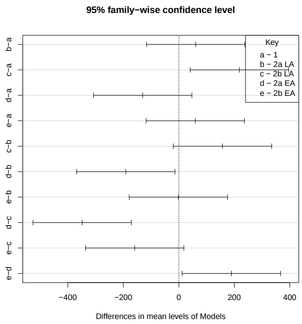
(c) 2 agent cluster PWFD



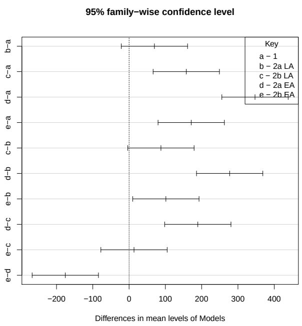
(d) 3 agent cluster GCFF



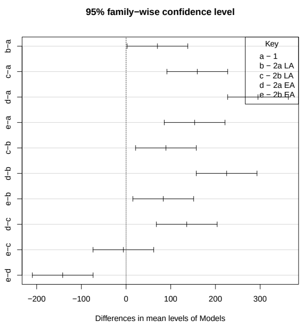
(e) 3 agent cluster HVFF



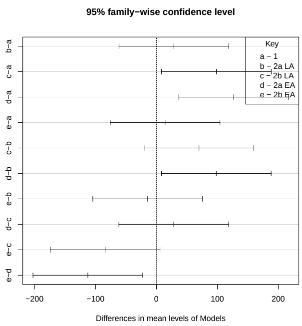
(f) 3 agent cluster PWFD



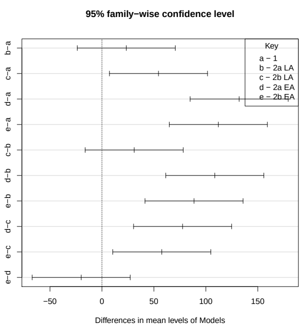
(g) 4 agent cluster GCFF



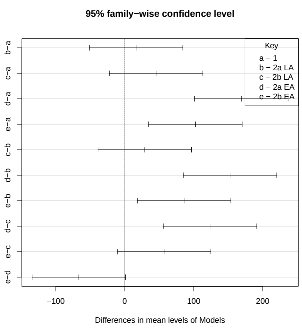
(h) 4 agent cluster HVFF



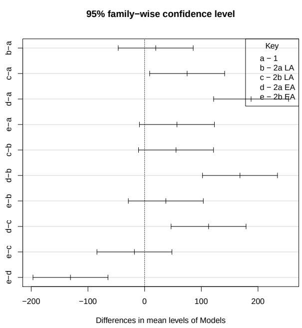
(i) 4 agent cluster PWFD



(j) 5 agent cluster GCFF



(k) 5 agent cluster HVFF



(l) 5 agent cluster PWFD

Figure A.1: Results of the Tukey Kramer test showing variations across models in the size of F-formations

Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	-16.4	-159.6052	126.80518	0.9967800*
Model 2b LA - Model 1	325.6	182.3948	468.80518	0.0000117
Model 2a EA - Model 1	-41.2	-184.4052	102.00518	0.9077105*
Model 2b EA - Model 1	-84.8	-228.0052	58.40518	0.4159887*
Model 2b LA - Model 2a LA	342.0	198.7948	485.20518	0.0000058
Model 2a EA - Model 2a LA	-24.8	-168.0052	118.40518	0.9844931*
Model 2b EA - Model 2a LA	-68.4	-211.6052	74.80518	0.6169802*
Model 2a EA - Model 2b LA	-366.8	-510.0052	-223.59482	0.0000021
Model 2b EA - Model 2b LA	-410.4	-553.6052	-267.19482	0.0000004
Model 2b EA - Model 2a EA	-43.6	-186.8052	99.60518	0.8893070*

Table A.6: Tukey multiple comparisons of means (HVFF) – models with obstacles

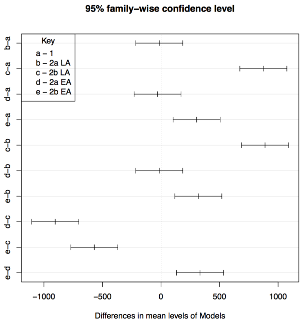
Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	-17.8	-155.1301	119.53011	0.9948113*
Model 2b LA - Model 1	354.6	217.2699	491.93011	0.0000018
Model 2a EA - Model 1	-33.2	-170.5301	104.13011	0.9485467*
Model 2b EA - Model 1	-57.6	-194.9301	79.73011	0.7202811*
Model 2b LA - Model 2a LA	372.4	235.0699	509.73011	0.0000009
Model 2a EA - Model 2a LA	-15.4	-152.7301	121.93011	0.9970318*
Model 2b EA - Model 2a LA	-39.8	-177.1301	97.53011	0.9054916*
Model 2a EA - Model 2b LA	-387.8	-525.1301	-250.46989	0.0000005
Model 2b EA - Model 2b LA	-412.2	-549.5301	-274.86989	0.0000002
Model 2b EA - Model 2a EA	-24.4	-161.7301	112.93011	0.9829492*

Table A.7: Tukey multiple comparisons of means (GCFF) – models with obstacles

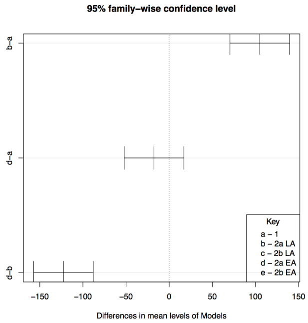
Models	diff	lwr	upr	p adj
Model 2a LA - Model 1	67.2	-159.7107	294.1107	0.8986592*
Model 2b LA - Model 1	397.8	170.8893	624.7107	0.0003395
Model 2a EA - Model 1	-351.0	-577.9107	-124.0893	0.0013626
Model 2b EA - Model 1	-5.8	-232.7107	221.1107	0.9999915*
Model 2b LA - Model 2a LA	330.6	103.6893	557.5107	0.0025043
Model 2a EA - Model 2a LA	-418.2	-645.1107	-191.2893	0.0001866
Model 2b EA - Model 2a LA	-73.0	-299.9107	153.9107	0.8684673*
Model 2a EA - Model 2b LA	-748.8	-975.7107	-521.8893	0.0000000
Model 2b EA - Model 2b LA	-403.6	-630.5107	-176.6893	0.0002862
Model 2b EA - Model 2a EA	345.2	118.2893	572.1107	0.0016200

Table A.8: Tukey multiple comparisons of means (PWFD) – models with obstacles

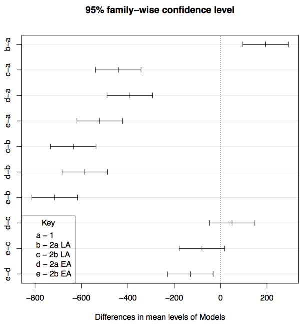
* p-value>0.05, hence not considered as significant. *** p-value marginally >0.05.



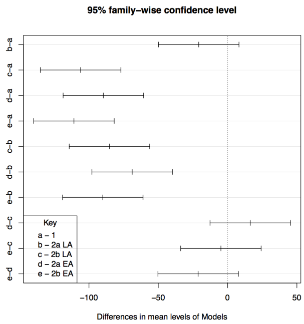
(a) H-shaped



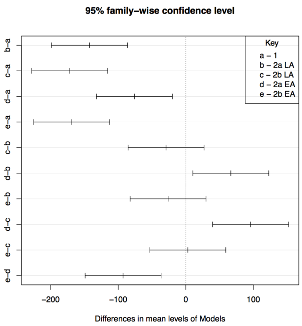
(b) N-shaped



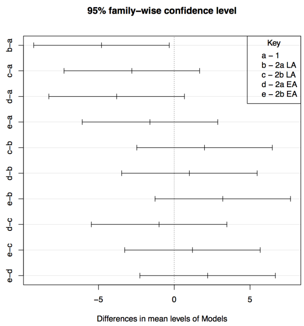
(c) V-shaped



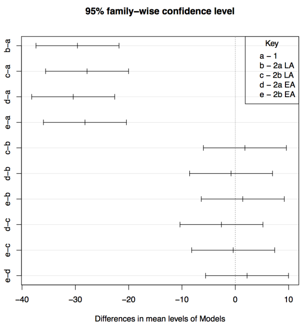
(d) L-shaped



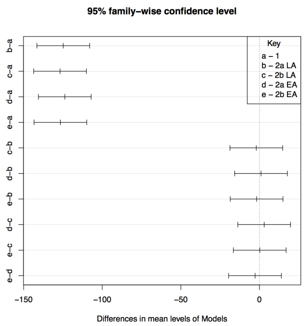
(e) C-shaped



(f) I-shaped

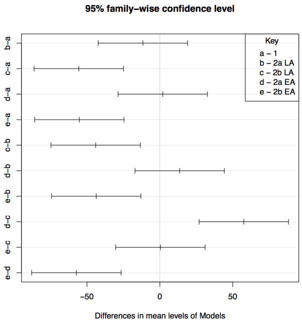


(g) Queue

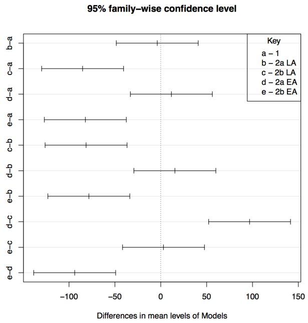


(h) Complicated queue

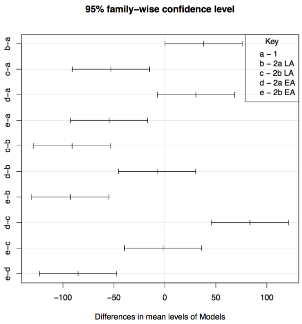
Figure A.2: Results of the Tukey Kramer test showing variations across models in dyadic arrangements



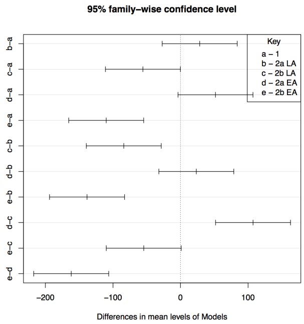
(a) 0.1-0.2



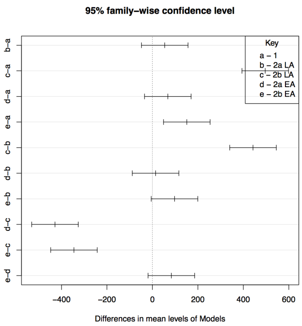
(b) 0.2-0.3



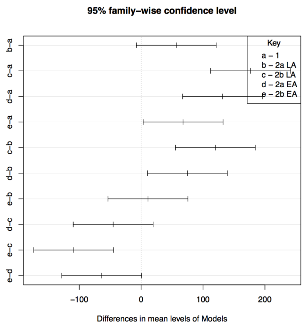
(c) 0.3-0.4



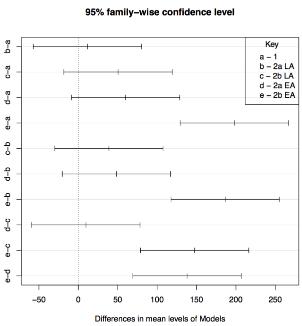
(d) 0.4-0.5



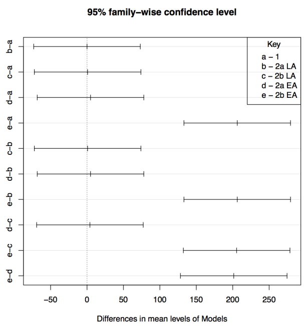
(e) 0.5-0.6



(f) 0.6-0.7



(g) 0.6-0.8



(h) 0.8-0.9

Figure A.3: Results of the Tukey Kramer test showing variations across models in multi-party F-formations

Comparisons of Column Proportions[b]

		Model					
		DR	M1	M2aEA	M2aLA	M2bEA	M2bLA
		(A)	(B)	(C)	(D)	(E)	(F)
Shape	C-shaped	D E F	A D E F	A B D E F	F	F	
	L-shaped	B C D E F	C E F	E F	E F		
	N-shaped	B F	F	B F	A B C F	.[a]	
	queue	D F	A C D E F	D F		D F	
	side-by-side	B C D E F	D	D		D F	
	V-shaped	E F	A E F	A E F	A B C E F		E
	vis-a-vis	B C D	D			A B C D	A B C D

Results are based on two-sided tests with significance level .05. For each significant pair, the key of the category with the smaller column proportion appears under the category with the larger column proportion.

a. This category is not used in comparisons because its column proportion is equal to zero or one.

b. Tests are adjusted for all pairwise comparisons within a row of each innermost subtable using the Bonferroni correction.

Figure A.4: Results of z-test comparing proportions of dyadic F-formations in the Party World and DR dataset

Bibliography

Argyle, M. and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 28(3):289–304.

Argyle, M. and Kendon, A. (1967). The experimental analysis of social performance. *Advances in experimental social psychology*, 3:55–98.

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7):819–833.

Bazzani, L., Cristani, M., Tosato, D., Farenzena, M., Paggetti, G., Menegaz, G., and Murino, V. (2013). Social interactions by visual focus of attention in a three-dimensional environment. *Expert Systems*, 30(2):115–127.

Becker, F. D. (1973). Study of spatial markers. *Journal of Personality and Social Psychology*, 26(3):439–445.

Birdwhistell, R. (1968). Kinesics. in international encyclopedia of the social sciences (vol. 8, pp. 379-385).

Birdwhistell, R. L. (1952). *Introduction to kinesics: an annotation system for analysis of body motion and gesture.* University of Louisville Press.

Birdwhistell, R. L. (1970). *Kinesics and context: essays on body motion communication.* University of Pennsylvania Press Philadelphia.

Bradski, D. G. R. (2000). Camera calibration with opencv. `http://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html`. [Online; accessed 29-August-2016].

Bradski, D. G. R. and Kaehler, A. (2008). *Learning Opencv, 1st Edition*. O'Reilly Media, Inc., first edition.

Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. MIT press.

Brom, C., Babor, P., Popelová, M., Bída, M., Tomek, J., and Gemrot, J. (2012). Controlling three agents in a quarrel: Lessons learnt. In *Motion in Games*, pages 158–169. Springer.

Cafaro, A., Ravenet, B., Ochs, M., Vilhjálmsson, H., and Pelachaud, C. (2016). The effects of interpersonal attitude of a group of agents on user's presence and proxemics behavior. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6(2):1–33.

Cafaro, A., Vilhjálmsson, H., Bickmore, T., Heylen, D., Jóhannsdóttir, K. R., and Valgarðsson, G. S. (2012). First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *Intelligent Virtual Agents*, pages 67–80. Springer.

Cassell, J. and Vilhjálmsson, H. (1999). Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous agents and multi-agent systems*, 2(1):45–64.

Cheyne, J. A. and Efran, M. G. (1972). The effect of spatial and interpersonal variables on the invasion of group controlled territories. *Sociometry*, 35(3):477–489.

Ciolek, T. (1983). The proxemics lexicon: A first approximation. *Journal of Non-verbal Behavior*, 8(1):55–79.

Ciolek, T. M. and Kendon, A. (1980). Environment and the spatial arrangement of conversational encounters. *Sociological Inquiry*, 50(3-4):237–271.

Condon, W. S. and Ogston, W. D. (1967). A segmentation of behavior. *Journal of psychiatric research*, 5(3):221–235.

Cook, M. (1970). Experiments on orientation and proxemics. *Human Relations*.

Cristani, M., Bazzani, L., Paggetti, G., Fossati, A., Tosato, D., Del Bue, A., Menegaz, G., and Murino, V. (2011a). Social interaction discovery by statistical analysis of f-formations. In *BMVC*, pages 1–12.

Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., and Murino, V. (2011b). Towards computational proxemics: Inferring social relations from interpersonal distances. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, pages 290–297. IEEE.

Damian, I., Endrass, B., Huber, P., Bee, N., and André, E. (2011). Individualized agent interactions. In *Motion in games*, pages 15–26. Springer.

Deutsch, R. D. (1977). *Spatial structurings in everyday face-to-face behavior: a neurocybernetic model*. Association for the Study of Man-Environment Relations.

Deutsch, R. D. (1979). On the isomorphic structure of endings: An example from everyday face-to-face interaction and balinese legong dance. *Ethology and Sociobiology*, 1(1):41–57.

Efran, M. G. and Cheyne, J. A. (1973). Shared space: The co-operative control of spatial areas by two interacting individuals. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 5(3):201.

Ennis, C. and O'Sullivan, C. (2012). Perceptually plausible formations for virtual conversers. *Computer Animation and Virtual Worlds*, 23(3-4):321–329.

Felipe, N. J. and Sommer, R. (1966). Invasions of personal space. *Social problems*, 14(2):206–214.

Ford, J. G., Knight, M., and Cramer, R. (1977). The phenomenological experience of interpersonal spacing. *Sociometry*, 40(4):387–390.

Friedman, D., Steed, A., and Slater, M. (2007). Spatial social behavior in second life. In *Intelligent virtual agents*, pages 252–263. Springer.

Gan, T., Wong, Y., Zhang, D., and Kankanhalli, M. S. (2013). Temporal encoded f-formation system for social interaction detection. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 937–946. ACM.

Gilbert, N. (1993). Computer simulation of social processes. *Social research update*, 6:1.

Gilbert, N. (2008). *Agent-based models*. Sage.

Gilbert, N. and Troitzsch, K. (2005). *Simulation for the social scientist*. McGraw-Hill International.

Gill, S. P., Kawamori, M., Katagiri, Y., and Shimojima, A. (2000). The role of body moves in dialogue. *International Journal of Language and Communication*, 12:89–114.

Gillies, M., Ballin, D., Pan, X., and Dodgson, N. (2008). Semi-autonomous avatars: A new direction for expressive user embodiment. In Cañamero, L. and Aylett, R., editors, *Animating Expressive Characters for Social Interaction*, pages 235–255. John Benjamins Publishing Company.

Goffman, E. (1961). *Encounters: Two studies in the sociology of interaction*. Bobbs-Merrill.

Goffman, E. (1963). *Behaviour in public places: notes on the social order of gatherings*. The Free Press, New York.

Goffman, E. (1971). *Relations in public: microstudies of the public order*. New York: Basic Books.

Goffman, E. (1983). The interaction order: American sociological association, 1982 presidential address. *American Sociological Review*, 48(1):1–17.

Hall, E. T. (1963). A system for the notation of proxemic behavior. *American Anthropologist*, 65(5):1003–1026.

Hall, E. T. (1964). Silent assumptions in social communication. *Research publications, Association for Research in Nervous and Mental Disease*, 42:41–55.

Hall, E. T. (1966). *Distances in man: The hidden dimension*. Double Day, Garden City, New York.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold Jr, A. R., Durbin, M., Edmonson, M. S., Fischer, J., Hymes, D., Kimball, S. T., et al. (1968). Proxemics [and comments and replies]. *Current anthropology*, pages 83–108.

Hayduk, L. A. (1978). Personal space: An evaluative and orienting overview. *Psychological Bulletin*, 85(1):117.

Hediger, H. (1961). The evolution of territorial behavior. *Social Life of Early Man*.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282–4286.

Helbing, D., Molnár, P., and Schweitzer, F. (1994). Computer simulations of pedestrian dynamics and trail formation. *Evolution of natural structures*, pages 229–234.

Huth, A. and Wissel, C. (1992). The simulation of the movement of fish schools. *Journal of theoretical biology*, 156(3):365–385.

Jan, D., Herrera, D., Martinovski, B., Novick, D., and Traum, D. (2007). A computational model of culture-specific conversational behavior. In *Intelligent virtual agents*, pages 45–56. Springer.

Jan, D. and Traum, D. R. (2005). Dialog simulation for background characters. In *Intelligent Virtual Agents*, pages 65–74. Springer.

Jan, D. and Traum, D. R. (2007). Dynamic movement and positioning of embodied agents in multiparty conversations. In *Proceedings of the Workshop on Embodied Language Processing*, pages 59–66. Association for Computational Linguistics.

Jarvis, R. A. (1973). On the identification of the convex hull of a finite set of points in the plane. *Information Processing Letters*, 2(1):18–21.

Jeffrey, P. and Mark, G. (1998). Constructing social spaces in virtual environments: A study of navigation and interaction. In *Workshop on personalised and social navigation in information space*, pages 24–38. Swedish Institute of Computer Science Stockholm.

Karimaghalou, N., Bernardet, U., and DiPaola, S. (2014). A model for social spatial behavior in virtual characters. *Computer Animation and Virtual Worlds*, 25(3-4):505–517.

Kendon, A. (1970). Movement coordination in social interaction: Some examples described. *Acta psychologica*, 32:101–125.

Kendon, A. (1973). The role of visible behaviour in the organization of social interaction. *Social communication and movement: Studies of interaction and expression in man and chimpanzee.*

Kendon, A. (1977). *Studies in the behavior of social interaction.* Humanities Press Intl.

Kendon, A. (1978). Spatial organization in social encounters: the facing-formation system. *Unpublished manuscript. Later published in Studies in the Behavior of Social Interaction.*

Kendon, A. (1979). Some emerging features of face-to-face interaction studies. *Sign language studies*, 22(1):7–22.

Kendon, A. (1988). Goffman's approach to face-to-face interaction. *Erving Goffman: Exploring the interaction order*, pages 14–40.

Kendon, A. (1990). Spatial organization in social encounters: The f-formation system. *Conducting interaction: Patterns of behavior in focused encounters*, pages 209–238.

Kendon, A. (2010). Spacing and orientation in co-present interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony*, pages 1–15. Springer.

Kistler, F., Endrass, B., Damian, I., Dang, C. T., and André, E. (2012). Natural interaction with culturally adaptive virtual characters. *Journal on Multimodal User Interfaces*, 6(1-2):39–47.

Knowles, E. S. (1973). Boundaries around group interaction: The effect of group size and member status on boundary permeability. *Journal of Personality and Social Psychology*, 26(3):327–331.

Koay, K. L., Syrdal, D. S., Ashgari-Oskoei, M., Walters, M. L., and Dautenhahn, K. (2014). Social roles and baseline proxemic preferences for a domestic service robot. *International Journal of Social Robotics*, 6(4):469–488.

Kuzuoka, H., Suzuki, Y., Yamashita, J., and Yamazaki, K. (2010). Reconfiguring spatial formation arrangement by robot body orientation. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pages 285–292. IEEE Press.

Laga, H. and Amaoka, T. (2009). Modeling the spatial behavior of virtual agents in groups for non-verbal communication in virtual worlds. In *Proceedings of the 3rd International Universal Communication Symposium*, pages 154–159. ACM.

Lakshika, E., Barlow, M., and Easton, A. (2012). Fidelity and complexity of standing group conversation simulations: A framework for the evolution of multi agent systems through bootstrapping human aesthetic judgments. In *Evolutionary Computation (CEC), 2012 IEEE Congress on*, pages 1–8. IEEE.

Lindner, F. and Eschenbach, C. (2011). Towards a formalization of social spaces for socially aware robots. In *Spatial Information Theory*, pages 283–303. Springer.

Lindskold, S., Albert, K. P., Baer, R., and Moore, W. C. (1976). Territorial boundaries of interacting groups and passive audiences. *Sociometry*, 39(1):71–76.

Llobera, J., Spanlang, B., Ruffini, G., and Slater, M. (2010). Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception (TAP)*, 8(1):3:1–3:12.

Lyman, S. M. and Scott, M. B. (1967). Territoriality: A neglected sociological dimension. *Social problems*, 15(2):236–249.

Macal, C. M. and North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of simulation*, 4(3):151–162.

Macy, M. W. and Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology*, 28:143–166.

Marshall, P., Rogers, Y., and Pantidi, N. (2011). Using f-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 445–454. ACM.

MATLAB and Toolbox, S. (2014). Matlab [computer software]. geometric transformations. `http://uk.mathworks.com/help/images/geometric-transformations.html`. [Online; accessed 29-August-2016].

McDermott, R. P. and Roth, D. R. (1978). The social organization of behavior: Interactional approaches. *Annual Review of Anthropology*, 7:321–345.

McDonald, J. H. (2009). *Handbook of biological statistics*. Sparky House Publishing Baltimore, MD.

McDowell, K. V. (1972). Violations of personal space. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 4(3):210.

Mead, R. and Matarić, M. J. (2011). An experimental design for studying proxemic behavior in human-robot interaction. Technical report, Technical Report CRES-11-001, USC Interaction Lab, Los Angeles.

Mead, R. and Mataric, M. J. (2012). A probabilistic framework for autonomous proxemic control in situated and mobile human-robot interaction. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '12, pages 193–194. ACM.

Mehrabian, A. (1968). Inference of attitudes from the posture, orientation, and distance of a communicator. *Journal of Consulting and Clinical Psychology*, 32(3):296–308.

Narasimhan, K. P. and White, G. (2014). Agent clusters: The usual vs. the unusual. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 244–255. Springer.

Nguyen, N. and Wachsmuth, I. (2011). From body space to interaction space: modeling spatial cooperation for virtual humans. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1047–1054. International Foundation for Autonomous Agents and Multiagent Systems.

Nguyen, N. and Wachsmuth, I. (2013). A computational model of cooperative spatial behaviour for virtual humans. *Representing Space in Cognition: Interrelations of behaviour, language, and formal models*, pages 147–168.

Patterson, M. and Edinger, J. A. (1987). A functional analysis of space in social interaction. *Nonverbal behavior and communication*, pages 523–562.

Pedica, C. and Vilhjálmsson, H. (2008). Social perception and steering for online avatars. In *Intelligent Virtual Agents*, pages 104–116. Springer.

Pedica, C. and Vilhjálmsson, H. (2010). Spontaneous avatar behavior for human territoriality. *Applied Artificial Intelligence*, 24(6):575–593.

Pedica, C. and Vilhjálmsson, H. (2012). Lifelike interactive characters with behavior trees for social territorial intelligence. In *ACM SIGGRAPH 2012 Posters*, SIGGRAPH '12. ACM.

Pedica, C., Vilhjálmsson, H., and Kristinsson, K. V. (2015). Study of nine people in a hallway. `http://www.ru.is/faculty/hannes/publications/MIG2015-Abstract.pdf`. [Online; accessed 10-August-2016].

Pedica, C., Vilhjálmsson, H., and Lárusdóttir, M. (2010). Avatars in conversation: The importance of simulating territorial behavior. In Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., and Safonova, A., editors, *Intelligent Virtual Agents*, volume 6356 of *Lecture Notes in Computer Science*, pages 336–342. Springer Berlin Heidelberg.

Railsback, S. F. and Grimm, V. (2011). *Agent-based and individual-based modeling: a practical introduction*. Princeton university press.

Reeves, W. T. (1983). Particle systems—a technique for modeling a class of fuzzy objects. *ACM Transactions on Graphics*, 2(2):91–108.

Rehm, M., André, E., and Nischt, M. (2005). Let's come together—social navigation behaviors of virtual and real humans. In *Intelligent Technologies for Interactive Entertainment*, pages 124–133. Springer.

Reynolds, C. (2001). Boids: Background and update. `URL:www.red3d.com/cwr/boids`. [Online; accessed 21-04-2014].

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *ACM Siggraph Computer Graphics*, volume 21, pages 25–34. ACM.

Reynolds, C. W. (1988). Not bumping into things. *SIGGRAPH course 27 notes: Developements in Physically-Based Modeling*, pages G1–G13.

Reynolds, C. W. (1999). Steering behaviors for autonomous characters. In *Game developers conference*, pages 763–782.

Reynolds, C. W. (2000). Interaction with groups of autonomous characters. In *Game Developers Conference*, volume 2000, pages 449–460.

Ricks, B. C. and Egbert, P. K. (2012). More realistic, flexible, and expressive social crowds using transactional analysis. *The Visual Computer*, 28(6-8):889–898.

Riesman, D., Potter, R. J., and Watson, J. (1960). The vanishing host. *Human Organization*, 19(1):17–27.

Rojas, F., Tarnogol, F., and Yang, H. S. (2016). Dynamic social formations of pedestrian groups navigating and using public transportation in a virtual city. *The Visual Computer*, 32(3):335–345.

Rojas, F. A., Yang, H. S., and Tarnogol, F. M. (2014). Safe navigation of pedestrians in social groups in a virtual urban environment. In *Cyberworlds (CW), 2014 International Conference on*, pages 31–38. IEEE.

Salem, B. and Earle, N. (2000). Designing a non-verbal language for expressive avatars. In *Proceedings of the third international conference on Collaborative virtual environments*, pages 93–101. ACM.

Scheflen, A. E. (1964). The significance of posture in communication systems. *Psychiatry*, 27(4):316–331.

Scheflen, A. E. (1975). Micro-territories in human interaction. *Organization of Behavior in Face-to-Face Interaction*, pages 159–173.

Scheflen, A. E. and Ashcraft, N. (1976). *Human territories: How we behave in space-time.* Prentice-Hall.

Setti, F., Hung, H., and Cristani, M. (2013). Group detection in still images by f-formation modeling: A comparative study. In *Image Analysis for Multimedia Interactive Services (WIAMIS), 2013 14th International Workshop on*, pages 1–4. IEEE.

Setti, F., Russell, C., Bassetti, C., and Cristani, M. (2014). F-formation detection: Individuating free-standing conversational groups in images. *PloS one*, 10(5):1–26.

Setti, F., Russell, C., Bassetti, C., and Cristani, M. (2015). Correction: F-formation detection: Individuating free-standing conversational groups in images. *PloS one*, 10(9).

Shi, C., Shimada, M., Kanda, T., Ishiguro, H., and Hagita, N. (2011). Spatial formation model for initiating conversation. In *Robotics: science and systems*, volume 11.

Sommer, R. (1962). The distance for comfortable conversation: A further study. *Sociometry*, 25(1):111–116.

Sommer, R. (1965). Further studies of small group ecology. *Sociometry*, 28(4):337–348.

Steinzor, B. (1950). The spatial factor in face to face discussion groups. *The Journal of Abnormal and Social Psychology*, 45(3):552–555.

Sussman, N. M. and Rosenfeld, H. M. (1982). Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology*, 42(1):66.

Swartout, W. R., Gratch, J., Hill, R. W., Hovy, E., Marsella, S., Rickel, J., and Traum, D. (2006). Toward virtual humans. *AI Magazine - Special issue on achieving human-level AI through integrated systems and research*, 27(2):96–108.

Talbot, C. and Youngblood, G. M. (2013a). Application of force-directed graphs on character positioning. In *Spatial Computing Workshop (colocated with AAMAS 2013)*, pages 53–58.

Talbot, C. and Youngblood, G. M. (2013b). Positioning characters using forces. In *Cognitive Agents for Virtual Environments (CAVE 2013)*.

Tosato, D., Spera, M., Cristani, M., and Murino, V. (2013). Characterizing humans on riemannian manifolds. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1972–1984.

Vázquez, M., Steinfeld, A., and Hudson, S. E. (2015). Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 3010–3017. IEEE.

Vilhjalmsson, H., Merchant, C., and Samtani, P. (2007). Social puppets: Towards modular social animation for agents and avatars. In Schuler, D., editor, *Online*

Communities and Social Computing, volume 4564 of *Lecture Notes in Computer Science*, pages 192–201. Springer Berlin Heidelberg.

Vine, I. (1975). Territoriality and the spatial regulation of interaction. *Organization of behavior in face-to-face interaction*, pages 357–388.

von Cranach, M. (1971). The role of orienting behavior in human interaction. In *Behavior and Environment*, pages 217–237. Springer.

Vroon, J., Joosse, M., Lohse, M., Kolkmeier, J., Kim, J., Truong, K., Englebienne, G., Heylen, D., and Evers, V. (2015). Dynamics of social positioning patterns in group-robot interactions. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*, pages 394–399. IEEE.

Watson, O. M. and Graves, T. D. (1966). Quantitative research in proxemic behavior1. *American Anthropologist*, 68(4):971–985.

Wescott, R. W. (1966). Introducing coenetics: A biosocial analysis of communication. *The American Scholar*, 35(2):342–356.

Wilensky, U. and Rand, W. (2007). Making models match: Replicating an agent-based model. *Journal of Artificial Societies and Social Simulation*, 10(4):2.

Yamaoka, F., Kanda, T., Ishiguro, H., and Hagita, N. (2010). A model of proximity control for information-presenting robots. *Robotics, IEEE Transactions on*, 26(1):187–195.

Zhang, L. and Hung, H. (2016). Beyond f-formations: Determining social involvement in free standing conversing groups from static images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1086–1095.