2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Banff Center, Banff, Canada, October 5-8, 2017

Hybrid Modeling and Simulation of Tactical Maneuvers in Computer Generated Force

Jang Won Bae
Smart Data Research Group
Electronics and Telecommunications Research Institute
Email: jwbae@etri.re.kr

Bowon Nam, Kee-Eung Kim
Dept. of Computer Science
Korea Advanced Institute of Science and Technology
Email: bwnam, kekim@ai.kaist.ac.kr

Junseok Lee and Il-Chul Moon
Dept. of Industrial and Systems Engineering
Korea Advanced Institute of Science and Technology
Email: junseok.lee, icmoon@kaist.ac.kr

Abstract—**Defense modeling and simulation (DM&S) offers insights into the efficient operations of combat entities, e.g., soldiers and weapon systems. Most DM&S aim at exact description of military doctrines, but often the doctrines fails to provide detail action procedures about how the combat entities conduct military operations. Such unspecified descriptions are filled with the rational behaviors of the combat entities in a battlefield, and thereby the combat effectiveness from these combat entities would differ. Also, by incorporating such rational factors, this could provide the insights that cannot be captured from the traditional works. To examine this postulation, this paper developed a computer generated force where the tactical maneuver of combat entities are realized by the combination of descriptive and prescriptive modeling. Specifically, the descriptive models describe the explicit action rules in military doctrines, and they are modeled using discrete event system specification (DEVS) formalism; the predictive models denoted the rational behavior of the combat entities under the military doctrines, and they are modeled using partially observable Markov decision process (POMDP). The provided results illustrated that the proposed approach helps to maintain a team formation effectively, and this formation maintenance lead to the better combat efficiency.**

1. Introduction

Social phenomena with rare occurrences but high impact, such as disaster response situations and war scenarios, are frequent the important subjects of simulation-based analysis that cope with the dynamism and complexity including rare observations. Furthermore, Modeling and simulating social phenomena inevitably involves social entities, i.e., human beings and organizations, so modeling the behavior of the social entities is a challenge compared to modeling system-oriented systems, e.g., a manufacturing process with simple state transitions.

Some researchers claimed that these social entities are modeled with domain knowledge, and an explicit description would be the best way to complete the modeling task. Descriptive modeling emphasizes such detail descriptions of system and individual behavior (e.g., unified modeling language (UML), discrete event system specification (DEVS) formalism, etc.). On the other hand, other researchers thought that the individual behavior should be modeled to decide the future action through the human reasoning process. Prescriptive modeling asserts that a framework can be a generalized framework to induce the behavior of an individual with optimized parameters (e.g., Partially Observable Markov decision process).

Even if these two lines of social entity modeling have made effects to the modeling and simulation of social phenomena, their usages have been different as they have different characteristics. Descriptive modeling shows an efficiency to macroscopic modes such as crowd simulation that creates social phenomena with simple agents (e.g., rule-based); Prescriptive modeling is more applicable to microscopic models consisting of smaller number of individuals who have high reasoning ability in their actions.

When we narrow the modeled social phenomena to the military domain, the above two modeling methods are applicable to defense modeling and simulation (DM&S). Descriptive modeling specifies the details of the behavioral rules described in military doctrines (e.g., irreducible semi-autonomous adaptive combat (ISAAC) [1] and map aware nonuniform automata (MANA) [2]). Even these descriptive models, which are limited in their reasoning capability, showed collective actions that cannot simply be modeled as the sum of individual behaviors. Meanwhile, Prescriptive modeling generates the complex and uncertain behavior from multiple individuals with a further reasoning process considering their neighbors, situated environment, and utilities.

The rational behavior from prescriptive model is quite different from the guided behavior from the descriptive model, such as unexpected behavior that we did not explicitly model. To identify modeling parts proper to the prescriptive modeling, we focused on the ambiguous parts in DM&S works. Specifically, while behaviors of military units are restricted by military doctrines, there would be a room between the military doctrines: the military doctrines macroscopically describe the behavior of combat entities, but they delegate the specific behaviors to the combat entities situated at the battlefield. For example, when a soldier approaching to the enemy, field manuals dictate utilizing covers during the approaching maneuver, so it is the soldier who choose a cover to hide himself behind. In this sense, we consider that the rational behavior could offer a clue for understanding and optimizing the detailed behavior that are not exist in the military doctrines.

This paper introduces a computer generated force where the maneuver behaviors of combat individuals are modeled via combining descriptive (describing behavior in military doctrines) and prescriptive (describing rational behavior) models. Specifically, the descriptive models describe the explicit action rules in military doctrines, and they are modeled using discrete event system specification (DEVS) formalism; the predictive models denoted the rational behavior of the combat entities under the military doctrines, and they are modeled using partially observable Markov decision process (POMDP). Note that this combined modeling has not been explored so far besides in a few works in the natural resource management [3]. Since the proposed model contains two separate behavioral models, this paper also suggests an interface supporting the interoperation of the two model. The case study using the developed CGF illustrates that the proposed approach helps to maintain a team formation effectively, and this formation maintenance lead to the better combat efficiency.

2. Previous Research in DM&S

DM&S has been utilized as a tool for the analysis of the military systems [4] [5]. This section provides the DM&S backgrounds from the descriptive and the prescriptive modeling.

Descriptive modeling is to capture systematical behaviors in the military systems, which are exemplified as the military doctrines. The state, operator, and result (SOAR) framework is an example of the descriptive modeling from the AI community [6]. The SOAR framework determines autonomous behavior by the combination of production rules and the state set of the model. Also, discrete event system specification (DEVS) [7] is used to develop DM&S works as a descriptive approach.

Having said that, some researchers criticizes that the descriptive modeling is insufficient to fill the unnoticed details in the military doctrines and the field manuals. The field manual generally dictates not details of behavior but abstract process, behavior objectives, and desired states after the behavior. Thus, combat entities have to reflect their experience and knowledge to their situation, i.e., human reasoning procedure, for deciding the optimized action during military missions. This limitation becomes a serious problem in descriptive modeling since the modelers would not be able to put every context and operation change in the state-transition rules. Hence, the demand for prescriptive modeling rises as we start better modeling the fine details of our combat entities.

Prescriptive modeling allows a flexible decision making process under uncertain circumstances. In particular, in DM&S area, the prescriptive modeling has been used to examine various combat strategies [8] [9] [10]. The partially observable Markov decision process (POMDP) [11] is an example of prescriptive modeling.

However, It is infeasible to apply the naive POMDP method to generate the rational behavior of multiple combat entities. Specifically, if the time horizon of such POMDP model is infinite in consideration of a future reward, the time complexity would become Polynomial Space (PSPACE)-complete [12], which induces a high cost even for a small number of actions and states. Although several literatures applied heuristic algorithms to resolve this intractability [13] [14], this scalability problem still hinders the POMDP model from being applied to large-scale models.

This paper begins with an idea that the descriptive and prescriptive methods can work in a complementary manner. Specifically, the prescriptive modeling supports to fulfill the vague parts from the descriptive modeling; on the other hand, the descriptive modeling helps to reduce the computational issues in the prescriptive modeling. To realize this notion, we focused on the characteristics of military domain: mission operations are performed by the guidance of the military manuals, so the behavior of combat entities should be bounded within the manuals. Therefore, those manuals would become references to the reduction of actions and states of the combat entities. To this end, this paper suggests to connect the descriptive and prescriptive modeling with a message-based interface.

3. Computer Generated Force via DEVS and POMDP

In DM&S, combat entities in a simulation model are generally called computer generated forces (CGF). Figure 1 depicts the overall structure of the proposed CGF. The rectangle and the trapezoid boxes represent the DEVS coupled model and the DEVS atomic model, respectively. The numbers next to the lines shows the multiplicity of models.

BlueForce and *RedForce* consists of various combat entity models: *Company* model, *Platoons* models, *MortarTeam*, *Squad* models, *Team* models, and *Soldier*. In particular, the behaviors of combat individuals are modeled by the separation of descriptive and prescriptive parts. The descriptive parts are *Decision-Making*, *Maneuver*, *Fire*, *Detection*, and *DamageEvaluation* models that describe the associated military doctrines; the prescriptive part (Maneuver POMDP) decide proper actions considering their current situation, which is connected with the corresponded *Team* model. The numbers next to a model indicate the number of its instances

3.1. Descriptive parts in CGF

Descriptive parts (DP) in CGF guide combat entities to behave with following the mission objectives and field manuals. This paper adopts DEVS formalism to develop DPs, which supports to model discrete event systems based on set theory [7]. DEVS formalism suggests two types of

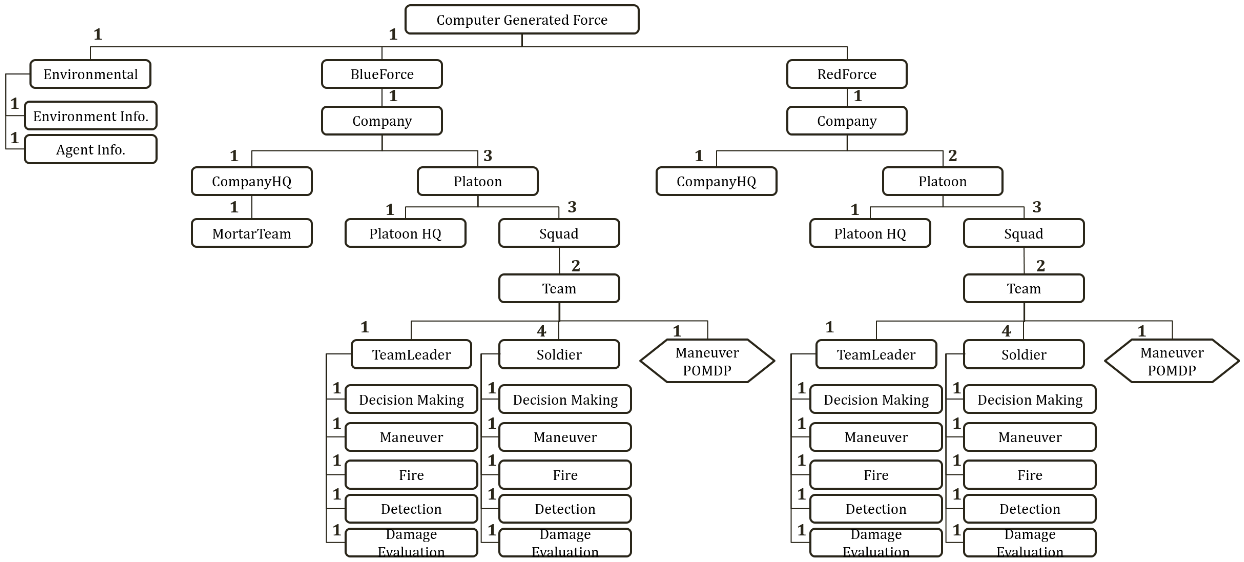

Figure 1. Model structure of the computer generated forces. The tree diagram shows the overall hierarchy of the models for the combat experiment.

submodels: atomic model (AM) for describing system behavior and coupled model (CM) for building up hierarchical model structure (see Figure 1).

Based on the modeling concept of DEVS, CGFs are developed in a bottom-up manner. For example, A *Team* model consists of one *Team leader*, one *Team POMDP*, and four *Soldiers*, and the internal structure of a *Soldier* is depicted in Figure 2. The five submodels in a soldier model are developed as DEVS AMs representing the basic combat actions of the soldier. Specifically, *Maneuver*, *Fire*, and *Detection* models represent moving, firing, and detecting behaviors of a soldier, and these models are controlled by *Decision-Making* model. Lastly, *Evaluation* model evaluates the soldier damage by other soldiers.

Among the solider atomic models, the model diagram of *Decision-Making* model is exemplified in Figure 3, and it is operated in the following way: Initially, the *Decision-Making* model starts from {STOP} state. The *Decision-Making* model receives an order from higher hierarchy level model (i.e., a leader model of its own squad), and this message contains course of action (COA) for the soldier. COA illustrates what a combat entity should perform during missions. The *Decision-Making* model conducts the received COA by activating/deactivating other behavioral models, i.e., the *Maneuver*, *Fire*, and *Detection* models, with sending output events. After changing to one of the {MOVE, FIRE, DISMANTLE} states, the *Decision-Making* model waits for a message claiming the action completion ({MOVE, FIRE}), or it cancels the action by the time advance of {DISMANTLE} state. Subsequently, the *Decision-Making* model delivers a report of the COA to the outside coupled models (Report_out event). If a new order comes during the action executions, the *Decision-Making* model returns to the {DECISION} state and get back to the beginning. Otherwise, if the *Decision-Making* model receives

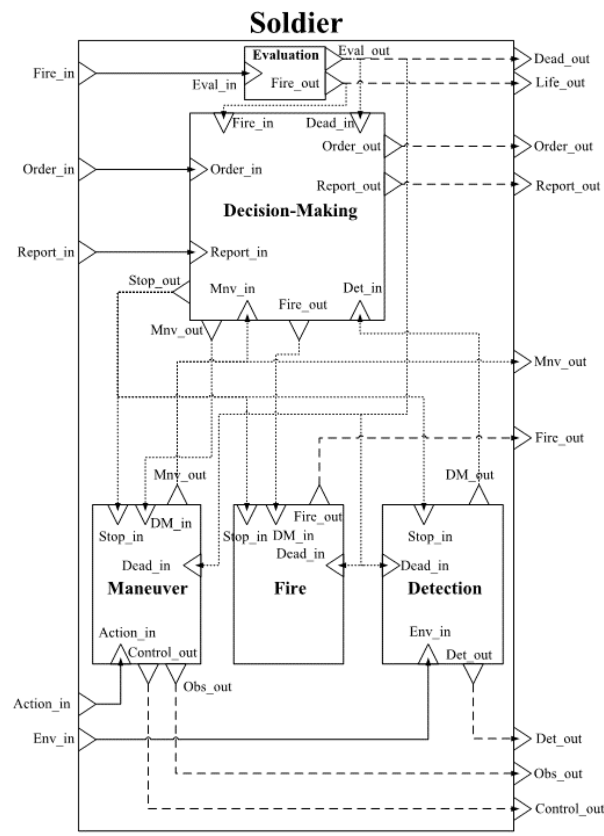

Figure 2. Coupling structure of *Soldier* model consisting of *Decision-Making*, *Maneuver*, *Fire*, *Detection*, and *Evaluation*.

the {*Dead_in*} message from *Evaluation* model, its state changes to {DEAD} state and ceases all functions of the soldier model.

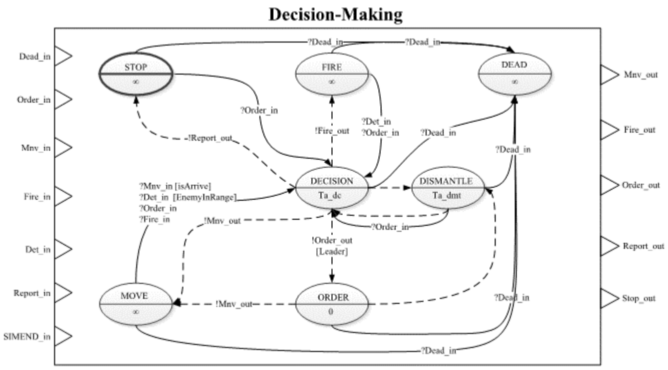

Figure 3. *Decision-Making* model for managing combat behaviors of a soldier model.

3.2. Prescriptive parts in CGF

Prescriptive parts (PP) in CGF decide the proper action of a combat individual with the current situation. To implement PPs, we applied partially observable Markov decision process (POMDP) method. POMDP models a sequential decision process for optimizing a policy, i.e., sequence of actions, under stochastic and partially observable environment [15]. In the proposed model, the prescriptive parts manage the *maneuver* behavior of combat entities. Specifically, the maneuver POMDP decides the moving direction and the step size of each combat entity in the light of the two objectives: maintaining a team formation and faster arrival at the mission point. The following provides detail illustration about how the maneuver POMDP was formulated.

Environmental state (S): environmental state s is defined as $< \vec{\alpha}, \vec{\beta}, \vec{g} >$. $\vec{\alpha}$ represents x, y coordinations of team soldiers ($\vec{\alpha} =< (\alpha_{1x}, \alpha_{1y}), (\alpha_{2x}, \alpha_{2y}), ..., (\alpha_{5x}, \alpha_{5y}) >$). $\vec{\beta} =< \beta_x, \beta_y >$ is the inferred centroid of enemy force's position from observation, and $\vec{g} =< g_x, g_y >$ is a x,y coordination of the team destination position.

Action (A): action a is defined as the step size(r) and the moving direction (Φ) of team soldiers ($< \vec{r}, \vec{\Phi} >$). The step size is a value within $(0, r_{MAX})$ where r_{MAX} means the maximum step size in a time step, and the moving direction is in the range of $(0, 2\pi)$.

Observation (Z): observation value z is defined as $< \vec{\alpha}, \vec{\beta}, \vec{g} >$, which is identical to the environmental state.

State transition function (T): state transition function deals with position changes of team soldiers according to the current state s and an action a, i.e., $(\alpha_{ix}, \alpha_{iy}) \rightarrow (\alpha_{ix} + r_i \cos\Phi_i, \alpha_{iy} + r_i \sin\Phi_i)$. Furthermore, the state transition function considers geographical features, such as elevation and terrain, and these features are expressed as a geographical noise factor ρ. Thus, the position change of a soldier is mathematically represented as follows:

$$(\alpha_{ix}, \alpha_{iy}) \rightarrow (\alpha_{ix} + \rho r_i \cos\Phi_i + N(0,1), \alpha_{iy} + \rho r_i \sin\Phi_i + N(0,1))$$

where ρ:geographical noise and $N(0,1)$: Gaussian noise

Observation function (O): observation value $\vec{\beta}$ depends on the existence of the line of sight (LOS) between team soldiers and opposing soldiers: If LOS is secured, the soldier detects the position of an opposing solider (b_ix,b_iy); otherwise, it does not acquire the position even if the enemy solider is within its detection range.

Reward function (R): each soldier has two objectives in the maneuver mission: the arrival at the goal position and the maintenance of team formation during maneuvering. The reward function $R(s, a, s')$ is defined as a combination of these two objectives.

$$R(s, a, s') = w_{goal} r_{goal}(s, a, s') + w_{formation} r_{formation}(s, a, s')$$

$r_{goal}(s, a, s')$ evaluates how close team soldiers approach the goal position. Hence, $r_{goal}(s, a, s')$ is defined by the potential function $\Phi_{goal}(s)$ that represents a Euclidean distance between the team leader's position and the goal position in state s:

$$r_{goal}(s, a, s') = -(\Phi_{goal}(s') - \Phi_{goal}(s))$$

$r_{formation}(s, a, s')$ is evaluate how perfectly team soldiers keep a team formation. In particular, team soldiers in the proposed model move in the wedge-shaped formation shown in Figure 4.

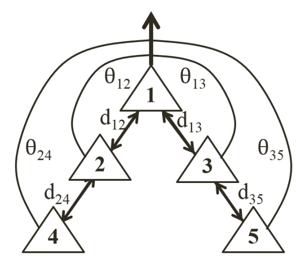

Figure 4. The wedge-shaped maneuver formation of team soldiers.

To maintain the team formation, a soldier calibrates their moving angle and distance with other soldiers. Therefore, the reward function $r_{formation}(s, a, s')$ is factorized for each formation factors: $r_{formation,d}(s, a, s')$ and $r_{formation,\theta}(s, a, s')$ as follows:

$$r_{formation,d}(s, a, s') = -e^{((d_{12}-d)^2+(d_{24}-d)^2+(d_{13}-d)^2+(d_{35}-d)^2)}$$

$$r_{formation,\theta}(s, a, s') = -e^{((\theta_{12}-\theta)^2+(\theta_{24}-\theta)^2+(\theta_{13}-\theta)^2+(\theta_{35}-\theta)^2)}$$

w_{goal} and $w_{formation}$ are weights for the two reward functions. The maneuver POMDP keeps the balance between fast arrival and the formation maintenance by adjusting these weight factors.

3.3. Interfacing descriptive and prescriptive parts

To generate the bounded-rational behavior of combat entities, the PP in the proposed model should be aware of the information from the DP. Although the PP and the DP are developed by different modeling methods, i.e., DEVS and POMDP, they can communicate each other through the interface that coupling them with the semantic relevance of their inputs and outputs.

The interface holds the surrogate models for the DP and PP sides, which are an atomic DEVS model (called Team POMDP) for the descriptive part and *Relational Dynamic influence Diagram Language (RDDL) composer* and *RDDL parser* for the prescriptive part. Team POMDP is a DEVS model for forwarding information from DPs to PPs and waiting its response from the PPs response. RDDL modules deal with planning domain definition language (PDDL) to specify a set of decision-making problems. In the interface, *RDDL Composer* receives the observation information from the connected DPs and *RDDL Parser* translates the received information into a set of planning problems for POMDP models.

The surrogate models exchange their information via three messages: *Control*, *Observation*, and *Action*. *Control* message specifies the mission objective, so it is sent from *Team leader* or high-level units. *Observation* message contains observed information by combat entities. These messages are generated from DPs and sent to PPs through the interface. Otherwise, *Action* message holds rational behaviors generated by PPs based on the received *Control* and *Observation* messages, and it is sent to the corresponded DPs through the interface. Table 1 presents detail information of the three messages.

4. Combat Experiments

We designed the combat scenario for the virtual experiments, which consists of maneuver, breakthrough, charge, and occupation phases (see Figure 5). At the maneuver phase, blue platoons move to their mission points, and *RedForces* mount guard at the occupied hill; at the breakthrough phase, while *BlueForce* engaged with *RedForces*, the 2nd platoon in *BlueTeam* makes a breakthrough in the obstacle area that is set up by *RedForces*; after the obstacle is dismantled, at the charge phase, and the 1st platoon in *BlueForce* moves to the breakthrough point and the 2nd and the 3rd platoons charges to *RedForces*. On the other hand, *RedForces* attack the approaching *BlueForce* soldiers until their combat ability falls below 30% of the initial status; If the combat abilities of all red soliders go below 30%, *RedForce* retreat out of the battlefield, which is the end of the scenario.

The combat experiments using the proposed model was conducted with following the above scenario. The experiments focus on the influence of the use of maneuver POMDP. Hence, we evaluated the combat efficiency from the proposed model and the no-use of maneuver POMDP, i.e., DEVS-only model.

We defined two performance measures to evaluate the effectiveness in the combat simulation. One measure is formation violation penalty (FVP), which indicates the fitness of the formation while a team maneuvers in the wedge formation (see Figure 4).

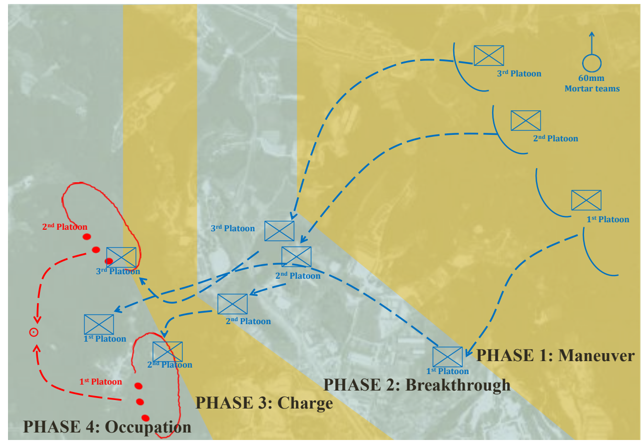

Figure 5. Combat scenario with four phases: maneuver, breakthrough, charge, and occupation

$$Formation\ Violation\ Penalty\ (FVP) = E[fvp]$$

$$fvp = fvp_d + C \times fvp_\theta$$

$$fvp_d = (d_{12} - d)^2 + (d_{24} - d)^2 + (d_{13} - d)^2 + (d_{35} - d)^2$$

$$fvp_\theta = (\theta_{12} - \theta)^2 + (\theta_{24} - \theta)^2 + (\theta_{13} - \theta)^2 + (\theta_{35} - \theta)^2$$

$$C : \text{scaling constant}$$

The other measure is damage ratio (DR). DR indicates the change of the damage state of soldiers in a platoon during simulation execution. Thus, higher DR represents lower combat efficiency. Mathematical definition of DR of a platoon (DR_P) is as follows:

$$DR_P = \sum_{s \in S_P} \frac{\frac{InitD_s - EndD_s}{InitD_s}}{|S_P|},$$

where S_P: a set of soldiers in platoon P, $InitD_s$: initial damage state of soldier s, and $EndD_s$: damage state of soldier s at the end of simulation,

Figure 6 illustrates FVP and DR values with respect to the use and the no-use of the maneuver POMDP. The results represents that the use case significantly reduces the FVP, which means the decisions provided from the maneuver POMDP is efficient to maintain the team formation. Moreover, this formation maintenance, although more experiments and statistical analysis are required, helps to reduce the DR of *BlueForce* and to increase the DR of *RedForce* (i.e., more combat effectiveness of *BlueForce*).

5. Conclusion

Descriptive and Prescriptive modeling have been applied to DM&S works, yet there raised concerns about the opponents' limitations from each side: the descriptive side raised concerns on the computational issues and the uninterpretable results from the prescriptive models; the prescriptive side

TABLE 1. Messages exchanged between descriptive (DEVS) and prescriptive (POMDP) models

Message Type	Message Contents	Message Sender	Message Receiver
Control	Current position	Team leader	Maneuver POMDP
	Destination position	Team leader	Maneuver POMDP
	Maximum speed	Team leader	Maneuver POMDP
Observation	Current position	Soldier	Maneuver POMDP
	Detected enemy	Soldier	Maneuver POMDP
	Geographic information	Soldier	Maneuver POMDP
	Detected enemy	CompanyHQ	Maneuver POMDP
Action	Moving step	Maneuver POMDP	Team leader/Soldier
	Moving Direction	Maneuver POMDP	Team leader/Soldier

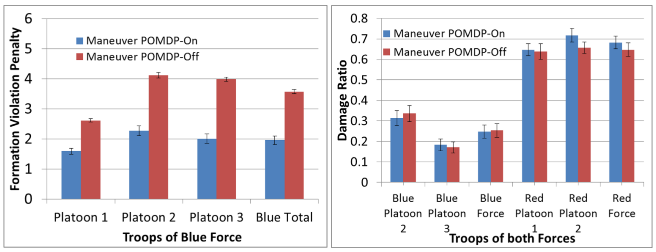

Figure 6. Formation violation penalty (FVP) and Damage ratio (DR) values with respect to use or no-use maneuver POMDP.

worried about the vague descriptions between the descriptive models. Having said that, this paper explained that these concerns can be resolved by combining the two approaches.

To this end, this paper proposes a computer generated force where the maneuver of combat entities are modeled by the separation of the descriptive and prescriptive parts. This separation was considered only on the maneuver operation, and the descriptive and the prescriptive parts are developed with DEVS and POMDP, respectively. Moreover, this paper suggests an interface that can communicate between those parts. The experimental results with the computer generated force developed by the proposed approach showed that the proposed method provides an efficiency to increase the combat effectiveness, even if the further statistical analysis is needed.

As further works, we considered extending the separated modeling approach to more combat behaviors and performing more sophisticated experiments to clearly identify the benefits and losses generated from the proposed method.

Acknowledgments

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (R7117-16-0219, Development of Predictive Analysis Technology on Socio-Economics using Self-Evolving Agent-Based Simulation Embedded with Incremental Machine Learning)

References

[1] A. Ilachinski, "Irreducible semi-autonomous adaptive combat (isaac): An artificial-life approach to land combat," *Military Operations Research*, vol. 5, no. 3, pp. 29–46, 2000.

[2] J. Moffat, J. Smith, and S. Witty, "Emergent behaviour: theory and experimentation using the mana model," *Advances in Decision Sciences*, vol. 2006, 2006.

[3] J.-C. Castella, S. Pheng Kam, D. Dinh Quang, P. H. Verburg, and C. Thai Hoanh, "Combining top-down and bottom-up modelling approaches of land use/cover change to support public policies: Application to sustainable management of natural resources in northern vietnam," *Land use policy*, vol. 24, no. 3, pp. 531–545, 2007.

[4] J. Kim, I.-C. Moon, and T. G. Kim, "New insight into doctrine via simulation interoperation of heterogeneous levels of models in battle experimentation," *Simulation*, vol. 88, no. 6, pp. 649–667, 2012.

[5] I.-C. Moon, K. M. Carley, and T. G. Kim, "Modeling and Simulating Command and Control for Naval Air Defense Operation," in *Modeling and Simulating Command and Control*. Springer, 2013, pp. 35–66.

[6] J. E. Laird, A. Newell, and P. S. Rosenbloom, "Soar: An architecture for general intelligence," *Artificial intelligence*, vol. 33, no. 1, pp. 1–64, 1987.

[7] B. P. Zeigler, H. Praehofer, and T. G. Kim, *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems*. Academic press, 2000.

[8] D. J. Bryant, F. M. J. Lichacz, J. G. Hollands, and J. V. Baranski, "Modeling situation awareness in an organisational context: Military command and control," *A cognitive approach to situation awareness: theory and application*, 2004.

[9] D. Aberdeen, S. Thiébaux, and L. Zhang, "Decision-Theoretic Military Operations Planning." in *ICAPS*, 2004, pp. 402–412.

[10] J. Schmitt and G. Klein, "A recognitional planning model," DTIC Document, Tech. Rep., 1999.

[11] E. J. Sondik, "The optimal control of partially observable Markov processes." DTIC Document, Tech. Rep., 1971.

[12] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Mathematics of operations research*, vol. 12, no. 3, pp. 441–450, 1987.

[13] J. Pineau, G. Gordon, S. Thrun, and Others, "Point-based value iteration: An anytime algorithm for POMDPs," in *IJCAI*, vol. 3, 2003, pp. 1025–1032.

[14] S. A. Miller, Z. A. Harris, and E. K. Chong, "Coordinated guidance of autonomous uavs via nominal belief-state optimization," in *American Control Conference, 2009. ACC'09*. IEEE, 2009, pp. 2811–2818.

[15] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1, pp. 99–134, 1998.

[16] T. G. Kim, C. H. Sung, S.-Y. Hong, J. H. Hong, C. B. Choi, J. H. Kim, K. M. Seo, and J. W. Bae, "DEVSim++ toolset for defense modeling and simulation and interoperation," *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 2010.