Using Qualitative Reasoning for Social Simulation of Crowds: A Preliminary Report

Gal A. Kaminka and Natalie Fridman*
The MAVERICK Group
Computer Science Department and Gonda Brain Research Center
Bar Ilan University, Israel

Abstract

We report on the use of qualitative reasoning (QR) for modeling the social behavior of large groups, in particular in demonstrations. We develop qualitative models consistent with the partial, qualitative social science literature, allowing us to model the interactions between different factors that influence violence in demonstrations. We then utilize qualitative simulation to predict the potential eruption of violence, at various levels, based on a description of the demographics, environmental settings, and police responses. In addition to providing predictions, the resulting qualitative simulation graph is analyzed to determine the factors that are most important in influencing the outcome. These factors can be used to support decision-makers. We make three separate contributions: first, we briefly show how the use of QR can be used to to contrast the predictions of different social science theories; second, we demonstrate that the QR technique has better explanatory power than a machine learning approach to prediction; and third, we use the analysis algorithm to determine important factors in specific real-world demonstrations. We show that the algorithm identifies factors that correspond to experts analysis of these events.

1 Introduction

We report on the use of qualitative reasoning (QR) for modeling the social behavior of large groups, in particular in demonstrations. Existing knowledge about how demonstrations develop dynamically, and the different factors that affect them, is unfortunately only partial, and always qualitative in nature. Different theories exist, typically analyzing specific real-world events to draw a conclusion as to the qualitative relation between a handful of factors. QR techniques can be used to tie together these *micro-theories* into a single qualitative model, which can be tested and reasoned about. Each theoretical component becomes a qualitative *quantity* (in QR terms). The qualitative relations between these components are modeled appropriately as *influences*.

We develop qualitative models consistent with the partial, qualitative social science literature, allowing us to model the interactions between different factors that influence violence in demonstrations. We focus on three separate models, incrementally increasing in complexity, and in the number of factors they consider. These three models are evaluated on real-world scenarios, using news reports and wikipedia entries as the source of information as to the values of different quantities.

By a simple technique, which considers the number of paths leading to different violence outcomes, we are able to provide an estimate of the likelihood of different outcomes, for each test case. We contrast the predictions of the different models, and thus demonstrate one important benefit of using QR for social simulation modeling, i.e., the ability to easily test social science theories on real-world data. Moreover, we show that while a machine-learning technique can be used to generate slightly more accurate predictions, it lacks the ability to support hypothetical "what-if" reasoning, because it does not have the explanatory power of a social science model.

Finally, we develop an algorithm which analyzes the qualitative simulation graph of each test-case, to determine the factors that are most important in influencing the outcomes of *the specific case* under consideration. The key to this algorithm is to determine simulation graph nodes with high outcome entropy, i.e., nodes which lead to different outcomes, at fairly uniform likelihood. In the states corresponding to such nodes, it is possible to identify actionable factors that can be used to influence the outcomes. We show that for real-world cases, the algorithm results in identifying causes also identified by experts.

2 Background and Related Work

Usage of computer simulation is considered to be a promising approach for modeling and reasoning regarding different social phenomena [6]. There are several micro and macro level techniques that enable such modeling, e.g., usage of agent based simulation, cellular automata and system dynamics. However, there are also techniques that do not require building a model to enable predictions, such as machine learning techniques, in particular a decision tree.

*We thank IMOD and ISF Grant 1357/07 for partial support of this research. Thanks to Avishay Zilka, and K. Ushi.

Agent-based simulation is a micro-level approach where by social behaviors are simulated by simulating each individual, and their interactions. By applying agents as an "intelligent" entity we have the ability to model complicated social interactions. Such simulations have been successfully used in modeling crowd behaviors [5; 7], economic phenomena [16], and more. However, it is a bottom-up approach in the sense that to receive a macro-level behavior we must model the micro-level interactions which necessitates detailed individual modeling, and when number of agents is scaled up it may provide significant computational barriers. Furthermore, there are domains such as predicting the likelihood of violence that modeling at the individual participant level may be too high a resolution and even unnecessary.

System dynamics approach [6] is a macro level approach in the sense that it models an entire system. It uses defined stocks, flows and feedback loops to model system behavior. The models are basically sets of differential equations that describe changes in the system. In our domain, such accurate and full definitions are not available.

Qualitative Reasoning (QR) is another macro level approach, allowing modeling and reasoning with partial and imprecise information. It has been used to allow for commonsense reasoning in physics [10; 3]. However, it has also been applied to other branches of science: ecology [13], social science [8], politics [4] etc. However, our use of QR to model and predict the violence level during demonstrations is novel.

Fuzzy Cognitive Maps (FCM) [9] is also a macro level approach which enables causal reasoning using fuzzy directed graphs. Similarly to QR, FCM enables imprecise and qualitative representation of the model. However, the output of FCM is a recommendation on a single action or goal, while QR returns the set of all possible behaviors that the model may manifest.

Machine learning techniques such as decision tree [12] enables reasoning regarding social phenomena without providing a model. Decision tree takes as an input set of properties and build a model, which is set of rules, that allows classification of the observed data according to the given properties. It is mostly used in domains that there no significance for the model and only the classification counts. However, as we show in this paper, prediction (classifications) accuracy is not the only requirement for policy decision-support.

3 Qualitative Simulation of Demonstrations

Qualitative simulation enables reasoning about possible system behaviors that can emerge from an initial world state. The simulation takes as input the initial state of the world which contains a structural description of the model and produces a simulation state transition graph, which captures all possible qualitative states that may manifest from the initial state. We refer to a sequence of states connected by state transitions where each state is the immediate successor of the one before, as a behavior path.

In each cycle and on each quantity, all influences (direct and indirect) are combined. When positive and negative influences are combined ambiguities may occur. The qualitative simulation considers all the possible combinations thus, when qualitative description is incomplete, it provides a non deterministic prediction.

QR modeling techniques are a good match for the current state of knowledge in social sciences regarding demonstrations. Existing theories are inaccurate nor complete. There are many micro-theories regarding the influencing factors on the violence level: Each such theory focuses on a small subset of factors. Integrating all of them into a single unified model is real challenge.

The Israeli police initiated a comprehensive study to address this challenge, resulting in a report [2] that provides a collection of factors and their influence on the violence level and also on each other. Their goal was to classify and analyze different kinds of demonstrations in order to propose appropriate methods for the police force in dealing with the mass. They studied 102 crowd events (in particular demonstrations) during the years 2000–2003 and interviews with 87 policemen and police officers. They analyzed a variety of factors that may affect violent behavior, as well as relevant literature. This report is a qualitative collection of factors which provide a challenge to the reasoning process. We use this report as a source of knowledge based on which we developed our models and by using qualitative simulation we provide an ability for reasoning regarding potential violence level.

Indeed, we developed three separate models, incrementally increasing in complexity and size, of the different components influencing violence in demonstrations. These are described below.

Base Model. The first (*Base*) model was developed based on the report's literature review [2] (see Figure 1). It was proposed there as a first attempt at building a baseline, purely based on literature review. According to the Base model the most influential factors on the violence level during the demonstration are (1) the crowd's a-priori hostility towards the police; (2) willingness to pay the personal price (such as willingness to be arrested); (3) low chance for punishment for violent actions (e.g., a belief that police will not respond strongly); (4) group cohesiveness; (5) previous history of violence. All of these directly increases the level of violence. However, not all have an opposite effect when reversed. For instance, the existence of previous history of violence among the specific group of demonstrators increases the potential violence level, but lack of such history does not decrease the violence level (i.e., has no effect).

Police Model. The *Police model*, described by Karmeli and Ravid-Yamin [2] (Figure 2), significantly expanded the Base model, based on interviews with police officers and their investigation into 102 demonstrations. This model adds 12 more variables, roughly divided into several groups. *Environmental factors* include weather, time of day, location sensitivity (e.g., for religious reasons), and time of year sensitivity (e.g., Christmas). *Participant factors* include the number of participants, the existence of violent core among the participants, the existence of group leader, and the cohesiveness of the group (e.g., if they all come from a particular ethnic minority). *Procedural factors* include a pre-demonstration

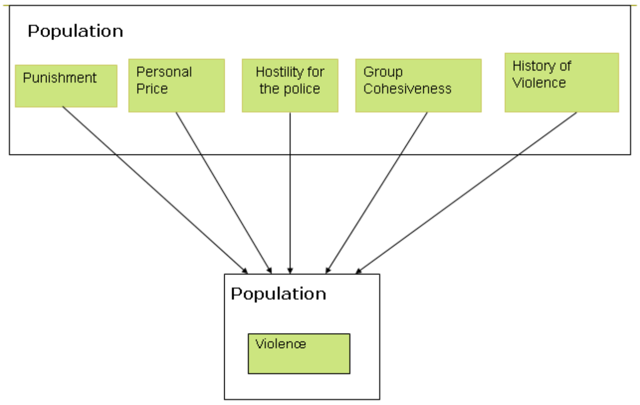


Figure 1: **Basic Model: Structure**

request for demonstration license, the purpose of the event (emotional or rational), and the timing and strength of police intervention.

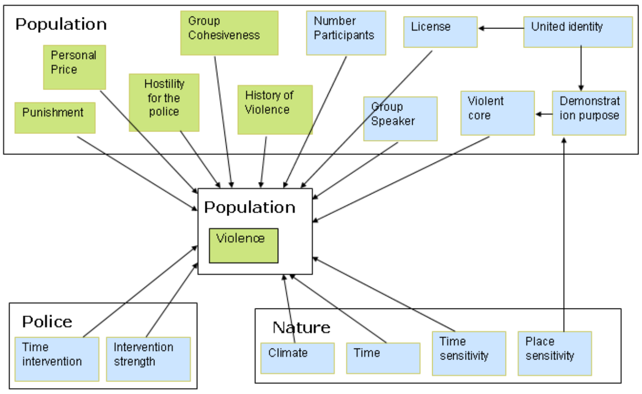


Figure 2: **Police Model: Structure.**

BIU Model. The third model, BIU model (Figure 3), is our own novel extension of the Police model. Based on interviews with social and cognitive scientists, as well as additional literature surveys, we added four additional variables, and updated 19 influences (relations) among the variables. The added factors are: (1) event order (indicates amount of preparation that was made following the event, such as delineation, disposition of the police forces etc.) (2) participants anonymity (indicates whether the participants can be recognized), (3) participants' visual cohesiveness (such as similar outfit as among football fans) and (4) light.

We provide here several examples for updated influences. First, we updated the influence of police's intervention strength, thus instead of direct impact on violence level as in the Police model, it impacts the participants' belief that they may be punished, and their hostility for the police. In BIU model, high intervention strength increases participants' hostility for the police and increases the participants' chance for punishment. However, low intervention strength just decreases the participants' chance for punishment without a change in hostility for the police factor. Another example is that existence of group speaker and the request (and acceptance) of a demonstration license increase the maintenance of order, which decreases the violence level. In contrast, in the Police model, license and group speaker variables had a direct influence on the violence level. Moreover, for the variable *number participants*, we no longer allow direct influence on the violence level as in Police model, but instead have it influence the participants' anonymity level ("the more participants around me the less recognizable I am"). Another example of addition to the BIU model is: participants visual cohesiveness has an impact on group cohesiveness, it actually increases the sense of belonging to the same group.

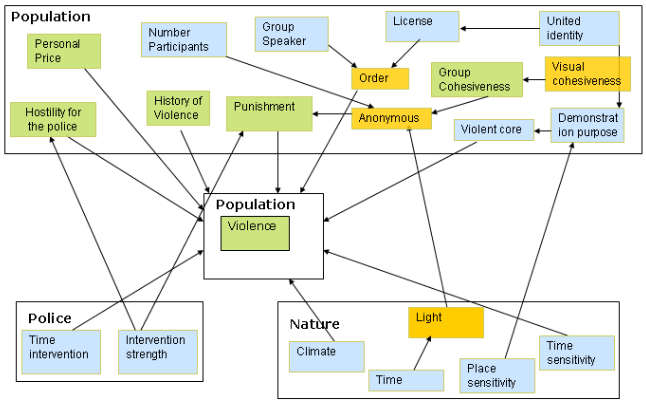


Figure 3: **BIU Model: Structure**

4 Prediction and Analysis

For different demonstration cases, one can set the initial state quantities to their qualitative values, based on the demographics and environment values as known at the time. Then qualitative simulation is used to expand all possible outcomes possible based on the initial values. The resulting violence outcomes are used as the basis for prediction. Then, the simulation graph itself is used to point out specific settings in which intervening is particularly important.

Estimating the Likelihoods of Different Outcomes. A qualitative simulator takes as input an initial setting of the world state (partial state information is acceptable) and produces a simulation state-transition graph. Each sequence of states, following transitions from the initial state and ending with a different outcome state is a possible system trajectory—a possible sequence of qualitative state changes that may occur given the initial state, and the qualitative dynamics specified. The end state in of each such path is where the system dynamics allow no further evolution (i.e., the system is stable). Taking the value of the outcome variables (in our case, violence level) in these final states allow categorical predictions.

The violence level variable can take three categorical values: zero, low and high. The zero value represents demonstrations that ended without any causalities and also without property damage. The low value represents demonstrations that ended with property damage but without any causalities, and the high value represents all those demonstrations that ended with causalities.

However, it is not enough to know whether a demonstration might be violent; in a sufficiently complex model, all three possible outcomes will have at least one stable state in which they appear. Instead, our goal is also estimate the likelihood of different outcomes. To do this, we use the received state-graph as an input and based on this developed graph we calculate the likelihoods of different outcomes by counting the number of behavior paths that lead to a specific violence level, and dividing it by the total number of paths. The result is a distribution over possible violence outcomes.

For instance, suppose that there are 345 total paths leading from the initial simulation state, to stable states (leaves in the simulation graph). Suppose further that 123 of these paths end up in leaves with violence level *high*, 121 of the paths end up in leaves with violence level *low*, and the remaining 101 paths end up with violence level *zero*. Then the distribution of predicted violence is $\langle high, low, zero\rangle = \langle 123/345, 121/345, 101/345\rangle = \langle 0.36, 0.35, 0.29\rangle$.

Determining Important Factors in Specific Settings. The a-priori predictions of the model, given initial values, do not provide decision-makers with information about factors that, *in the particular case*, influence the level of violence. Thus we do not know, out of the many different factors that may increase the level of violence, which are important in the specific case being simulated.

For instance, a perception of anonymity by the demonstrators may reduce their fear of being punished for breaking the law, and this in turn can increase the chances of violence erupting during a demonstration. Perception of anonymity can be addressed by the police in various means: segregating the demonstrators into smaller disconnected groups, shining bright lights (if the demonstration is held when dark), etc. However, a-priori, there are few indicators of the potential anonymity perceived by the crowd.Moreover, we do not know whether tackling such perception can be effective: It could be that there are so many factors increasing the violence, that anonymity (being an indirect influencing factor) is just not worth treating. Or likewise, it could be that violence is highly unlikely, and thus bringing in bright lights is just an overkill that may incite the crowd. Thus anonymity should be addressed only in specific settings, where it becomes a determining factor in promoting violence.

To aid in this decision-making, we describe an algorithm for determining the k most important factors in influencing the outcomes of the simulation, and also determining the conditions under which they should be addressed. This is carried out as follows.

First, we traverse the simulation graph bottom up (from leaves to the root, which is the initial state). In each node, we count the number of paths resulting from it, which end up in high violence, low violence, or zero violence. This process is in fact a generalization of the prediction process described above for making predictions. The number of paths of each type which is associated with the initial state is exactly the outcome distribution which we describe above. Here, we are simply generating the same count for all nodes in the graph.

Now, we identify the k nodes with the highest level of *outcome entropy*, who have more than a single child[1]. The outcome entropy measures the uniformity of the distribution of different potential violence outcomes. A perfectly-uniform distribution $\langle 0.33, 0.33, 0.33\rangle$ will have maximal entropy; a perfectly non-uniform distribution where all paths lead to the same outcome will have minimal entropy (0).

The reason for seeking simulation nodes with high entropy is that these our nodes where there is a difference to be made, i.e., they are *actionable*. Nodes with low entropy are those in which outcome is essentially decided already. Changing their outcome will necessarily involve making multiple changes to the state, i.e., they involve more complex intervention. In contrast, nodes with high entropy are nodes whose outcome is far from decided, and thus offer a good opportunity for relatively simple intervention.

Given the k highest-entropy nodes, we can now identify the factors that influence the outcomes. We do this by examining the simulation information saved at the node, and contrasting it with that of its children. We thus determine which qualitative relations are at work at the node, and how they interact to lead towards the different outcome. This significantly narrows the list of factors that are relevant to the different outcomes, and also unravels the conditions in which these factors are important.

For instance, we may see that a node splits into different children because of the interaction between two opposing forces: Low anonymity which decreases violence (it increases the chances of punishment, as perceived by the crowd), and the lack of police responses to events (i.e., the police is responding too late, or too weakly) which increases violence. Both these factors interact to cause multiple possible outcomes. Acting (e.g., by increasing police force) can countermand the interaction, and cause the outcomes leading from this node to converge towards low or zero violence. Moreover, the state represented by the simulation node tells us the conditions under which increasing the police force will be affective (as this is not always the correct response to violence!).

5 Evaluation

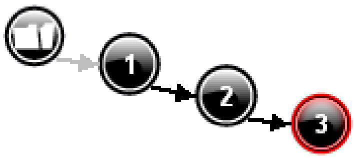
To evaluate the approach described above we implemented the three models in GARP, a QR engine which enables building and simulating qualitative models and was successfully used in many domains [13; 1]. We also developed 24 test-cases, real-life demonstrations reported on by a variety of sources. 22 of these were taken from Hebrew Wikipedia under category demonstrations. The cases were taken both from the main category, and from the subcategories: "demonstrations in Israel" and "massacres in demonstrators". We disqualified general descriptions which did not describe a specific event (e.g., descriptions of recurring demonstrations) and also omitted two cases due to lack of information (for a total of twenty cases). Additional three cases are well known events which where extensively analyzed and described [2; 11; 15; 14] by experts. The last event was a peaceful demonstration that we video-taped.

[1] A parent with a single child will have the same count of paths going through it as its child, and thus the same entropy. But we seek the state where the divergence into multiple outcomes occurs, hence we prefer the child.

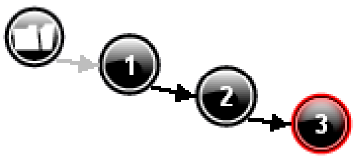
To initialize the test cases, we utilized the information appearing in their descriptions in the literature and in Wikipedia. We initialized only the quantities for which we had explicit information. Quantities for which we had no information or ambiguous information were removed from the initial set. Qualitative simulation can work with such partial information.

5.1 Prediction accuracy

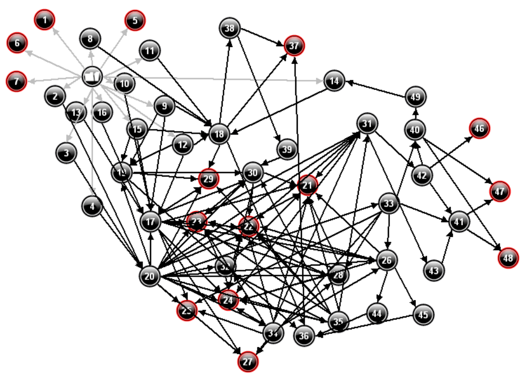
Each model was examined on the twenty four test cases described above. We use the simulation state graph for our calculation of the numeric probability as presented earlier. Figure 4 represents the example of transitions state-graph built by GARP of one of the experiment. Figure 4(a) represents the Base model built state-graph, Figure 4(b) represents the Police model state-graph in same experiment and Figure 4(c) represents the BIU model state-graph in the same experiment. The circles represents states and the arrows represent state transitions. The end path circles are the final states with one of the possible outcomes: zero violence, low violence or high violence.



(a)



(b)



(c)

Figure 4: Transitions state-graph

In evaluating the predictions of the different models, we look at the maximum likelihood prediction of each model, for the 24 different cases. If the maximum-likelihood prediction corresponds to the outcome of the event in the real world, we count this as an error of 0. Otherwise, we examine how far off was the prediction from the actual outcome: Predicting low violence is closer to zero violence than a prediction of high violence. Thus a 1-level error corresponds to a mistake by one ordinal level. The maximal error is a 2-level error (e.g., predicting high violence when the actual result is zero violence).

Figure 5 summarizes the experiment results across the 24 cases. The horizontal axis separates the different models. The vertical axis measures the number of cases. The results of the three models are presented as stacked bars. Their total height is always equal (24 cases), but they are internally divided into 0-level errors, 1-level errors, and 2-level errors. The *base* and *police* models each have 19 successes, and 5 2-level errors. In contrast, the *Bar Ilan* model replaces 4 of these 2-level errors (predicting high levels of violence where there was none) with 1-level errors (predicting low levels of violence where there was none). Its predictions are thus noticeably closer to the actual outcomes.

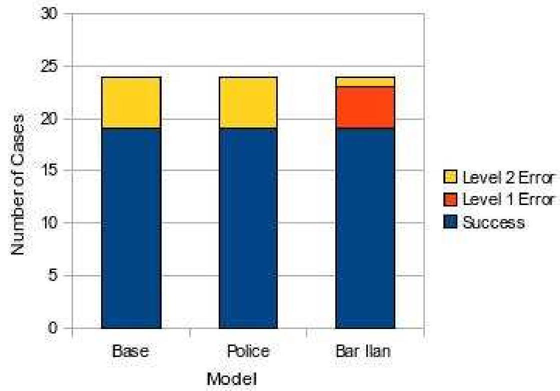
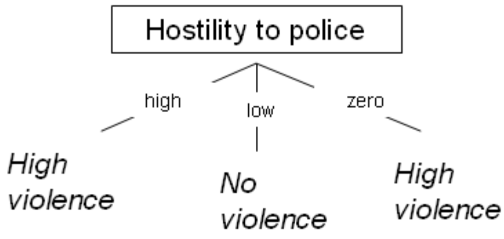


Figure 5: Model prediction results.

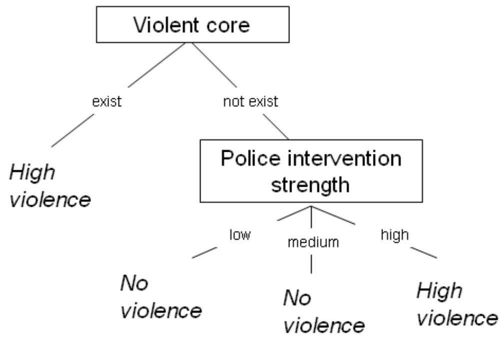
5.2 Comparison to the Machine Learning Techniques

We wanted to examine whether the machine learning techniques such as decision tree may provide a better prediction than our models. We used Weka, an open source software that contains collection of machine learning algorithms and used the J48, decision tree algorithm. We built three files that were used as an input to Weka. Each file contains collection of attributes with their values and was built based on the quantities initialization set of each QR model (Base model, Police model and BIU model). The target class value of attribute violence in each file was set based on the real-life event outcome. The output of J48 algorithm is the learned decision tree and classification statistics.

Figure 6 present the decision trees that were learned based on the each QR model initial quantity set. Figure 6(a) presents the tree that was learned based on the quantity set of the Base model (Base tree), Figure 6(b) presents the tree that was learned based on the quantity set of the Police model (Police tree) and the same tree was learned based on the quantity set of the BIU model (BIU tree).



(a)



(b)

Figure 6: Decision trees

The results show that Police tree and BIU tree have 100% of success in classification, while Base tree has 70.8% of success. While the machine learning techniques provide an accurate prediction, a slightly better prediction than the BIU model with QR approach, we will claim, in the next section, that QR approach is much more sensitive to changes and can account for *what if* scenarios. Thus, using QR approach is better for reasoning regarding the potential violence level to improve the police decision making process.

5.3 Sensitivity Analysis

In the following experiments we want to demonstrate the use of QR approach and machine learning techniques for variety of hypothetical changes. According to experts [11; 15; 14] in several of the events we modeled (Exp. 15–17), the police intervention strength was found to be one of the important factors for the violence eruption. Thus, in this section, we want to examine the presented QR model's prediction and the machine learning techniques in *what if* scenarios.

First we want to examine whether the presented models with QR approach and machine learning techniques are sensitive enough for the changes in term they can account for *what if* scenarios. Moreover, we want to examine what influence has the police intervention strength on the event outcome, could it be the main factor than can prevent the violence or the event essence to be violent and the police intervention strength has little to do with it? Then we want to examine hypothetical situation of changing the chance for violence in several test cases scenarios by changing different controlled factors and not just the police intervention strength.

Sensitivity Analysis: Experiment 1

In this experiment we want to examine whether the presented models built with QR approach and the machine learning technique, may account for changes in the police intervention strength. We used the same twenty four test cases as described in Section 5 and examined the police intervention strength attribute with it's all possible values. As in Section 5 we estimated the likelihood of different event outcomes. The model will consider to be sensitive to the changes if for different values in examined attribute, it will provide different outcome. The change can be one of the following: different distribution values but no change in classification and different distribution values with change in classification.

We compared the BIU and Police models built with QR techniques to decision tree that was built with BIU initialization set. The Base model built with QR techniques is irrelevant for this experiment since the Base model not accounts for the factor of police intervention strength therefore there are no change in the model's predictions.

The results show that Police model changes its distribution in five test cases (from twenty four) and in two of them it also changes its classification. The BIU model changes its distribution in all of the examined test cases and in seven of them it also changes its classification. The decision tree cannot provide distribution for all possible outcomes, it can only provide a final classification, thus unless there was a change in classification the prediction remains the same. From twenty four examined test cases, the decision three change its classification on six of them. Thus, the results show that BIU model is more sensitive to changes than the Police model and the decision tree. However, the question is whereas these models provide a correct changes in the predictions. We answer this below.

We used the three test cases which were explored by experts and we modeled. The first event, Exp. #15, is the 1985 Heysel Stadium Disaster, during the European Cup final. According to Lewis [11] who analyzed this event, one of the reasons for this violent outcome was the police's lack of intervention to prevent the developing violence.

The second event, Exp. #16, is the Los Angeles Riots which occurred in 1991 (55 killed, and over 2000 injured). Useem [15] who analyzed this event, argued that the police were not properly organized and did not react in time with appropriate force to prevent the eruption. This allowed a violent core to grow.

The third event, Exp. #17 , is the London Riot Disaster which occurred in 1990. As opposed to the previous two events, here the police used enormous force against the protests without distinguishing between anarchists and peaceful marchers [14]. What started as a peaceful protest turned to a very violent event with many casualties.

Table 1 presents the experiment results. The first column corresponds to the examined test case. The second column corresponds to recommended police intervention strength. Then we present the models' predictions for each possible outcome: no violence, low violence and high violence. Below each prediction, we present the change, if any, in the recommended prediction. *Dist. Change* denotes a change in the distribution, but not in overall prediction; *Classif. Change* signifies a change in the classification.

The results demonstrate that the decision tree technique is not sensitive to the examined changes that were claimed by the experts. The Police model performed a slightly better than the decision tree (changed the distribution in Exp. #15 but failed in two others test cases). However, the BIU model provided good results which shows that it can account for *what if* scenarios.

Sensitivity Analysis: Experiment 2

In this experiment we want to examine the hypothetical situation of changing the likelihood violence in several test cases scenarios. Specifically, we wanted to examine whether we can decrease even more the violence level in test case 15 (Heysel Stadium disaster) and 16 (LA riots). We used same initializations with several updates as explained below. Some factors such as weather or history of violence cannot be changed, while others can be controlled. For example, police's intervention strength, anonymity, order are examples for features that can be manipulated in the sense that there are actions that can be done to change their values. Police may increase the intervention strength by using more men power or by using different kind of weapon. The existence of projectors and cameras in the gathering zone decrease the perception of anonymity of the participants.

Table 2 presents the experiment results. In this experiment we examined the BIU model and the decision tree. First column corresponds to the examined test case and the second column corresponds to the changed initial values of the quantities. Then we present the models predictions before the change and after.

Here again the results demonstrate that the decision tree technique is not sensitive to changes. This is not surprising, since the only components of the learned tree which can change its classification is the existence of violence core and the police intervention strength. However, the BIU model is found to sensitive to the changes.

5.4 Determining Influential Factors

The previous section demonstrates that while QR models are sensitive to hypothetical changes to the simulated quantities, not all changes can cause a qualitative change in the predictions of the system. In other words, not all factors have equal weight in affecting the outcome of a particular case. Trying all factors, in the hope of identifying those that are important, is a long computer-intense process, that does not scale.

We now turn to evaluating the use of the algorithm described in Section 4 for determining important factors influencing the outcome of the demonstrations. We ran the algorithm on the resulting simulation graphs for the three cases (Exp 15–17) for which we have expert analysis in addition to the predictions of the different models. We requested the 5 highest-entropy nodes. The algorithm analyzed the information associated with them to determine which factors were interacting to cause the different outcomes to form (or more accurately, to create children leading to the different outcomes).

In Table 3 we report on the top factors increasing violence in cases Exp. 15–17. The second and third columns in the table show the factors determined by the algorithm, and the factors determined by experts in the field, who have analyzed these cases.

In case Exp. 15 we see complete agreement between the algorithm and the expert. In case Exp 16. we see partial agreement: Both the algorithm and the expert agree that the police responded with too little strength, but the expert also points out that its intervention occurred too late. The algorithm, in contrast, reports the number of participants as a significant factor in the violence. This is a factor that cannot typically be changed dynamically, but of course the algorithm cannot differentiate static from dynamically-controllable factors (we leave such extension to future work). Note, however, that the algorithm does recognize that the eruption of violence occurs here when, in addition to responding too weakly, the police is too late. However, it does not report on it as a key factor in the eruption of violence. In this, it differs from the expert opinion.

Finally, in case Exp. 17 there is an apparent disagreement between the expert and the algorithm: The expert believes that the main factor accounting for the violence is that the police acted too harshly, while the algorithm points out the existence of a core of demonstrators, and a low perceived chance of punishment, as being key factors. Note, of course, that the algorithm and experts do not provide contradictory results. It could be that both are correct: our algorithm's goal is to discover opportunities for intervention, and it could be that the expert's analysis accounts for a large portion of the state space, in which no intervention is possible (since there, the police acted too harshly, but this cannot be taken back).

Indeed, almost as a side-effect of this analysis, we not only discover which factors are important, but also under what circumstances to act upon them. These circumstances are easily determined by examining the state of the qualitative behavior, as denoted by the node in question. For instance, for Exp. 15, the highest-entropy state (where the algorithm recommends increasing the police response) has the following attributes: Moderate weather, high cohesion of the demonstrators, emotional event, a hard core of demonstrators is present, between 100 and 1000 participants, weak police strength applied, too quickly, lack of a spokesman or representative for the demonstrators, evening hours and dark, property damage already caused by the demonstrators (i.e., low violence already erupted). Under these specific settings, the corrective action to take would be to immediately increase the strength of the police response, in hopes of preventing the violence from escalating.

6 Summary and Future Work

In this paper we described a method for modeling and reasoning about social behavior of large groups, and applied it to

Exp. #	Recommended Change	Model Outcome	**Police Model**	**BIU Model**	**Decision tree**
15	Increase strength [11]	High violence Low violence No violence	66% 0 34%	83% 6% 10%	
	Prediction Change	High	High **Dist. Change**	High **Dist. Change**	High No Change
16	Increase strength [15]	High violence Low violence No violence	66% 0 34%	87% 3% 10%	
	Prediction Change		High No Change	High **Dist. Change**	High No Change
17	Decrease strength [14]	High violence Low violence No violence	80% 0 20%	19% 45% 36%	
	Prediction Change		High No Change	Low **Classif. Change**	High No Change

Table 1: **Experiments results: changed police intervention strength.**

Exp. #	Changed initializations		**BIU Model before change**	**BIU Model after change**	**Dec. Tree before change**	**Dec. Tree after change**
Exp15	Police strength: medium Punishment: high Anonymity: low	High v. Low v. No v.	96% 3% 1%	80% 6% 14%		
	Prediction Change			High **Dist. Change**	High	High No Change
Exp16	Police strength: medium Punishment: high Order: high	High v. Low v. No v.	99% 1% 0%	80% 6% 14%		
	Prediction No Change		High	High **Dist. Change**	High	High No Change

Table 2: **Experiments results: hypothetical manipulations.**

the problem of predicting potential violence during demonstrations. We used qualitative reasoning (QR) techniques, which to our knowledge have never been applied for modeling crowd behaviors, nor in particular demonstrations. Based on social science research, we incrementally presented and compared three QR models for predicting the level of violence in demonstrations: A Base model, Police model and BIU model. We evaluated these models on twenty four real life test cases scenarios. The results show that BIU model makes better predictions than the compared models and it also was found out to be sensitive to changes. We also compared our performances to the machine learning method, a decision tree. While, the machine learning method made an accurate predictions, it fails in the sensitivity analysis. Thus, the BIU model built with QR approach can account for *what if* scenarios is opposed to the decision tree and is more preferable for reasoning regarding the potential violence level to improve the police decision making process In our future work we plan to expand our model to account for bidirectional influences (feedback loops). For example, in the BIU model the ”hostility for the police” quantity increases the violence level. However, increasing the violence level has no impact on hostility. We believe that such expansion is necessary to provide a more accurate prediction. We also plan to tackle the next logical step in the use of QR for social simulation, which is to move beyond determining the important factors, to determining plans of action that utilize them.

References

[1] B. Bredeweg and P. Salles. Mediating conceptual knowledge using qualitative reasoning. In S. Jrgensen, T.-S. Chon, and F. E. Recknagel, editors, *Handbook of Ecological Modelling and Informatics*, pages 351–398. Wit Press, Southampton, UK, 2009.

[2] A. Carmeli and I. Ravid-Yamin. Research report on the subject of crowd events and public order. Technical report, Ministry of public security, Bureau of the chief scientist, 2006.

[3] K. D. Forbus. Qualitative reasoning. In *CRC Handbook of Computer Science and Eng.,*. CRC Press, 1996.

[4] K. D. Forbus and S. E. Kuehne. Towards a qualitative model of everyday political reasoning. In *Proceedings of the Nineteenth International Qualitative Reasoning Workshop*, 2005.

Exp. #	Factors (Algorithm)	Factors (Experts)
15	Police Strength Too Low	Police Strength Too Low [11]
16	Police Strength Too Low, Number of participants > 100	Police Strength Too Low, Too Late [15]
17	Hard Core of Demonstrators, Low Perceived Chance of Punishment	Police Strength Too Much [14]

Table 3: Important Factors in Test Cases Exp. 15–17.

[5] N. Fridman and G. A. Kaminka. Towards a cognitive model of crowd behavior based on social comparison theory. In *AAAI-07*, 2007.

[6] N. Gilbert and K. G. Troitzsch. *Simulation for the social scientist*. Open University Press, 2005.

[7] W. Jager, R. Popping, and H. van de Sande. Clustering and fighting in two-party crowds: simulating the approach-avoidance conflict. *Journal of Artificial Societies and Social Simulation*, 4(3), 2001.

[8] J. Kamps and G. Péli. Qualitative reasoning beyond the physics domain: The density dependence theory of organizational ecology. In *Proceedings of QR95*, 1995.

[9] B. Kosko. Fuzzy cognitive maps: advances in theory, methodologies, tools and applications. In M. Glykas, editor, *studies in fuzziness and soft computing*. Springer, 2010.

[10] B. Kuipers. Qualitative reasoning: modeling and simulation with incomplete knowledge. *Automatica*, 25(4), 1989.

[11] J. M. Lewis. A value-added analysis of the heysel stadium soccer riot. *Current psychology*, 1989.

[12] T. M. Mitchell. *Machine learning*. McGraw-Hill, 1997.

[13] P. Salles and B. Bredeweg. Modelling population and community dynamics with qualitative reasoning. *Ecological Modelling*, 195:114–128, 2006.

[14] C. Stott and J. Drury. Crowds, context and identity: Dynamic categorization process in the ṕoll tax riot. *Human Relations*, 2000.

[15] B. Useem. The state and collective disorders: The los angeles riot/protest of april 1992. *Social Forces*, 76, 1997.

[16] J. Wander. *Modelling consumer behavior*. PhD thesis, University of Groningen, 2000.