Modeling Agent Social Joint Actions via Micro and Macro Coordination Strategies

Seung In Park, Francis Quek, Yong Cao
Department of Computer Science
Virginia Tech
Blacksburg, USA
Email: {spark80, quek, yongcao}@vt.edu

Abstract—This paper presents a multi-agent model for large crowd simulations that addresses the need for socially plausible coordination behavior. A computational model for multi-agent coordination informed by well-established common ground theory is proposed. We introduce the idea of macro- and micro-coordination strategies that allow agent-based simulations to adapt to different domains. Our agent model allows the selection of appropriate behaviors based on the spatiotemporal conditions of the agent-group's environment. By showing that different micro-coordination strategies of individual groups has an influence on the overall distribution of a crowd, we demonstrate the importance of incorporating such models into multi-agent simulations of large crowd behaviors.

Keywords-coordination, crowd simulation, common ground theory.

I. Introduction

This paper presents a multi-agent model for large crowd simulations that addresses the need for socially plausible coordination behavior. Given that crowd dynamics emerges from patterns of social interactions among groups of pedestrians [1], [2], employing such a coordination model in crowd simulations may produce significant differences in overall crowd circulation patterns. The 'intelligence' of groups differs in a sense from individual intelligence because it is an emergent phenomena driven by the need of the groups to maintain cohesive activity.

We employ a model from social-psychology and linguistics due to Herbert Clark in which members of a group must negotiate and maintain a state of common ground (CG) as a precondition of joint activity [3], [4]. CG maintenance is not solely a linguistic activity. Behavior comes into play whereby agents 'display their intentions' visually, and mutual behavior awareness plays an important role.

Our model includes two kinds of strategies by which behaviors of groups may be embedded into a larger crowd simulation in a space rich with interaction possibilities for the agents. These strategies are macro-coordination and micro-coordination strategies. Macro strategies relate to overall action plans to accomplish the groups' goals, and is dependent on the domain (e.g., a group of soldiers may select a particular strategy influenced by training and doctrine, while a group of friends at a sporting activity may decide how to meet at a predetermined location). Micro coordination strategies relate to the dynamics of human communicative behavior, and is determined to a degree by the constraints of human perception, the physics of sound in voice communication, and cultural concerns. Micro coordination strategies are also influenced by the state of the immediate environment (e.g., voice communication may be suppressed in a movie theater, or visual displays and uptake may be constrained in a very dense crowd).

In our previous work [5], we introduced the CG model and showed how each agent's high-level choices on the route and walking strategies are affected by the group coordination. In this paper, we extend the CG model and operationalize it for multi-agent coordination. An agent evaluates the spatial and temporal conditions with respect to its group members and the state of immediate environment, and selects appropriate micro-coordination strategies. By showing that different micro-coordination strategies of individual groups has an influence on the overall distribution of a crowd, we demonstrate the importance of incorporating the coordination model into multi-agent simulations of large crowd behaviors.

The remainder of this paper is organized as follows. In section II, we discuss related works and give a short overview on Clark's common ground theory. Section III presents the design of our multi-agent model. The description of our coordination model for crowd simulation is provided in Section IV. Section V presents our implementation choices and the results of experiments. We draw conclusions and provide possible future research directions in Section VI.

II. Related Work

A. Crowd Dynamics Simulation Models

Numerous approaches of modeling crowd dynamics have been proposed in the literature. One common method to the simulation of mass flow is to model pedestrian in a particle-like manner [6], [7]. To the degree that humans are physical entities, and behave like particles, e.g., forming vortices as they move through an aperture [8], these flow-based models have been able to account for certain crowd characteristics. However, humans are more than particles, and physics-only flow models are incapable for accounting for human volitional behavior such as individual interest, social activity maintenance, and individual and group preferences.

A natural way to represent individual pedestrian is an agent based approach, in which each agent is modeled as an autonomous entity capable of perceiving its environment and interacting with other agents. Modeling of agent behavior has been heavily studied using variants of agent-based methods [9], [10] incorporated with social force models [11], [2], and velocity models [12], [13]. To achieve more realistic simulations, further sophistication to agent-based approaches often addresses cognitive [14], psychological [15], [16], emotional [17], and cultural factors [18], [19] in modeling of decision-making process of agents. However these models focus on the behavior selection of individual agents and not the effect of employing coordination model into a large scale simulation. In our work, we account how the ongoing communication and coordination among agents may affect the overall crowd flows in the simulation.

B. Common Ground Theory

People engage in a joint activity when they act in co-ordination with others to pursue a common goal. The CG concept as the basis of joint activity has seen application in AI agents [20], joint robot activity [21], and Computer-Supported Cooperative Work (CSCW) [22], [23], [24].

In Clark's model, CG is negotiated among participants in a joint activity, and has to be maintained for coordinated action to be successful. It considers the mutual knowledge, beliefs, and assumptions among individuals in a collaborative process. According to Clark, p is common ground for members of group $\mathcal{G}$ if and only if [3]:

1. members of $\mathcal{G}$ know that p;
2. members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that p;
3. members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that members of $\mathcal{G}$ know that p.

Suppose that A and B are heading to the seats in a football stadium. As they pass a concession stand, A decides to get some refreshments. A informs B of his plan to visit the stand and to return to their current location, x. We denote the plan to split up and reunite at x as P. For the plan to succeed, A needs to know that B knows the plan P. This, however, is insufficient for coordination. B needs to know that A knows that she is privy to P, otherwise she might think A would be left looking for her. Furthermore, if the agreement ends here, A may not know that B knows that he knows the plan, and may, therefore not be confident to execute the plan. Hence, A needs to know that B knows that he knows the plan.

This common ground may be communicated verbally, or may be enacted through action. In our example, A may signal his intention by pointing toward the concession stand and pointing to their current location x. This requires that B be within the range of sight and be looking at A. A needs to see that B is looking at A, and has signaled agreement (e.g., by nodding). B needs to see that A sees her nodding.

Finally A needs to see that B sees that he has seen and acknowledged the plan.

The CG model comprises logically separable components and thus lends itself to a direct implementation of coordination process. With Clark's CG model, we can simulate the coordination process at a higher behavioral level without modeling the process of social interaction at a lower level of communicative intent to reasoning. This allows the simulation to be more computationally tractable while maintaining realism.

III. MULTI-AGENT FRAMEWORK

We design our multi-agent framework as a layered architecture. In the lower layer, an individual agent acquires understanding of its immediate surroundings or situation that includes the state of member of the group to which it belongs, and the surrounding environment and selects proper behaviors for successful coordination among group members. In the upper layer, group level attributes for different social groups are maintained.

A. Agent Model

An agent with its personal identifier i is denoted as A_i. In order to function with group members, an agent should be able to interpret status and intentional signals of the members and adapt its behaviors. The agent also needs to understand its surrounding situation and respond to the environment conditions. Hence, our agent is modeled with sensory capabilities for speech, vision, and touch. Figure 1 shows the proposed sensory model for the agent with a perception geometry. In our model, touch can be sensed within range of agent radius r, hearing can be omnidirectional with range limitation d_h, and vision is directional and is effective up to a range, d_v, along its gaze direction (for simplicity, body orientation is synonymous to gaze direction in our simulation) and within a field of view defined by an angle, α. An agent can also sense the level of congestion in surrounding unit area. In a crowded area, d_h and d_v might decrease with respect to the level of congestion.
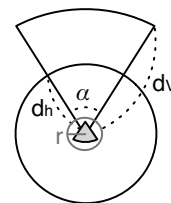

Figure 1.   Agent perception geometry

Our agent model is able to handle 'goal interrupts' through the introduction of new intermediate sub-goal determined either stochastically or through interaction with the environment. For instance, a members of groups may need to visit the restroom (stochastically generated sub-goal), or a member of a group may found a store that sells

something he is interested in (environmentally driven goal). To determine whether an agent is attracted by certain places during navigation, an agent A_i maintains a list of interests $I_{i,\alpha}$, $\alpha = 1, \ldots, K$ and corresponding propensity-to-visit values ranging from 0 to 1.0. When passing by some point of interest, an agent compares its propensity-to-visit value to the attraction intensity of the place, and selects potential sub-goals, s_h, accordingly. To prevent an agent proposing the same attraction point as a sub-goal repeatedly, we endow an agent with ability to keep the list of visited places in memory.

An agent is associated with a specific role $\mathcal{R}$ in the context of the current coordination mode. For example, an agent takes a role of initiator if it proposes a new sub-goal, and other members become respondents to the agent. These roles are used to determine which communicative and joint actions are taken by whom, and in what orders in micro- and macro-coordination strategies.

B. Group Structure

Our model assumes that the group memberships and goal of groups are known *a priori* and the relationships of agents are not subject to change throughout the simulation. Note that a group in our approach is formulated to provide social ties among agents, such as friends or family members. Group members tend to walk together and maintain the group cohesiveness. We model a group k as

$$\mathcal{G}_k = [k, A_i, p_k^0, g_k, Pr_k]$$

such that $\mathcal{G}_k$ has initial position p_k^0, final goal g_k and a roughly planned global path from p_k^0 to g_k. To add a variability in group characteristics, each group may have different preference probability Pr_k for macro-coordination strategies.

IV. COORDINATION MODEL

When a sub-goal is introduced by an agent, this necessitates the communication and update of CG with the rest of members to ensure the group cohesiveness. Our model for coordinated human group activity consists of three steps of agent execution: *situation assessment*, *micro-coordination*, and *macro-coordination*. Depending on the results of the coordination process, the group of agent can generate different sequences of actions. Figure 2 shows a conceptual agent architecture addressing the three execution steps in our model.

A. Situation Assessment

An agent constantly observes the surroundings and updates its awareness of the situation. The sensory information that an agent acquires includes:

- Event objects: interest points, final destination.
- States of group members: gaze direction, distance.
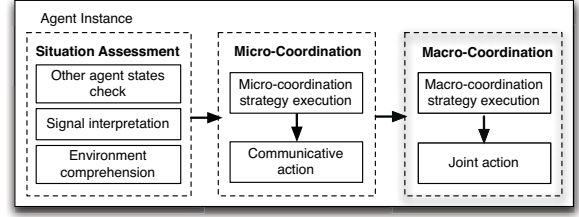- Signals from group members: communicative actions.

Figure 2. A conceptual architecture of our agent model

- Environment factors: congestion level.

In a normal navigation mode, an agent checks the presence of points of interest or proximity to the final goal point. When it finds an interest place, it proposes the place as a sub-goal to group members by choosing and initiating an appropriate micro-coordination strategy (microCS) according to its group's physical configuration.

To select a microCS, an initiator agent $A_i \in \mathcal{G}_k$ collects the following sensory inputs with respect to its group members that we model as the situational relationships (sit_{ij}) of A_i with all agents $A_j \in \mathcal{G}_k, \forall j \neq i$.

$$sit_{ij} : [d_{ij}, gd_j, V_{ji}|\alpha_i, E_{\mathcal{G}_k}]; \quad A_j \in \mathcal{G}_k, \forall j \neq i, \quad (1)$$

where d_{ij} is a distance between A_i and A_j, gd_j is a gaze direction of A_j, $V_{ji}|\alpha_i$ is the visibility of A_j to A_i subject to A_i's field of view α_i, and $E_{\mathcal{G}_k}$ describes the local environment of $\mathcal{G}_k$ as the level of congestion around the group $\mathcal{G}_k$. The effect of $E_{\mathcal{G}_k}$ is to determine the value of α_i and the visibility and audibility thresholds for d_{ij}. In a sparse crowd, longer-range gestures and speeches may be allowed while agents must be much closer together to communicate in a dense crowd. Furthermore, each agent may have its own evaluation function as a property of the agent (to determine its tolerance for crowdedness, for example). Then it selects an available microCS depending on the sit_{ij}; $\forall A_j \in \mathcal{G}_k, j \neq i$.

Figure 3 shows the six possible relations some agent A_j may have to A_i that will influence the micro-coordination choices made by A_i. In each case, we assume that A_i is already oriented toward A_j (i.e. V_{ji} is TRUE), otherwise no micro-coordination initiation is possible because the initiator would not know where the recipient is. A_i could already be in the visual field of A_j (conditions c_1 and c_2). A_i could be near to the visual field of A_j (conditions c_3 and c_4). We denote the near visual field condition as V'_{ij}. A_i could be outside V'_{ij} (conditions c_5 and c_6). For notation, we call this state $\overline{V}_{ij}$. Additionally we model the audibility range of A_j to A_i (meaning if A_i can gain the attention of A_j by using a vocal signal). Conditions c_2, c_4, and c_6 are in the audibility range, which we denote $\mathcal{H}_{ij}$, and c_1, c_3, and c_5 are in the $\overline{\mathcal{H}}_{ij}$ condition. In summary, the condition to state mapping are: $c_1 \leftarrow (V_{ij}, \overline{\mathcal{H}}_{ij})$, $c_2 \leftarrow (V_{ij}, \mathcal{H}_{ij})$, $c_3 \leftarrow (V'_{ij}, \overline{\mathcal{H}}_{ij})$, $c_4 \leftarrow (V'_{ij}, \mathcal{H}_{ij})$, $c_5 \leftarrow (\overline{V}_{ij}, \overline{\mathcal{H}}_{ij})$, and $c_6 \leftarrow (\overline{V}_{ij}, \mathcal{H}_{ij})$.

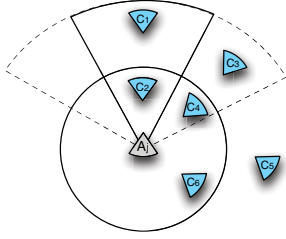Hence, Equation 1 can be evaluated for each A_j to assign it with one of our six configuration labels.

Figure 3. (a) Six possible relations of group members

B. Micro-coordination Strategy

A micro-coordination strategy (microCS) is a set of communicative actions that simulates how group members get each other to understand what they intend, so that some particular macro-coordination strategy (macroCS) may be initiated. Consider the situation where agent A_i receives a sub-goal event trigger (e.g., to visit the restroom). Depending on the evaluation of Equation 1, it will select an appropriate microCS, μ_γ to perform with the closest A_j. For example, if Equation 1 evaluates to $c_3 \leftarrow (V'_{ij}, \overline{\mathcal{H}}_{ij})$, it may select the microCS described in Table I. For our example, suppose the macroCS $\mathcal{M}_\beta$ in line 5 is a 'divide-and-wait' strategy, A_j will be tasked to communicate an 'inform-to-wait' activity with all other group members. This will require A_j to propagate the message to other group members by performing a microCS with each one to initiate the 'inform-and-wait' macroCS. The macroCS activity of A_i, in this case, is to proceed to its destination and return.

In our c_3 example, a second microCS may have been selected to move into hearing range (satisfying $\mathcal{H}_{ij}$) and simulating a speech interaction. Yet another c_3 microCS may be to have A_i walk within hearing range, and signal A_j to look at her, and then to proceed with a visual/gestural interaction. In our model each of these are coded separately as alternatives from which A_i may choose. This range of possible choices picked either randomly or based on some agent-specific preference may serve to make the simulation seem less mechanistic or fully deterministic, thereby adding to realism. The key is that the design of the microCS behaviors conform to the rules for CG negotiation outlined in Section II-B. Similarly, each configuration $c_1, \ldots, c_6$ can initiate a set of appropriate microCSs.

The failure or success of a microCS is dependent on two factors. The first is that the respondent agent A_j is somehow unable to comply with the signaling request of A_i (e.g., because it is simultaneously attending to another microCS from some other agent in the group), or if any step in the microCS script is unattainable (e.g., A_i's way is blocked to make it visible to A_j). The second is that the final random evaluation function of the microCS returns

Table I
AN EXECUTION OF A MICROCS, $\mu_\gamma. (V'_{ij}, \overline{\mathcal{H}}_{ij})$ CONDITION

step	Action Description
1	A_i moves to satisfy V_{ij}
2	A_i performs a *signaling action*, $\mathcal{S}$.
3	A_j gives attention to A_i.
4	A_i proposes a macroCS, $\mathcal{M}_\beta$ (i.e. select $\mathcal{M}_\beta$)
5	A_j signifies acknowledgement for $\mathcal{M}_\beta$.
6	If μ_γ is successful, Return TRUE (i.e. execute $\mathcal{M}_\beta$), Else Return FALSE (coordination failed)

a FALSE value. We do this to add some variability to the simulation. The probability of this random function returning FALSE is generally set to a very low number, meaning that most microCS negotiations result in a TRUE result. The associated macroCS script is activated only if the microCS is successful. The resulting action when a microCS fails is dependent on domain of simulation, and intended action. A_i could decide to abandon the sub-goal, or it may retry with a different microCS.

C. Macro-coordination Strategy

Given the ability to signal to other group member to adapt their actions, agents jointly select and perform a macro-level actions to achieve the group's goal (or sub-goal). A macro-coordination strategy (macroCS) refers to the macro-level action plan. MacroCSs need to be initiated whenever some sub-goal inducing event takes place that requires joint action.

The most common kind of sub-goal event is when a group member has to go to some point of interest and the entire group has to eventually proceed to its original goal. We call this the 'point-subgoal-and-resume' situation. The 'divide-and-wait' macroCS described earlier satisfies this kind of sub-goal. Another macroCS that can satisfy this sub-goal is the 'divide-and-proceed' strategy (see Section V-A2). Other sub-goal types may include an agent abandoning the group and leaving ('agent-leaves-group'), and the group abandoning their original goal and evacuating (e.g., when an emergency occurs – 'emergency-exit'). Depending on application domain, various other sub-goal types can be implemented. For example, in a military scenario, a sub-goal may be to send out a reconnaissance and wait, to divide and proceed to different goals, or to spread out and head to original goal position. The design of macroCSs may be based on common sense, typical domain-specific strategies, or from ethnographic observations of a population of people whose strategies are being modeled [18], [19].

When selecting a macroCS in certain situation, an initiator agent A_i of $\mathcal{G}_k$ may select a specific $\mathcal{M}_\beta$ from the set of macroCSs that satisfies the current sub-goal. This set may be constrained by the group's preference Pr_k. A simpler approach might be choosing a relevant macroCS to the sub-goal randomly.

V. Experiment and Results

In this section we first discuss the implementation of our CG-based crowd simulation, and present a series of experiments to show the quantitative differences in simulation results with and without incorporation of the CG model.

A. Implementation

1) MicroCS: A series of communicative actions to simulate the micro-coordination is defined for each situational relationship, c_1, ..., c_6, which are shown in Figure 3. Multiple options of microCS are available at each case, and the initiating agent must select one of them. For example, in c_1, a visual gestural coordination strategy may be activated in some cases, while in others, the agent may move to meet the audibility range constraint and communicate verbally.

Table II
A LIST OF MICROCSS IN OUR SIMULATION

Action block 1	
Sit.	Action Description
c_1	Immediate Comm. Visual Gesture
	Move to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_2	Immediate Comm. Visual Gesture
	Immediate Comm. Verbal
c_3	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
	A_i moves to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_4	Immediate Comm. Verbal
	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
c_5	A_i moves to satisfy V_{ij} & Comm. Visual Gesture
	A_i moves to satisfy $\mathcal{H}_{ij}$ & Comm. Verbal
c_6	A_j calls A_j, A_j turns around & Comm. Verbal
	A_i approaches and taps A_j, A_j turns around & Comm. Verbal
Action block 2	
A_i performs a *signaling action*, $\mathcal{S}$.	
A_j gives attention to A_i.	
A_i proposes a macroCS, $\mathcal{M}_\beta$ (i.e. select $\mathcal{M}_\beta$).	
A_j signifies acknowledgement for $\mathcal{M}_\beta$.	
If μ_γ is successful,	
Return TRUE (i.e. execute $\mathcal{M}_\beta$),	
Else Return FALSE (coordination failed).	

As previously shown in Table I, a microCS consists of several steps of action. Actions in step 1 are pertinent to a specific situational relationship of agents. Actions described in step 2 through step 6 in the table are general and commonly applicable to any microCS. Thus, we frame a microCS specification with two action blocks. Action block 1 contains a set of particular actions subject to a kind of microCS, and action block 2 is the generalizable part of microCS. Once agents complete the actions specified in action block 1, they perform the actions stated in the action block 2. A *signaling action* $\mathcal{S}$ implements the communication mode selected in action block 1. Table II summarizes a list of available microCSs in our simulation, in which A_i; $(A_i \in \mathcal{G}_k)$ is an initiator, and A_j; $(A_j \in \mathcal{G}_k, \forall j \neq i)$ is a respondent agent.

2) MacroCS: Since our implementation focuses on generating cohesive group movements in a crowd, we implement macroCSs only for the most common sub-goal type: 'point-subgoal-and-resume'. We support three macroCSs, 'divide-and-proceed', 'divide-and-wait', and 'detour-together' as listed in Table III. A_i is a member of $\mathcal{G}_k$ and initiator, and A_j; $(A_j \in \mathcal{G}_k, \forall j \neq i)$ is a respondent agent. Suppose the 'divide-and-wait' is selected. A_i heads for the sub-goal by itself while the rest of group members stay at the current location. Once it achieves the sub-goal, A_i returns to where it left the group members. When all of the involved group members reunite, they resume the original navigation. At the selection of 'detour-together', A_i leads all the members together to the sub-goal. When the sub-goal is accomplished, they resume the original navigation from the sub-goal location. In essence, during the detour, the group is in a 'follow-the-leader' macroCS with A_i in the role of the leader. Therefore a global path plan is updated first, and the group proceed to the final destination along the path. In our simulation, an agent choose a macroCS using the group preference characteristic, Pr_k.

Table III
A LIST OF MACROCSS IN OUR SIMULATION

	Action Description
divide -and-proceed	A_i proceeds to a sub-goal s_h, If the sub-goal is accomplished, proceed to a goal g_k
	A_j proceeds to a goal g_k
divide-and-wait	A_i proceeds to a sub-goal s_h If the sub-goal is accomplished, A_i returns to its previous location
	A_j stays and waits until A_i returns If A_i returns, $\forall A \in \mathcal{G}_k$ proceed to a goal g_k
detour-together	A_i leads $\mathcal{G}_k$ to the sub-goal s_h All $A_j \in \mathcal{G}_k$ follow A_i If the sub-goal is accomplished, $\forall A \in \mathcal{G}_k$ proceed to a goal g_k

3) Agent Model: For simplicity, we apply the perception model described in Figure 1 in the same way to all agents. The distance parameters d_h and d_v are subject to an environmental influence function modeled as a sigmoid function:

$$d'_\epsilon = d_\epsilon, \quad x \leq t,$$
$$d'_\epsilon = d_\epsilon(\frac{1}{1 + e^{\frac{(x-t)}{w_\epsilon}}}), \quad x > t, \tag{2}$$

where x is a number of agents in the unit cell, t is a tolerance for the crowdedness, w_ϵ determines the width of the sigmoid for each perceptual mode.

4) Environment Model: A virtual shopping mall is designed as shown in Figure 4 (top). The shopping mall model contains 8 restrooms (green squares) and 25 shops (yellow to red squares) as potential sub-goals. To make analysis results across conditions comparable in the current study,
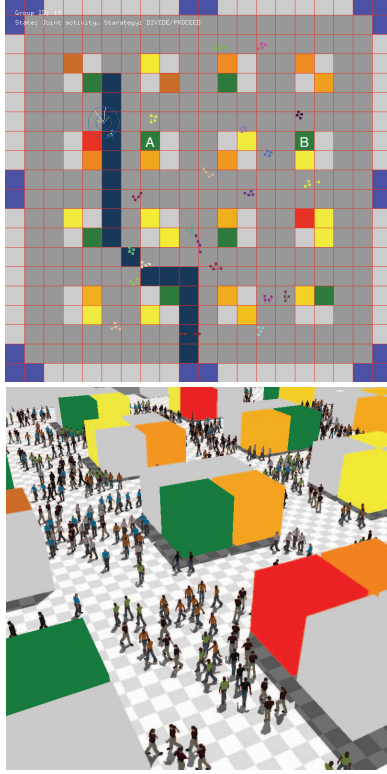
Figure 4. (Top) A snapshot of the shopping mall and (Bottom) a simulation running with animated characters

we assume all agents have homogeneous interests. The color intensity of shops represents the level of attraction. For example, shops in red are the points of highest attraction and those in yellow are the least attractive points. Eight exits (blue squares) are generated as possible final goals for agents. Each cell in the 2D grid describes a 5m×5m area.

Given the initial position, p_i^0, and final goal, g_k, of group k, the A* algorithm is used to generate a global path. The global path of $\mathcal{G}_{19}$ is drawn in navy blue in top of Figure 4. An animation of a simulation in a 3D environment is shown in bottom of Figure 4.

B. Experiments

We set four simulation configurations and compared the congestion levels at key points to show the impact of the 'coordination overhead' in the simulation. Given the total number of agents n, the agents are organized as (1) n of solely individual agents, (2) $n/4$ simple groups of 4 members without considering the CG model, (3) $n/4$ groups of 4 members which instantly choose a macroCS without processing micro-coordination, and (4) $n/4$ groups of 4 members with incorporating the CG model. We shall call each of the four configurations CI (for Individuals), CDT (for always 'Detour-Together'), CM (for Macro only), and

CμM (for both micro and Macro CSs), respectively. To simulate simple groups in CDT, we set the groups to always choose 'detour-together' (i.e., all group members will satisfy all sub-goals together, before proceeding to the final goal).

1) Scenario and Simulation Parameters: For all the individual agents in CI and agents of groups in CDT, CM and CμM, one of the eight exits is selected as a final goal g_k at random. Initial positions of agents are randomly distributed in a shopping mall, taking care that members within groups are collocated. Though agents in CI are all individuals, we initialized the simulation with agents clustered in groups of 4 to have the same initial conditions as CDT, CM, and CμM.

Starting from initial positions, agents walk around the shopping mall and eventually proceed to the exit. A random event generator triggers agents to visit the nearest restroom. When agents pass by shops, they may be attracted to some shop within a range (e.g., 15m×15m in our implementation).

We used 1000, 2000, and 3000 for the total number of agents n, and hence they are organized as 250, 500, and 750 groups, respectively. Default perceptual geometry of agent is set to: $[d_v, d_h, \alpha] = [15m, 5m, \angle 60]$. For the environmental influence introduced in Equation 2, $t = 12$ and $w = 35$ are used (for simplicity, we use the same w for d_h and d_v). w is determined assuming that the maximum agent capacity of a unit cell of 5m×5m being around 35.

C. Results and Discussion

Our hypothesis for the experiment was that the pragmatic need to maintain common ground incurs costs at the level of the entire simulation. For instance, if we find a non-CG simulation to always performs better than a CG-based model with regard to level of congestion, this might suggest that models that do not add the cost of coordination may systematically under-estimate real crowd effects.

In order to compare the congestion levels among the four simulation configurations, we counted the number of agents in 3×3 unit cells (hence 15m×15m) centered at key points such as shops or restrooms per frame. We present the results measured in area A and B (marked in Figure 4(a)) in Figure 5. All the results are averaged over 10 independent simulations. Stochastic events invoking agents to visit a restroom are indicated with gray vertical lines. Figures 5(a)-(c) are the level of congestion measured in area A for 250, 500, and 750 groups, respectively. Figures 5(d)-(f) are the results for area B. The x-axis in each graph represents time and the y-axis measures the level of congestion in # of agents per unit area.

The results presented in the graphs indicates the emergent effects of individual agent and small agent group behavior on the overall crowd simulation. From the graphs, it can be seen that the CI case almost always results in the lowest congestion because there is no coordination or joint action overhead in the individual action plans. CDT consistently results in the highest congestion across crowd size simply
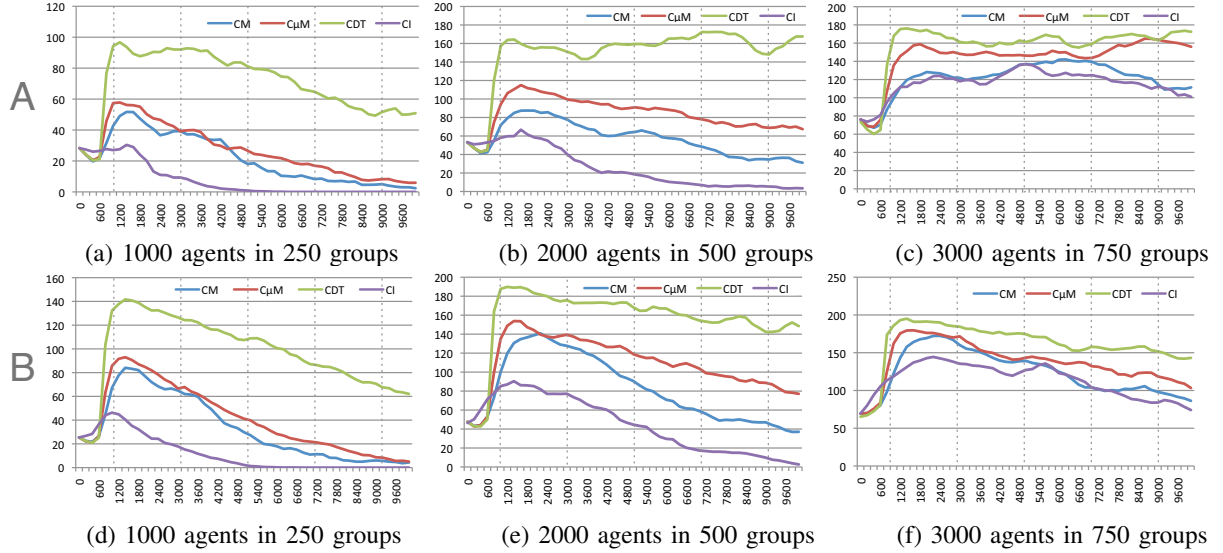
Figure 5. Congestion levels measured as the number of agents in unit areas of CI, CDT, CM, and CμM

because the total distances traversed by all agents is the highest in this condition. The interesting differences come in the CM and CμM conditions as these vary depending on the overall crowd size. Notice that across all crowd sizes, congestion tends to fall over time as some agents and agent groups leave the mall, leading to overall lower congestion everywhere. Since the sub-goal events are generated stochastically, not all groups in the CDT, CμM, and CM cases will have members that receive sub-goal events at any particular sub-goal event cycle.

When we compare the conditions with respect to overall crowd size, we notice that apart from CDT and CI, CM and CμM are quite similar when the crowd is sparse (1000 agents in 250 groups), with CμM resulting in slightly higher congestion followed by CM and then CI. This shows that the joint action and the micro-coordination overheads have some effect at this crowd level.

This order of difference becomes accentuated in the middle level of crowd size (2000 agents in 500 groups). The general difference between locations A and B in the simulation is that A is near the middle of the environment where the effects across conditions are accentuated, and B is near the exit where congestion tends to fall more quickly as agents exit the environment over time. In both areas A and B, the micro-coordination overhead causes higher congestion than the macroCS-only (CM) condition. This shows that the behavioral cost of performing the necessary micro-coordination activity cannot be ignored in overall crowd simulation with multiple agents.

When the overall crowd density is very large (3000 agents in 750 groups), the congestion across the CDT, CμM, CM, and CI conditions become relatively similar because we have arrived at a ceiling effect of overall over-congestion.

One way to think of this is that when the crowd density gets too high, no strategy will work effectively. However, even in this condition the order of congestion levels across groups (CDT, CμM, CM, CI) is generally maintained, and the micro-coordination process adds higher congestion than the macroCS-only condition. In area A, especially the congestion of CμM even approaches that of CDT (which is an unrealistic condition in real life), while the CM congestion level falls to near CI levels after 8800 time steps as agent groups exit the environment.

The difference of the congestion level among the different crowd models in our experiment demonstrates the significance of incorporating realistic human social interactions into an agent design to achieve valid simulation results. For example, suppose that we were designing a mall with certain target number and placement of shops for a target number of visitors. Underestimations of crowd complexity may cause designers to be overly optimistic about the size of corridors or of evacuation rates in an emergency.

VI. CONCLUSION

We presented a computational model informed by common ground theory to incorporate the impact of social interaction in multi-agent simulations. In our approach the CG model is framed as micro- and macro- coordination strategies, which allows the multi-agent system to adapt to varying domains. Our approach focuses on simulating the coordination process established by CG theory rather than modeling the process of social interaction at a lower level of communicative intent to reasoning. We believe that this intermediate level simulation is viable for two reasons. First, it is computationally more tractable than trying to simulate the micro processes of communication

and negotiation of joint action itself. Second, CG theory is a well-vetted scientific theory, and already takes the effect of real human micro-level negotiation of joint activity into consideration. In fact, because the theory is well established, any bottom-up micro-negotiation model will be hard pressed to match the realism of a CG model.

We conducted a series of studies and showed notable differences resulted in the simulations when the common ground model is enabled. Since CG is a well-researched interaction model, such differences provide insight concerning the effect of employing such a model into a large scale simulation may be important.

Our future research direction includes evaluating the realism of crowd behaviors and dynamics by conducting user studies. The congestion level in the simulation reflects only a quantitative aspect of a simulation, and we want to validate our CG model through a qualitative measurement. In order to do that, a complete integration of the simulation engine with animated characters supporting subtle communicative gestures are also needed.

REFERENCES

[1] C. Peters and C. Ennis, "Modeling groups of plausible virtual pedestrians," *IEEE Comput. Graph. Appl.*, vol. 29, no. 4, pp. 54–63, Jul. 2009.

[2] M. Moussaïd, N. Perozo, et al "The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics," *PLoS ONE*, vol. 5, no. 4, pp. e10047+, Apr. 2010.

[3] H. H. Clark, *Using Language*. Cambridge U. Press, 1996.

[4] H. H. Clark and S. Brennan, "Grounding in communication," in Perspectives on socially shared cognition, American Psychological Association, 1991, pp. 127–149.

[5] S. I. Park, F. Quek, and Y. Cao, "Modeling social groups in crowds using common ground theory," To appear in Proc. of the 2012 *Winter Simulation Conference (WSC)*.

[6] R. L. Hughes, "A continuum theory for the flow of pedestrians," *Transportation Res. Part B: Methodological*, vol. 36, no. 6, pp. 507 – 535, 2002.

[7] R. Narain, A. Golas, S. Curtis, and M. C. Lin, "Aggregate dynamics for dense crowd simulation," *ACM Trans. Graph.*, vol. 28, pp. 122:1–122:8, 2009.

[8] A. Treuille, S. Cooper, and Z. Popović, "Continuum crowds," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 1160–1168, 2006.

[9] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *SIGGRAPH Comput. Graph.*, vol. 21, pp. 25–34, 1987.

[10] S. R. Musse and D. Thalmann, "A Model of Human Crowd Behavior: Group Inter-Relationship and Collision Detection Analysis," in *Wksp. Computer Animation and Simulation of Eurographics*, 1997, pp. 39–52.

[11] D. Helbing, P. Moln, I. J. Farkas, and K. Bolay, "Self-organizing pedestrian movement," *Environment and Planning B: Planning and Design*, vol. 28, no. 3, pp. 361–383, 2001.

[12] P. Fiorini and Z. Shillert, "Motion planning in dynamic environments using velocity obstacles," *Intl J. of Robotics Res.*, vol. 17, pp. 760–772, 1998.

[13] J. van den Berg, M. Lin, et al "Reciprocal velocity obstacles for real-time multi-agent navigation," in *IEEE Intl. Conf. Robotics and Automation, 2008. ICRA 2008*, pp. 1928 –1935.

[14] A. S. Rao and M. P. Georgeff, "BDI-agents: from theory to practice," in *Proc. 1st Intl. Conf. on Multiagent Systems*, 1995.

[15] N. Pelechano, "Crowd simulation incorporating agent psychological models, roles and communication," in *First International Workshop on Crowd Simulation*, 2005, pp. 21–30.

[16] M. Hoogendoorn and J. Soumokil, "Evaluation of virtual agents utilizing theory of mind in a real time action game," in *Proc. 9th Intl Conf. Autonomous Agents and Multiagent Systems: vol. 1*, AAMAS '10, 2010, pp. 59–66.

[17] L. Luo, S. Zhou, el al "Modeling human-like decision making for virtual agents in time-critical situations," in *Proc. 2010 Intl. Conf. Cyberworlds*, CW '10, 2010, pp. 360–367.

[18] J. Tsai, N. Fridman, et al "Escapes: evacuation simulation with children, authorities, parents, emotions, and social comparison," in *10th Intl Conf. Autonomous Agents and Multiagent Systems - Vol. 2*, AAMAS '11, 2011, pp. 457–464.

[19] B. Endrass, E. André, et al "Culture-related differences in aspects of behavior for virtual characters across germany and japan,"in *10th Intl Conf. Autonomous Agents and Multiagent Systems - Vol. 2*, AAMAS '11, 2011, pp. 441–448.

[20] G. Klein, P. J. Feltovich, J. M. Bradshaw, and D. D. Woods, *Common Ground and Coordination in Joint Activity*. John Wiley and Sons, Inc., 2005, pp. 139–184.

[21] R. Kirby, R. Simmons, and J. Forlizzi, "Companion: A constraint-optimizing method for person-acceptable navigation," in *18th IEEE Intl Symp. Robot and Human Interactive Communication, RO-MAN 2009*, pp. 607 –612.

[22] A. Monk, "Common Ground in Electronically Mediated Communication: Clark's Theory of Language Use", in HCI Models, Theories and Frameworks, Morgan Kaufmann, Morgan Kaufmann, 2003, pp. 265–290.

[23] D. C. Neale, J. M. Carroll, and M. B. Rosson, "Evaluating computer-supported cooperative work: models and frameworks," in *ACM CSCW '04*. 2004 pp. 112–121.

[24] G. Convertino, H. M. Mentis, M. B. Rosson, J. M. Carroll, A. Slavkovic, and C. H. Ganoe, "Articulating common ground in cooperative work: content and process," in *Proc. 26th CHI*, CHI '08, 2008, pp. 1637–1646.