Layered Behavior Modeling via Combining Descriptive and Prescriptive Approaches: A Case Study of Infantry Company Engagement

Jang Won Bae, Junseok Lee, Do-Hyung Kim, Kanghoon Lee, Jongmin Lee, Kee-Eung Kim, *Member, IEEE*, and Il-Chul Moon

***Abstract*—Defense modeling and simulation (DM&S) has brought insights into how to efficiently operate combat entities, such as soldiers and weapon systems. Most DM&S works have been developed to reflect accurate descriptions of military doctrines, yet these doctrines provide only guidelines of military operations, not details about how the combat entities should behave. Because such vague parts are often fulfilled with the appropriate behavior of combat entities in a battlefield, one part argues that DM&S should consider individual combat behaviors as well. However, it is known as an infeasible problem discovering best individual actions from infinite searching space, such as the battlefield. This paper proposes a layered behavior modeling to practically resolve this issue. The proposed method applies descriptive modeling to reduce the searching space by employing domain-specific knowledge; and prescriptive modeling to discover best individual actions in the reduced space. For the generalization, the proposed method adapts both modeling methods being modularized, and then the proposed method suggested an interface between them that is based on their semantic analogies. Both modeling methods are modularized, so they are interacted through an interface defined in the proposed method. This paper presents a realization of the proposed method through a case study of infantry company-level operations. In the case study, the proposed method is implemented with discrete event system specification formalism as the descriptive part and Markov decision process as the prescriptive part. The experimental results illustrated that the combat effectiveness resulted from the proposed method is statistically better than that from the descriptive-only modeling, and the difference would be guided by the objective of the combat behavior. Through the presented experimental results and the discussion, this paper argues that future DM&S should consider a broad spectrum from the battlefield incorporating the rational behavior of military individuals.**

***Index Terms*—Computer generated force (CGF), defense modeling and simulation (DM&S), descriptive modeling, layered modeling, prescriptive modeling.**

I. Introduction

MODELING and simulation has been utilized as an efficient method for analyzing and evaluating military systems [1]–[3], and this field is generally referred to as defense modeling and simulation (DM&S). The current success in DM&S is partially rooted in the applicability of the modeling and simulation techniques that support analyzing the situations that have not yet occurred, but are expected to in the future (e.g., anticipating the outcome of potential combat scenarios).

These scenarios consist of the cooperation among multiple combat individuals in military systems, such as weapon system and military rank structure. To improve the combat effectiveness of military systems, military experts suggested improving their systematic efficiencies, in particular, considering their hierarchical command structure [4]. These experts designed a guideline for managing the military systems efficiently, and this experience has been documented and refined up to now. Such documents are generally referred to as military doctrines or field manuals. From this perspective, DM&S works have been developed to evaluate military systems by models based on the military doctrines [5]–[7]. In this sense, therefore, descriptive modeling was considered as an appropriate method for DM&S.

Descriptive modeling aims at explicitly illustrating the system mechanisms that steer the system behavior, and the military doctrines are regarded as the core of those mechanisms in DM&S. Specifically, descriptive modeling intends to describe military doctrines as they are, and the developed model is used to evaluate and analyze combat effectiveness. For example, Yun *et al.* [8] presented agent-based modeling and simulation to identify the key tasks of company-level commanders by analyzing the command and control systems of the infantry company.

Meanwhile, other researchers raised concerns about applying only systematic analysis to military systems. They instead focused on individual modeling in the military systems that

Manuscript received January 18, 2018; accepted March 18, 2018. Date of publication April 23, 2018; date of current version June 16, 2020. This work was supported by Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korea government(MSIT) (1711065140, Development of Predictive Analysis Technology on Socio-Economics Using Self-Evolving Agent-Based Simulation Embedded with Incremental Machine Learning). This paper was recommended by Associate Editor R. Qiu. *(Corresponding author: Il-Chul Moon.)*

J. W. Bae is with the Smart Data Research Group, Electronics and Telecommunications Research Institute, Daejeon 34129, South Korea (e-mail: jwbae@etri.re.kr).

J. Lee, D.-H. Kim, and I.-C. Moon are with the Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, South Korea (e-mail: jslee@kaist.ac.kr; dhkim@kaist.ac.kr; icmoon@kaist.ac.kr).

K. Lee, J. Lee, and K.-E. Kim are with the Department of Computer Science, Korea Advanced Institute of Science and Technology, Daejeon 305-701, South Korea (e-mail: khleecs.kaist.ac.kr; jmleecs.kaist.ac.kr; kekim@cs.kaist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2018.2822312

actually carries out combat behavior, so they suggested considering human reasoning process behaviors in DM&S [9], [10]. Furthermore, we focused on another reason for the above concerns originated from inherent limitations in the military doctrines: the military doctrines describe the behaviors of the general combat situations, but this generality often fails to provide specific behaviors in a certain situation. For example, due to the lack of specific representations of the rules in the military doctrines, combat individuals should determine or infer their future behaviors based on their experiences and situated environments. It also means that there might be room that the military doctrines does not fully cover.

Prescriptive modeling is to model human reasoning processes in several decision-making fields, such as operation research and artificial intelligence. Through this reasoning process, an individual could determine his rational behavior considering his current situation. The rational behavior implies the behavior patterns based on the utility theory, often mentioned by economists [11], [12]. For example, a rational agent modeled with Markov Decision Process (MDP) would select their behaviors as maximizing the expected sum of rewards [13], [14]. Generally, this rational behavior differs from the guided behavior from the descriptive modeling: prescriptive modeling can generate unexpected behaviors that are not explicitly coded in the model. In DM&S cases, for instance, prescriptive modeling is able to provide the rational behavior that a combat entity could autonomously behave when the military doctrines fails to offer explicit guides. Having said that, the utilization of prescriptive modeling in DM&S has been inherently limited, because it requires infeasible computation time for discovering infinite searching space on the combat situation [15].

One way to make the prescriptive modeling being more practical is reducing its searching space. This reduction, however, is an intricate problem because it is at risk for losing or distorting the original information. Having said that, in some domain-specific problems, such as military systems, we assumed that well-refined domain knowledge, such as field manuals, would be used to reasonably reduce the searching space. To implement this idea, we applied the descriptive modeling with domain knowledge for reducing searching space and the prescriptive modeling on the reduced searching space to find approximately best actions in the reduced searching space (see Fig. 1). Having said that, we also note that this approximation holds tradeoff relationships between the applicability by the reduced searching space and the effectiveness of the selected actions from the reduced space, which can be measured in an empirical manner.

In terms of military systems, we can rephrase the above assumption with detail examples: combat individuals in the military systems are generally highly trained to follow the military doctrines prior to their individual thoughts. This distinctive feature backs up the assumption that the best actions of combat individuals in a battlefield would fall into the boundary of the field manuals. In other words, their combat behaviors would be placed within the boundary of the military doctrines, and human reasoning process can find out their best behaviors for the current situations. We consider that this approach is

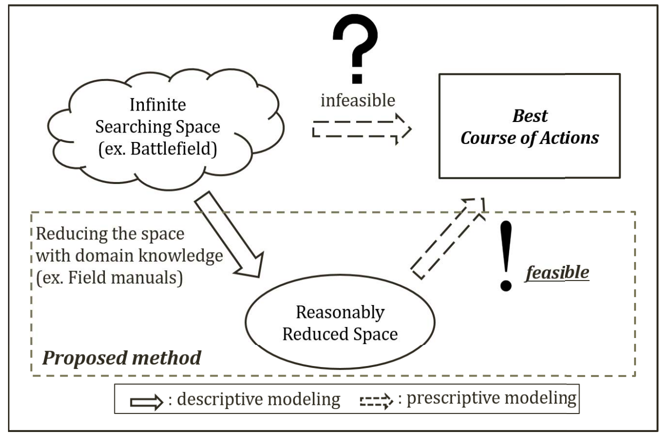


Fig. 1. Conceptual view of the proposed method: finding approximated course of actions by utilizing prescriptive method on the reduced searching space by descriptive method.

expected to provide more efficiency rather than the descriptive-only method by filling the gap among the military doctrines with the human reasoning process. Therefore, we also expect that this idea would give an answer to military experts standing on both descriptive and prescriptive modeling sides.

To this end, this paper proposes layered behavior modeling by interfacing the descriptive and the prescriptive modeling methods. Both modeling methods are modularized, so the proposed method combines them with an interface mapping their semantic analogies. In particular, we note that this loosely coupled structure makes the proposed method being generalized, which will be covered in the discussion section. We note that the proposed layered modeling has rarely been applied so far in DM&S domain, but it has in a few works in the natural resource management [16]. This paper also presents a case study of a company-level engagement model. The case study introduces an agent-based model conducting a company-level occupation mission, and combat individuals in the model conduct the mission with considering the environmental factors that influence their combat behaviors, such as maneuver and fire. To build a layered behavior structure for the case study, we employed discrete event system specification (DEVS) formalism as descriptive modeling (for describing the military doctrines) and MDP with continuous state and action spaces as prescriptive modeling (for describing the rational behavior of combat individuals). From the result analysis, the proposed method statistically showed better combat efficiency rather than the conventional method (i.e., descriptive-only model). For further insights, this paper also discussed the strong and weak points, the distinctive benefits, and the generality of the proposed method. We expect that the proposed method will provide a modeling way of embedding individual rational behavior to systematical behavior.

II. Previous Research in DM&S

DM&S has been used as a tool for researchers and practitioners in the analysis of the military domain [17], [18], especially for training [19], acquisition [20], [21], and analysis [22], [23]. Among the DM&S works, in particular, computer generated forces (CGF) have been developed to evaluate combat efficiency. CGF is defined as "A generic term used to refer to computer representations of entities in simulations which attempts to model human behavior sufficiently so that the forces will take actions automatically (without requiring man in-the-loop interaction)" [24].

By its definition, CGFs contains both systematic (e.g., mission operations) and individual (e.g., human behavior) features at once. From the systematic view, CGFs focus on military doctrines that explicitly denote the basic principles and guidelines for military operations, and combat entities in the CGFs are designed to follow the military doctrines. One method of developing such CGFs is a descriptive modeling method. Descriptive modeling is specialized in presenting the explicit rules about system behavior [25]–[27]. Hence, descriptive modeling is applied to illustrate military doctrines and field manuals, which dictate basic rules of military missions, so it has been applied to various defense models [28], [29]. Examples of descriptive modeling are the state, operator, and result (SOAR) framework [30], [31] and DEVS formalism [32], [33]. The SOAR framework determines autonomous behavior by the combination of production rules and the state set of the model; the DEVS formalism uses algebraic representation consisting of sets and functions to represent systematic behaviors so that it enables to realize system of systems concept [34].

On the other hand, some researchers said that the descriptive modeling was insufficient for fulfilling the unnoticed details between the military doctrines [9], [10], [35]. They argued that the military doctrines generally care about macroscopic rules at war, but a combat is conducted through the cooperation of multiple combat entities. Specifically, combat entities have to reflect their experience and knowledge according to their situations for investigating the best action during military missions [36]–[38].

Prescriptive modeling considers for decision-making process under uncertain circumstances. In particular, in the DM&S area, prescriptive modeling has been used to examine various combat strategies [39]–[43]. MDP is an example of prescriptive modeling [44]–[46]. MDP generates a sequence of behaviors in the stochastic environment, so it enables the suggestion of the rational behavior of combat entities with respect to their battlefields, which is not described in the military manuals.

Most of the current practices in DM&S are either macroscopic (e.g., military doctrines) or microscopic (e.g., combat entity behavior) only. Due to the characteristics of the military, DM&S emphasizes systematic behaviors based on the field manuals. However, it is required to consider that due to the vague regions between the military doctrines, the combat effectiveness could significantly be changed.

To investigate and clarify those vague regions, DM&S should be performed from the multiple perspectives. Few cases considering both factors together have been conducted [47]–[54], but they still hold limitations in terms of model-scale, applicability, and extendability issues. For example, [47] developed an agent-based model applying human reasoning process via human-in-the-loop manner, but it is

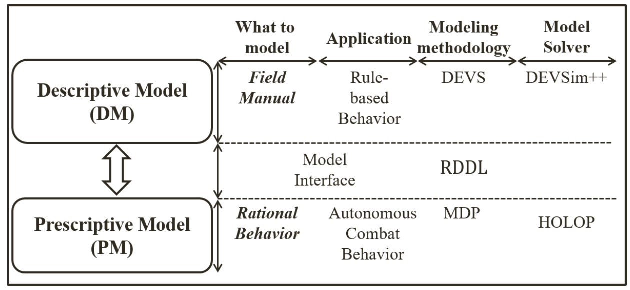


Fig. 2. Structure of the layered combat behavior modeling applied to the company-level CGF. The behaviors of the combat entities are described by the combination of two modeling methods: the DM and PM represents the field manual (i.e., what to do) and rational behavior (i.e., how to do), respectively.

considered for small number of scenarios and implemented as *ad-hoc* manner. Furthermore, [48], [49], [51], [53], [54] developed their agent-based models embedding reasoning process with simulation frameworks, such as MANA, Pythagoras, JANUS, ACOMSIM, and Repast, but they covered squad-level or individual military unit operations which still remains a question of the applicability. Hocaoglu [50] and Elkosantini [52] proposed a simulation framework for decision process but did not present a practical example.

This paper proposes layered behavior modeling consisting of the descriptive and prescriptive methods. Specifically, we employed formal methods, i.e., DEVS and MDP, in the proposed approach and interfaced them by connecting their elements. The following sections introduce the details of the proposed method with an explicit example and also analyze its efficiency through virtual experiments.

III. Layered Behavior Modeling Using Company-Level Engagement Operations

Layered behavior modeling was suggested to embed rational behavior into artificially designed systems (e.g., embedding rational combat behavior into military systems). In particular, we designed the proposed method as a practical one meaning not to discover optimal values in infinite time, but to find acceptable and applicable solutions in reasonable time. Hence, the proposed layered method was designed to hierarchical form and its parent and child part have different roles: the parent part, as descriptive model, reduces problem regions with domain-specific knowledge; and the child part, as prescriptive model (PM), discovers the best strategy (or behavior) on the reduced problem regions. By explicitly separating and interacting two modeling method, we note that the proposed method is capable of being applied to domain problems with the similar characteristics.

Based on this design concept, this section illustrates how this concept would be implemented through the model development of a company-level CGF. For better understanding, it starts with introducing the objective and structure of the CGF.

A. Company-Level CGF With Layered Behavior Model

The company-level CGF describes an occupation mission where one infantry company of the blue force intends to take over an enemy spot where one infantry company of the red

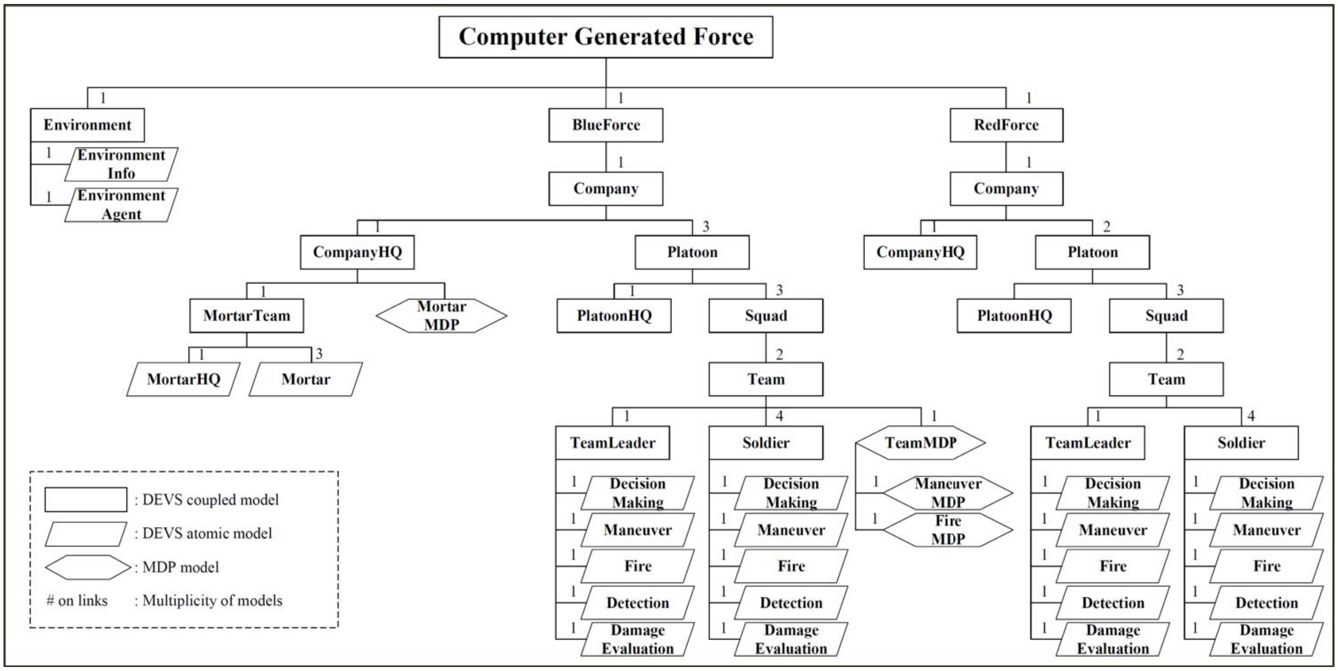


Fig. 3. Model structure of the company-level CGF. The rectangle and the trapezoid boxes represent the DEVS CM and AM, respectively. The numbers next to the lines show the multiplicity of the models. In particular, *BlueForce* uses the MDP models to reason combat entity behavior.

force is located. The combat entities in the company-level CGF are soldiers, mortar teams, team leaders, and mortar headquarters: the soldiers and mortar teams are engagement units, and the others are command and control (C2) units.

To incorporate behavior reasoning process in the blue force, we designed it with layered behavior modeling (see Fig. 2). The layered behavior modeling consists of two main parts: descriptive model (DM) and PM. Specifically, the DEVS model, as a DM, describes military doctrines regarding the maneuvers, direct fire, and mortar fire of company units; MDP, as a PM, reasons the combat entity behaviors of the maneuver, direct fire, and mortar fire on the conditions of the military doctrines and the surrounded environment. Also, to communicate between DM and PM, there is an interface between DM and PM.

Fig. 3 depicts the overall structure of the company-level CGF. The *BlueForce* model consists of one *Company* model, and the *Company* model has *CompanyHQ* and three *Platoon* models. *CompanyHQ* delivers missions to the *Platoon* models and to the *MortarTeam* models, and the bombardment order to *MortarTeam* by controlling the *Mortar MDP* model, which determines the bombing target for each mortar. The *Platoon* model has a similar structure as the *Company* model, i.e., it has one *PlatoonHQ* and three *Squad* models. The *Squad* model has two *Team* models, and each *Team* model has five combat entities, such as *Team Leader* and *Soldier*, and one *Team MDP* model. *Team leader* generates an order to *Soldiers* according to the combat mission from the *PlatoonHQ* and reports the result of the mission to the *PlatoonHQ*. Each combat entity consists of five behavioral models: 1) decision-making; 2) maneuver; 3) fire; 4) detection; and 5) damage evaluation. *Team MDP* model has linkages with two MDP models: 1) maneuver MDP and 2) fire MDP. Each MDP model determines the associated behaviors of the combat entities. Also, the company-level CGF contains the *Environment* model, which provides battlefield information (e.g., elevation and terrain type) used in the behavior reasoning process. In Fig. 3, the rectangle and the trapezoid boxes represent the DEVS coupled model (CM) and atomic model (AM), respectively. The numbers next to the lines shows the multiplicity of models.

B. Descriptive Model in the Company-Level CGF

DMs in the company-level CGF guide the combat entities to behave considering the mission objectives and field manuals. We applied DEVS formalism to descriptive modeling. DEVS formalism suggests two types of submodels: 1) the AM for describing system component behavior and 2) the CM for building up a hierarchical model structure. Table I presents the formal specifications of the two models [for reference, $M.e$ in Table I indicates an event port (e) associated with a model (M)].

In the development of the company-level CGF, DEVS CM illustrates the hierarchical structure of friendly and enemy troops, and DEVS AM represents individual-level combat behaviors, such as decision-making, maneuvers, fire, detection, and damage evaluation. Through this building process, CGFs are developed in a bottom-up manner: for example, one soldier on a team consists of five AMs regarding combat behaviors, and that team is also a component model of a squad. Fig. 4 shows the coupling structures of the *BlueForce* model. Fig. 4 (left) is the composition structure of the troop hierarchy of the *BlueForce*. The numbers next to a model indicate the number of its instances; for example, there is one company model in the *BlueForce*, and a *Company* model has three *Platoon* models. In particular, a *Team* model consists of one *Team leader*,

TABLE I
MATHEMATICAL REPRESENTATIONS FOR DEVS MODELS: AM AND COUPLED MODEL (CM)

Atomic Model: AM = $< X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta >$	Coupled Model: CM = $< X, Y, M, EIC, EOC, IC, Select >$
X: a set of input events Y: a set of output events S: a set of states δ_{ext}:$Q \times X \rightarrow S$, an external transition function $(Q = \{(s, e) \mid s \in S, 0 \leq< e \leq ta(s)\})$ $\delta_{int} : S \rightarrow S$, an internal transition function $\lambda : S \rightarrow Y$, an output function $ta : S \rightarrow R^+$, a time advance function	X: a set of input events Y: a set of output events M: a set of DEVS models in CM $(\forall M_i, M_j \in M(i \neq j))$ $EIC \subseteq CM.X \times \cup M_i.X$: external input coupling relations $EOC \subseteq \cup M_i.Y \times CM.Y$: external output coupling relations $IC \subseteq \cup M_i.Y \times \cup M_j.X$: internal coupling relations $Select$: tie-breaking function

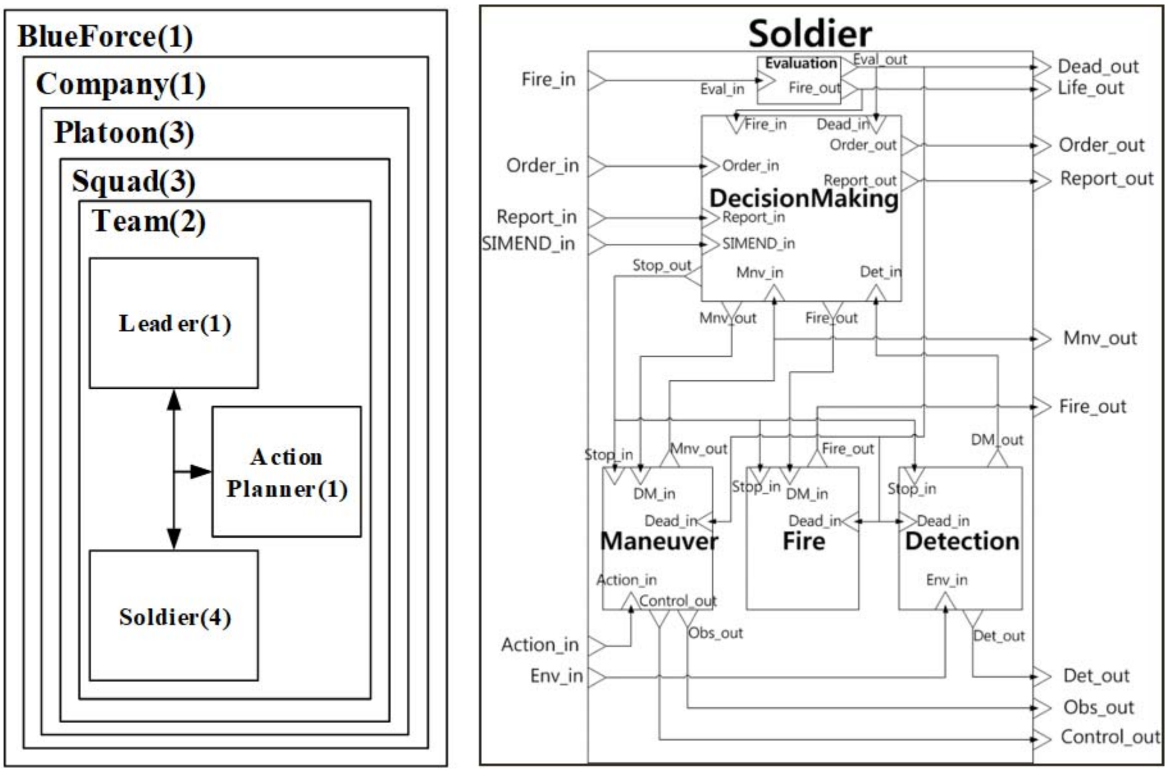


Fig. 4. Coupling structure of the blue force in the company-level CGF. Left: Composition of the troop hierarchy of *BlueForce* in the simulation (The numbers next to the model names indicate the number of the model instances). Right: *Soldier* model by DEVS CM consists of five DEVS AMs for combat behaviors, such as *Decision-Making*, *Maneuver*, *Fire*, *Detection*, and *Evaluation*.

one *Team MDP*, and four *Soldiers*, and the internal structure of a *Soldier* is depicted in Fig. 4 (right).

The five submodels in a soldier model represent the basic combat behaviors of a soldier. The *Maneuver*, *Fire*, and *Detection* models are controlled by a central model called the *Decision-Making* model. *Evaluation* model evaluates the soldier damage caused by other soldiers. For example, we present the diagram of *Decision-Making* model (see Fig. 5), which represents the decision-making procedure of a soldier.

The company-level CGF contains an environment model, where environmental factors, such as the terrain and elevation of the battle scene, and the location of blue and red agents. When making a decision, all combat entities obtains agent information at time t from the environment model with their limited scopes (e.g., detection range). Also, terrain factors in the environmental factors are categorized into three types: 1) mountain; 2) open field; and 3) road, and they have an influence on maneuvering and firing behaviors. The elevation factors, on the other hand, affects to detecting behavior by interfering in securing the line of sight (LOS) between blue and red units. Such interferences are modeled as noise factors in MDP models.

C. Prescriptive Model in the Company-Level CGF

The PM is proposed for determining a proper action of a combat entity considering the currently situated environments. To implement PMs, we applied MDP. MDP provides an mathematical framework for modeling decision making based on the Markovian property. The formal definition of MDP is represented as a tuple $<S, A, T, R, \gamma>$ (see Table II).

MDP is in a certain state $s_t \in S$ at each time t, and there are several actions $a_t \in A$ that are available in state s_t. By conducting an action, the next state (s' or $s_{t+1} \in S$) is stochastically

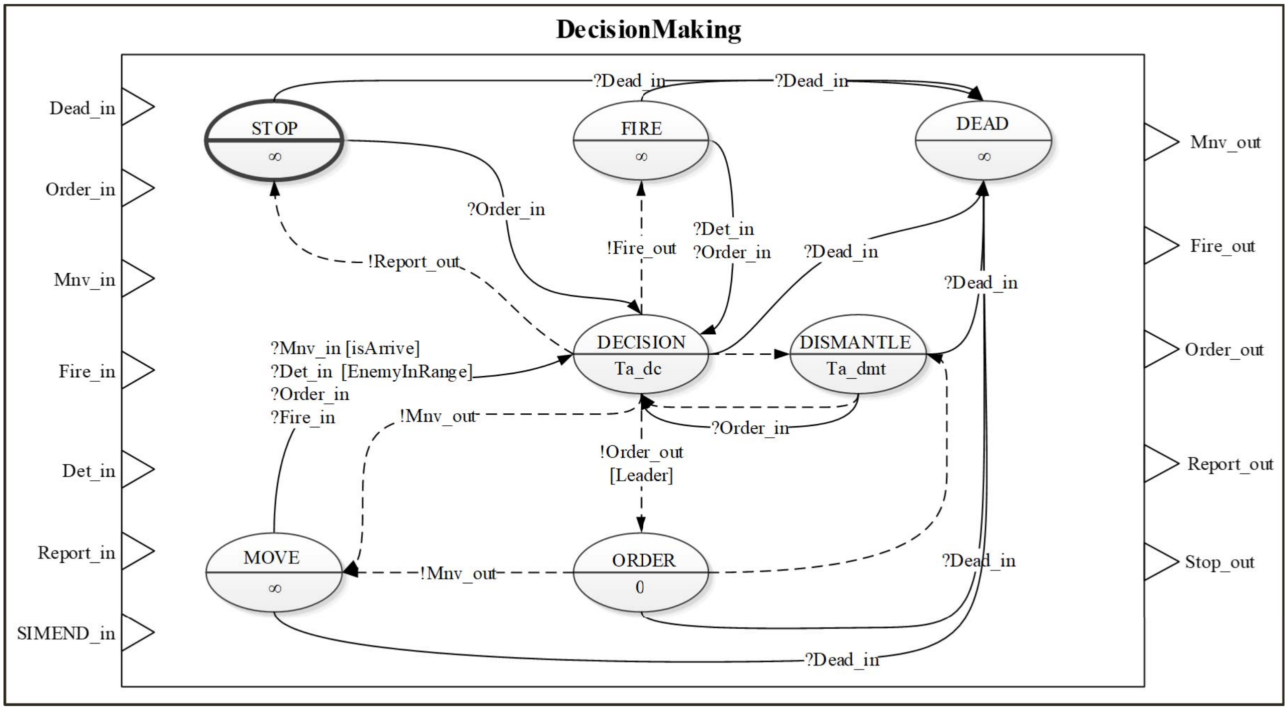


Fig. 5. DEVS AM diagram of *Decision-Making* model, managing combat behaviors of a soldier model.

TABLE II
MATHEMATICAL DEFINITION OF MDP MODELS

$MDP = < S, A, T, R, \gamma >$
S: a set of environmental states
A: a set of possible actions of agents
T: $S \times A \times S \rightarrow [0, 1]$, a transition function of a state ($s \in S$) being changed to a state $s' \in S$ as a result of an action $a \in A$
R: $S \times A \times S \rightarrow \mathbb{R}$, a reward function giving a reward of an action $a \in A$ at the current and the next state (s and $s' \in S$)
$\gamma \in$ [0, 1], discount rate for the future rewards

determined by state transition function T, and the corresponding reward is determined by reward function R. Also, the rewards for future states are discounted by the factor of γ.

The company-level CGF contains MDP models to manage *maneuver*, *fire*, and *mortar* behaviors of combat entities. The following section introduces how each MDP model is designed in the company-level CGF by specifying its formal description.

1) Maneuver MDP: Maneuver MDP decides the moving direction and speed of each soldier considering the balance between two objectives: 1) keeping the team formation where the soldier is involved and 2) arriving at the mission point as soon as possible. This concept was designed as the following MDP formulation.

Environmental State (S): Environmental state s is defined as $<\vec{\alpha}, \vec{\beta}, \vec{g}>$. $\vec{\alpha}$ represents x, y coordinations of five soldiers on a team [i.e., $\vec{\alpha} = <(\alpha_{1x}, \alpha_{1y}), (\alpha_{2x}, \alpha_{2y}), \ldots, (\alpha_{5x}, \alpha_{5y})>$]. $\vec{\beta} = <\beta_x, \beta_y>$ is the centroid of the enemy force's position, and $\vec{g} =< g_x, g_y >$ represents the team destination position that the red team occupied in the scenario.

Action (A): Action a is defined as moving speed (v) and direction (Φ) of team soldiers ($< \vec{v}, \vec{\Phi} >$). The moving speed is a value within $[0, v_{\text{MAX}}]$ where v_{MAX} means the maximum speed, and the moving direction is in the range of $[0, 2\pi)$.

State Transition Function (T): The state transition function deals with the position changes of team soldiers according to the current state s and an action a. Specifically, the state transition function provides the speed of each soldier considering the situated environments. The geographical noise, ρ, consists of two components: 1) terrain (ρ_t) and 2) elevation (ρ_e) noise. ρ_t is determined by three terrain types; *mountain* as 0.6, *open terrain* as 0.8, and *road* as 1.0; ρ_e is calculated with respect to the elevation degree, θ_e [i.e., $\rho_e = 1 - (c * \theta_e)$, c is a constant from the ROK field manual]. ρ is therefore calculated by the multiplication of those environmental factors. With the provided speed of each soldier, their next positions are calculated

$$\left(\alpha'_{ix}, \alpha'_{iy}\right) \rightarrow \left(\alpha_{ix} + \rho r_i \cos \Phi_i, \alpha_{iy} + \rho r_i \sin \Phi_i\right) \tag{1}$$

where $\rho = \rho_t * \rho_e$:geographical noise, r_i : moving distance ($r_i = s_i * t$, s_i as soldier speed and t as moving duration).

Reward Function (R): Each soldier has two objectives in the maneuver mission: 1) the arrival at the mission position as soon as possible and 2) the maintenance of team formation during maneuvering. These objectives are designed as coefficient factors in the reward function

$$R\left(s, a, s'\right) = w_A * r_A\left(s, a, s'\right) + w_F * r_F\left(s, a, s'\right) \tag{2}$$

where w_A and w_F are the weights for the arrival and the formation reward, respectively.

$r_A(s, a, s')$ evaluates how close team soldiers approach the mission point. Mathematically, $r_A(s, a, s')$ is defined as the potential function $\Psi_{goal}(s)$, which represents a Euclidean distance between the team leader's position and the goal position in state s [i.e., $r_A(s, a, s') = -(\Psi_{goal}(s') - \Psi_{goal}(s))$].

TABLE III
MESSAGES EXCHANGED BETWEEN THE SURROGATE MODELS OF DM AND PM: CONTROL, OBSERVATION, AND ACTION; MESSAGE CONTENTS ARE BASED ON THE SEMANTIC RELEVANCE BETWEEN DEVS AND MDP

Message Type	Message Contents	Message Sender	Message Receiver	Corresponding Value
Control	Current position	Team leader	Maneuver MDP	$\vec{\alpha}$
	Current life status	Team leader	Fire MDP	$\vec{\alpha_s}$
	Destination position	Team leader	Maneuver MDP	$\vec{g}$
	Maximum speed	Team leader	Maneuver MDP	$\vec{r}$
	Current life status	CompanyHQ	Mortar MDP	$\vec{\alpha_s}$
Observation	Current position	Soldier	Maneuver MDP	$\vec{\alpha}$
	Red life status	Soldier	Fire MDP	$\vec{\beta_s}$
	Detected enemy	Soldier	Maneuver/Fire MDP	$\vec{\beta}$
	Geographic information	Soldier	Maneuver MDP	ρ
	Red life status	CompanyHQ	Mortar MDP	$\vec{\beta_s}$
	Detected enemy	CompanyHQ	Mortar MDP	$\vec{\beta}$
Action	Next location	Maneuver MDP	Team Leader/ Soldier	$\vec{\alpha}'$
	Fire target ID	Fire MDP	Team Leader/ Soldier	$R_i,\ (i \in 1, , 5)$
	Fire target point	Mortar MDP	Mortar	$R_i,\ (i \in 1, , 3)$

$r_F(s, a, s')$ evaluates how perfectly team soldiers maintains the team formation. (e.g., the wedge-shaped formation).

To maintain the team formation, a soldier calibrates his current position with other soldiers, which means $r_F(s, a, s')$ is factorized by the x and y coordination of the team soldiers [i.e., $r_F(s, a, s') = \sum_{i=2}^{5} \sqrt{(\tilde{x_i} - x_i)^2 + (\tilde{y_i} - y_i)^2}$, where $\tilde{x_i}$ and $\tilde{y_i}$ are the expected x and y coordination of soldier i to maintain the team formation according to soldier 1's position]. Also, we assumed the discount rate, γ, of Maneuver MDP to be 0.95;

2) Fire MDP and Mortar MDP: Fire MDP and Mortar MDP determines the fire target of each soldier and the target point to support the infantry with mortar fires, respectively, to maximize the combat effectiveness. The following description illustrates how Fire MDP was formulated, and due to their behavioral similarity, Mortar MDP is explained by noting the differences with Fire MDP.

Environmental State (S): Environmental state s is defined as $<\vec{\alpha}, \vec{\alpha_d}, \vec{\beta}, \vec{\beta_d}>$. $\vec{\alpha}$ and $\vec{\alpha_d}$ are the position and the damage state of team soldiers, respectively. The damage state of a soldier ($\alpha_{d,i}$) is initially set to 100 and decreases by the fire damage from the opposing force. If $\alpha_{d,i}$ goes down to below 30 or becomes zero, the soldier becomes disabled or dead. Similarly, $\vec{\beta}$ and $\vec{\beta_d}$ are the position and the damage state of the detected soldiers in the red force, respectively. In Mortar MDP, $\alpha_{d,i}$ $(i = 1, 2, 3)$ does not change because the mortar model is designed not to be damaged by the enemy fire.

Action (A): An action $a = <R_1, R_2, \ldots, R_5>$ represents the fire targets of team soldiers. Specifically, R_i indicates a red soldier or a location targeted by a blue soldier i (in Fire MDP case), or a bombing target of mortar i (in Mortar MDP case).

State Transition Function (T): The damage states of combat entities (α_d and β_d) are changed by the transition function of Fire and Mortar MDP

$$\beta'_{d,i} = \beta_{d,i} - D_i \tag{3}$$

where $\beta_{d,i} \in \vec{\beta_d}$ and D_i is fire damages to soldier $i, D_i \leq 100$.

Damage from the Fire MDP is determined by *hit probability* and *fire power*. Hit probability is a random variable derived from Bernoulli distribution with the accuracy rate of a rifle. The rifle accuracy rate is assumed as 90% within its fire range; otherwise 0%. Having said that, when the LOS between the shooter and target is not obtainable by the elevation, the accuracy rate is also 0%. Fire power is modeled as a parameter, called probability of killing (see Table IV). In the mortar MDP case, we modeled that damage by the mortar fire are affected by terrain type and elevation the elevation degree between the impact area and the force garrison.

Reward Function (R): The objective of fire and mortar MDP is to increase the number of the disabled red force. Thus, these models are designed to get more reward when more opposing soldiers' life status (i.e., β_s) go below 30. Also, we set the discount rate, γ, of fire and mortar MDP as 0.95.

D. Model Interface Between DM and PM

While DM describes the military doctrines and current combat situations, PM can infer rational behavior regarding maneuvering and firing operations within the military doctrines. To support such interoperation, we designed an interface for exchanging information from the DEVS and MDP models. Specifically, we resolved this issue by coupling the two models with the semantic relevance of the inputs and outputs of those models.

The protocols used in the suggested interface consist of three messages: *Control*, *Observation*, and *Action*. *Control* specifies: 1) the mission objectives generated from *Team leader* or high-level units and 2) the current life status of the team (α_s). *Observation* contains the observed information, including red force units and the situated environments. These two messages are sent from DM to PM through the interface. On the other hand, *Action* message, forwarded from PM to DM, indicates the behaviors of the associated combat entities according to the received *Control* and *Observation* messages. Table III presents the details and the semantic relevance between the DEVS and MDP of the three messages.

Fig. 6 illustrates an interface applied to the *Team* model. The interface holds surrogate models for DM and PM, which are *Team MDP* and *relational dynamic influence diagram language (RDDL) Composer* and *RDDL Parser*. *Team MDP* is

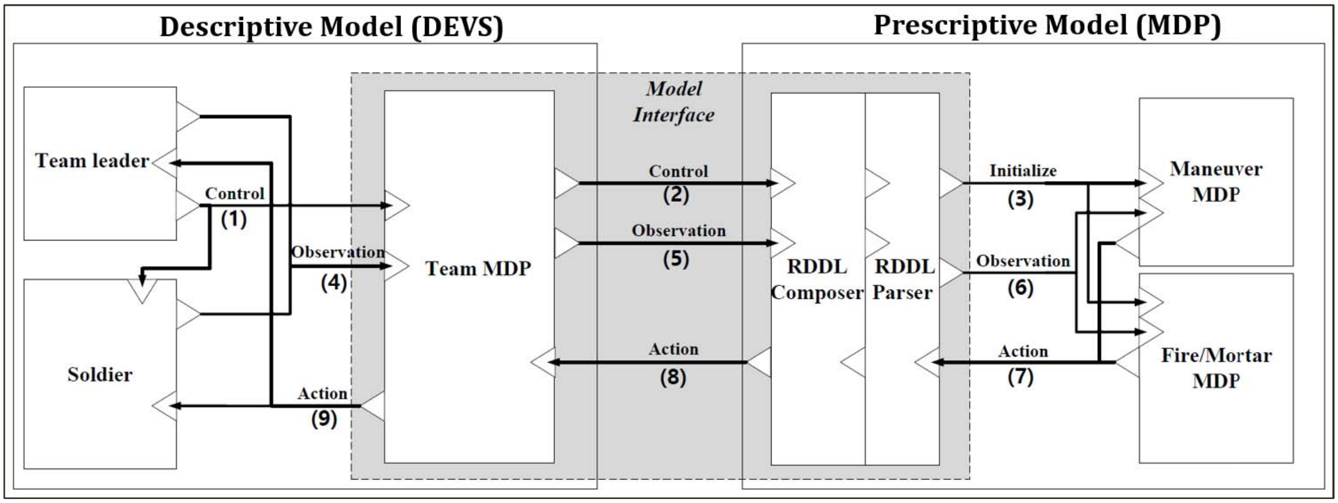


Fig. 6. Interface structure (shaded area in the figure) between DM and PM: the interface consists of two surrogate models in both side, and they exchanges their information through three messages, such as control, observation, and action.

a DEVS AM for transforming model states into the interface protocols (see Table III) and forwarding them to PMs. Using RDDL [55], which is a planning domain definition language [56], to specify a set of decision-making problems, *RDDL Composer* receives the observation information, and *RDDL Parser* translates the received information into a set of planning problems for MDP models.

Let us elaborate the interface procedure between DM and PM: (1) when a *Team leader* decides to conduct a course of action, it sends a *Control* message to its *Team MDP* and other *Soldiers*; (2) the *Team MDP* model forwards the *Control* message to RDDL composer; (3) *RDDL Composer* extracts the required information from the *Control* message and translates it into RDDL format. Then, *RDDL Parser* interprets the translated message and initializes either the *Maneuver MDP* or *Fire/Mortar MDP* solver according to the message contents; (4)–(6) meanwhile, *Soldier* models that received the *Control* message (including the *Team Leader* model) send an *Observation* message through the previous sequence; and (7)–(9) using the received information, *Maneuver MDP* and *Fire/Mortar MDP* calculate each soldier's next action and return the calculated result to the *Soldier* models by the *Action* message. Team soldiers repeat steps (4)–(9) until they complete the current mission.

We used DEVSim++ [7] as the development tool for the DMs and applied hierarchical open-loop optimistic planning algorithm as an MDP solver because it is able to describe continuous state and action spaces and multiple-horizon MDP planning [57]. The interface between DM and PM was implemented using TCP/IP communication.

IV. Virtual Experiments

We performed virtual experiments with the developed company-level CGF. The objective of the experiments is to evaluate the change of the combat effectiveness when the reasoning process is considered. The following section introduces the performance measures and the virtual experimental designs.

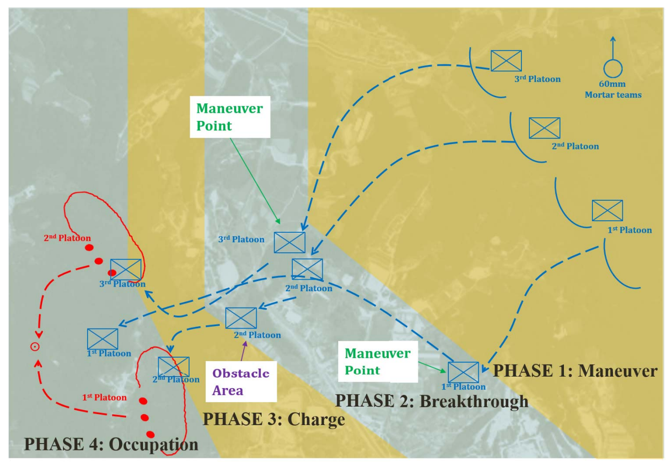


Fig. 7. Initial deployment and maneuver routes of blue and red forces according to four scenario phases: maneuver, breakthrough, charge, and occupation.

A. Scenario and Performance Measures

Before introducing our virtual experimental design, we illustrate an infantry engagement scenario using the company-level CGF. The scenario proceeds with four phases, such as *maneuver*, *breakthrough*, *charge*, and *occupation* (see Fig. 7). Initially, there is a hill occupied by *RedForce* with two *Platoons*, and *BlueForce* with three *Platoons* and *Mortar teams* are deployed to take over the hill.

1) At the maneuver phase, Platoons in *BlueForce* move to their maneuver points, and *RedForce* mounts guard at the occupied hill.
2) At the breakthrough phase, the first Platoon and *Mortar* team of *BlueForce* engaged with *RedForce*, while the second Platoon makes a breakthrough in the obstacle area set up by *RedForce*
3) After the obstacle is dismantled, at the charge phase, the second and the third Platoons charges to *RedForces*, and the first Platoon follows them a little way off.

TABLE IV
VARIABLE LIST IN THE COMPANY-LEVEL CGF; VARIABLES ARE CATEGORIZED BY INPUT VARIABLE, OUTPUT VARIABLE, AND MODEL PARAMETERS

Variable Type	Variable Name	Value (unit)	Implications
Input Variable	Soldier: Fire damage of rifle	Varied in scenarios	Damage of rifle bullet when it hits a soldier
	Maneuver MDP model		Whether or not to apply the MDP model for maneuver action
	Fire MDP model		Whether or not to apply the MDP model for fire action
	Mortar MDP model		Whether or not to apply the MDP model for mortar fire action
	Arrival weight		Arrival weight of maneuver MDP(W_A); $W_F = 1 - W_A$
Output Variable	Formation violation penalty (FVP)	-	Performance measure for evaluating the deviation of formation
	Damage ratio (DR)		Performance measure for evaluating the engagement result
Model Parameters	Dispersion angle of formation	120(°)	Angle between soldiers to form a wedge formation
	Dispersion distance of formation	2.2(m)	Distance between soldiers to form a wedge formation
	Soldier: Maximum speed	3.33(m/s)	Maximum moving speed of a soldier
	Soldier: Probability of killing	70(%)	Probability of a bullet hitting the target soldier
	Soldier: Fire range of rifle	50(m)	Effective firing range of rifle
	Soldier: Detection range	150(m)	Detection range at which a soldier can perceive an enemy soldier
	Mortar: Fire damage of mortar	100(%)	Damage of mortar ammunition at the center
	Mortar: Explosion radius of mortar	13.5(m)	Any soldiers in the explosion radius of mortar get damage

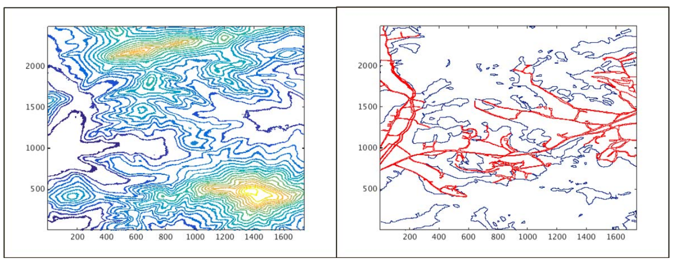


Fig. 8. Contour map and terrain types (red line: road, blue line: mountain line) of the battlefield in the scenario: around *Mt. Bongseo* in South Korea.

4) When the damage states of *RedForces* fall down to 30%, the occupation phase begins, and *BlueForce* takes over the hill.

The behavior of the combat entities are affected by their situated environments, so the battlefield of the scenario should be considered. For practical experiments, we employed a real-world environment as the battlefiled: around the *Mt. Bongseo* in South Korea (lat: 37.8377, long: 126.8180). Fig. 8 depicts the elevation and terrain types of the battlefield.

Based on the above scenario, we defined two performance measures: 1) formation violation penalty (FVP) and 2) damage ratio (DR). FVP indicates the fitness of the formation while a team maneuvers in the wedge formation, by calculating the violation of the x and y coordination of team soldiers to maintain the wedge formation. Therefore, FVP is mathematically identical to the reward function of formation (r_F) in Maneuver MDP

$$FVP = \sum_{i=2}^{5} \sqrt{(\tilde{x_i} - x_i)^2 + (\tilde{y_i} - y_i)^2} \tag{4}$$

where $\tilde{x_i}$ and $\tilde{y_i}$ is the expected x and y coordination of soldier i, respectively, to maintain the team formation according to soldier 1's position.

DR represents a ratio for how much damage a Platoon suffered during simulation execution. Thus, a higher DR represents that soldiers in the corresponding platoon received more damage during engagement, which indicates lower combat effectiveness. Mathematically, the DR of a Platoon (DR_P) is defined as follows:

$$\mathrm{DR}_P = \frac{1}{|S_P|} \sum_{s \in S_P} \frac{100 - \mathrm{End}D_s}{100} \tag{5}$$

where S_P: soldiers in platoon P, and EndD_s: damage state of soldier s at the end of simulation.

B. Experimental Design

Table IV presents variables used in the company-level CGF. Excluding the output variable defined in the above section, there are two kinds of variables; one group of the variables comes from real data or field manuals, which is presented as the parameter (e.g., fire range of rifle); the other is added to test the performance of our treatments, which is presented as the input variable (e.g., applying Maneuver MDP model or not).

Using the input variables, we designed virtual experiments. The experiments were conducted to see the influence of the proposed layered structure of the behavior model on combat effectiveness. Specifically, combat behaviors in the company-level CGF are developed with either the proposed layered modeling or the descriptive-only modeling (i.e., the typical modeling practice). Table V shows the experimental design of the battle experiment. The use of Maneuver, Fire, and Mortar MDP models indicates the usage of the MDP models in the combat behaviors. Also, we varied the arrival weight (W_A) in Maneuver MDP by 0.05, 0.5, and 0.95, which means formation weight (W_F) also changed by 0.95, 0.5, and 0.05. Fire damage represents the rifle damage of soldiers. Through the full-factorial experimental design, a total of 36 experimental cells are created, and each cell was replicated for 10 times. The experiments were performed in the following environment: Windows 7, Intel i7 6700K 4.0GHz, 32GB RAM.

V. RESULT ANALYSIS

This section provides analyses on the simulation result from the experimental design. The result analysis consists of

TABLE V
VIRTUAL EXPERIMENTAL DESIGN OF THE BATTLE EXPERIMENT. THE USE OF MANEUVER, FIRE, AND MORTAR MDP MODELS, ARRIVAL WEIGHT (W_A) IN THE MANEUVER MDP, AND FIRE DAMAGE ARE CONSIDERED

Variables	Experiment Design	Implication
Maneuver MDP	Use/Not Use (2 cases)	Whether or not to use the MDP model in maneuvers
Fire MDP	Use/Not Use (2 cases)	Whether or not to use the MDP model in fire
Mortar MDP	Use/Not Use (2 cases)	Whether or not to use the MDP model in mortar
Arrival Weight	0.05, 0.5, 0.95 (3 cases)	Weight of arrival reward of maneuver MDP; $W_F = 1 - W_A$
Fire damage	10, 60, 90 (3 cases)	Fire damage of soldier's rifle
Total cases	2×2×2×3×3=72 cells	Each cell is replicated for 10 times

TABLE VI
n-WAY ANOVA RESULTS ON THE FVP AND THE DR OF BLUE AND RED FORCES (BOLDED FIGURES INDICATE p-VALUE UNDER 0.05)

Variables	d.f.	FVP		BlueDR		RedDR	
		F	p-value	F	p-value	F	p-value
Maneuver MDP	1	36.25	**0.000**	174.23	**0.000**	124.4	**0.000**
Fire MDP	1	0.00	0.993	0.85	0.358	1.06	0.305
Mortar MDP	1	0.07	0.794	51.20	**0.000**	3.19	0.076
Fire Damage	2	0.16	0.853	232.9	**0.000**	50.89	**0.000**
Maneuver MDP * Fire MDP	1	0.03	0.862	0.74	0.391	2.97	0.087
Maneuver MDP * Mortar MDP	1	0.06	0.812	0.23	0.635	0.06	0.800
Maneuver MDP * Fire Damage	2	0.06	0.946	7.59	**0.001**	2.77	0.066
Fire MDP * Mortar MDP	1	0.04	0.841	2.56	0.112	2.2	0.141
Fire MDP * Fire Damage	2	0.00	0.998	0.06	0.945	0.41	0.665
Mortar MDP * Fire Damage	2	0.01	0.991	0.49	0.611	0.8	0.452

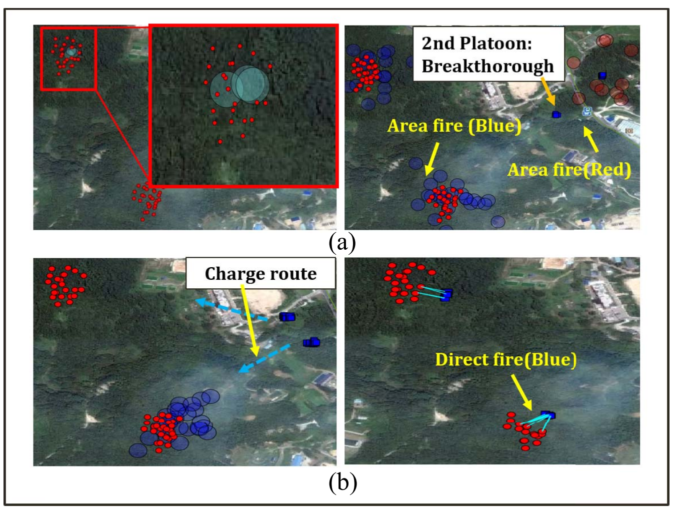


Fig. 9. Visualization of the infantry engagement scenario. (a) Breakthrough and (b) charge.

statistical analysis on the performance measure and illustration of maneuvering with MDP modules.

A. Statistical Analysis on the Performance Measure

Fig. 9 shows screen captures from the provided scenario: breakthrough and charge. Small blue squares and red circles represent blue and red soldiers, respectively. Blurred and large circles represent the coverage of the area fire by each force. Also, mortar and direct fire engagements are represented by bigger circles and lines.

Statistical analysis is performed with two objectives: 1) statistical differences in the performance measures by introducing the combat behavior reasoning process (i.e., MDP models) and 2) significant factors among the input variables that affected the performance measures. To this end, we conducted n-way analysis of variance (ANOVA), t-test, and meta-modeling with regression analysis to the simulation results.

N-way ANOVA is applied to see whether statistical means of the performance measures differ by the variation of the input variables, such as MDP models usages or not and fire damages. The null hypothesis of the ANOVA is that statistical means among the variations are same (i.e., $\mu_1 = \mu_2 = \cdots = \mu_n$, where $1, 2, \ldots, n$ indicate different experimental cases), and it could be rejected at the 95% confidence interval when the associated p-value is under 0.05. Table VI presents the ANOVA results where: 1) the usage of Maneuver MDP causes a significant mean difference to FVP, the DR of the blue force (BlueDR) and the red force (RedDR); 2) the usage of Mortar MDP has an influence on the mean difference of BlueDR, but not RedDR; and 3) fire damage is also a significant factor to the DR measures in the ANOVA test.

The ANOVA test presents which input variables can significantly affect the mean of the combat effectiveness, but further works are required to see the details of the differences. For example, the ANOVA results show that there is a significant difference via the usage of Maneuver MDP, but that difference does not indicate whether the mean of Maneuver MDP usage is significantly greater than that of the no usage case, and *vice versa*. Therefore, one-sided t-test was performed. The null hypothesis of the right-tailed t-test is $\mu_1 \leq \mu_2$, so it can present a significant comparison of the means of the two groups.

One-sided t-tests were conducted with considering the usage of MDP models: Maneuver MDP (arrival weight $W_A = 0.95$, formation weight $W_F = 0.95$, and no MDP), Fire MDP (use and no MDP), and Mortar MDP (use and no MDP). In Table VII, fire

TABLE VII

t-TEST RESULTS ON THE FVP AND THE DR OF BLUE (BLUEDR) AND RED FORCES (REDDR) (BOLDED FIGURES INDICATE THE SIGNIFICANTLY BIGGER GREATER RATHER THAN THE OPPOSITE ONE): ARRIVAL: $W_A = 0.95$, FORMATION: $W_F = 0.95$; NOTE THAT t-STATS AND p-VALUES ARE CALCULATED FROM THE RIGHT-TAILED t-TEST

	Maneuver MDP						Fire MDP		Mortar MDP	
FVP	Arrival	Formation	Arrival	No MDP	No MDP	Formation	Use MDP	No MDP	Use MDP	No MDP
Mean	**4249.30**	540.38	**4249.30**	1547.20	**1547.20**	504.38	2162.10	2160.40	2134.60	2187.90
Std	393.13	263.39	393.13	140.21	140.21	263.39	1352.70	1292.40	1354.60	1289.90
t-stat	38.40		49.96		23.95		0.01		-0.24	
p-value	0.00		0.00		0.00		0.50		0.60	
BlueDR	Arrival	Formation	Arrival	No MDP	No MDP	Formation	Use MDP	No MDP	Use MDP	No MDP
Mean	**0.27**	0.21	**0.27**	0.17	0.17	**0.21**	0.20	0.21	**0.19**	0.22
Std	0.06	0.06	0.06	0.07	0.07	0.06	0.07	0.07	0.07	0.07
t-stat	3.30		5.96		-2.48		-0.38		-3.03	
p-value	0.00		0.00		0.99		0.65		1.00	
RedDR	Arrival	Formation	Arrival	No MDP	No MDP	Formation	Use MDP	No MDP	Use MDP	No MDP
Mean	0.68	**0.75**	**0.68**	0.64	0.64	**0.75**	0.67	0.68	0.67	0.68
Std	0.45	0.30	0.45	0.04	0.04	0.30	0.05	0.06	0.06	0.06
t-stat	-6.83		3.50		-11.90		-0.64		-1.11	
p-value	1.00		0.00		1.00		0.74		0.87	

damage was dropped in the t-test because its influence on the DR measures is quite conjecturable. Table VII shows the results from the t-test. Bold figures in the table represents that those figures are greater than the opposite one; for example, the FVP mean of arrival weight = 0.95 cases (i.e., 4249.30) is significantly greater than that of formation weight = 0.95 cases (i.e., 540.38) at the 0.05 significance level. For a better understanding of the result interpretation, we note that better combat effectiveness is relevant to lower FVP, lower BlueDR, and higher RedDR by the definition of the performance measures. The result table explains that: 1) in the maneuver MDP case, using an MDP model with a higher formation weight shows the best combat performance (i.e., lower FVP, lower BlueDR, and higher RedDR), but using an MDP with higher arrival weight causes lower combat effectiveness even rather than no MDP case; 2) the usage of Mortar MDP shows significantly better performance only for BlueDR; and 3) there is no significant difference between the use and no use of Fire MDP.

The t-test results show that the FVP can be reduced by utilizing the formation-weighted (W_F) MDP hybrid model compared to the pure DEVS model and the arrival-weighted (W_A) MDP hybrid model. This result indicates two advantages of utilizing the layered model. First, the layered model can minimize the unexpected losses adaptively compared to the pure DM. The FVP can only be minimized by adaptively repositioning soldier maneuver given the terrain environment. The DM may not produce adaptation rules without extensive scenario and environment-based modeling while the MDP part of the layered model can provide such adaptations. Second, the layered model can diversify the behavior model through parameter changes. If a model is purely descriptive (e.g., DEVS model only), the model is often extended through further descriptions. This often complicates the model, and eventually this limit the diversification on the model behavior. However, the MDP part can be controlled by the reward setting parameter (e.g., W_A and W_F in the Maneuver MDP), so the implicit policy from the MDP can be diversified, and eventually the model behavior. This adaptive nature is one characteristic that the DM cannot obtain.

TABLE VIII

STANDARDIZED COEFFICIENTS OF THE LINEAR REGRESSION MODELS (ROWS ARE INDEPENDENT VARIABLES, AND COLUMNS ARE DEPENDENT VARIABLES, AND BOLD FIGURES INDICATE p-VALUE UNDER 0.05)

	Standardized Coefficients		
	FVP	BlueDR	RedDR
Arrival weight	**0.397**	**0.122**	**-0.279**
Maneuver MDP	**0.276**	**0.405**	**-0.467**
Fire MDP	0.004	-0.020	-0.056
Mortar MDP	-0.002	**-0.214**	**-0.079**
Fire damage	-0.27	**0.647**	**0.412**
Adjusted R^2	0.228	0.644	0.470

We also conducted meta modeling analysis to determine the effectiveness and robustness of each input variable for the performance measures. Specifically, the meta modeling analysis was conducted with the linear regression method, and its result is presented in Table VIII. The standardized coefficients in the table represent the effectiveness of the independent variable to the associated dependent variable, and their p-value under 0.05 (i.e., bold figures in the table) indicates their statistical significance. Hence, the following interpretations comes from the result table: 1) FVP depends only on the maneuvering variables (e.g., arrival weight and the use of Maneuver MDP), and, in particular, arrival weight is the strongest affecting factor for FVP; 2) the use of Mortar MDP significantly affects on the DR measures, but the use of Fire MDP does not; and 3) by considering the relationship of $W_F = 1 - W_A$, maintaining the formation statistically leads to better performance on DR measures (i.e., decreasing BlueDR and increasing RedDR).

B. Illustration of Platoon Maneuvering With MDP

The combat effectiveness generated from applying MDP modules to DM&S is analyzed in the previous section, but the proposed method also provides detail individual behaviors during the simulation executions. For example, by the layered structure of descriptive and prescriptive methods, the proposed method provides the detail behaviors illustrating how individuals autonomously react to the situated environment

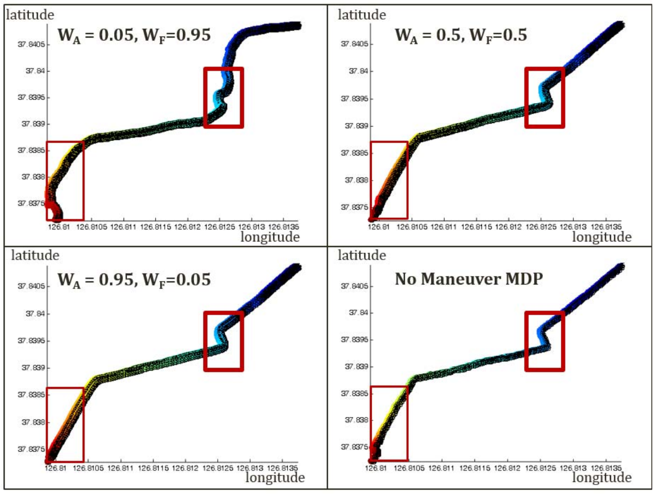


Fig. 10. Illustration of the maneuver routes of the first team of the second Platoon in the Blueforce as the maneuver coefficients (W_A and W_F) changes: red boxes highlight noticeable differences among four figures.

changes, which eventually makes the developed model being more adaptive or robust to environmental changes. We note that such illustrations are hardly obtained from the previous approach (i.e., descriptive-only methods), and they would be used to analyze current systems and design next systems.

In the case study, the proposed method enables us to incorporate Maneuver MDP that determines individuals' movement considering their terrain information, such as terrain type and elevation,which eventually makes the developed model being more adaptive to the situated environments. Specifically, this section focuses on the influence on the platoon maneuver by changing reward coefficients in Maneuver MDP [i.e., weight for formation (W_F) and arrival (W_A, $W_F = 1 - W_A$)].

Fig. 10 depicts the maneuvering route variations of the first team of the second platoon in the Blueforce as the maneuver coefficients changed. Furthermore, we highlighted the major difference among the subfigures with red boxes. With the *No Manever MDP* case as the baseline, the subfigures represent that the moving trajectory changes as the Maneuver MDP was applied. Furthermore, the maneuver routes are also varied depending on the reward coefficients. Specifically, ($W_A =$ 0.05, $W_F = 0.95$ case (the top-left figure in Fig. 10) denotes that combat individuals in the team are modeled to concentrate on the maintenance of a formation against the situated environments, such as elevation and terrain type (see Fig. 8), and it leads to their maneuver routes, being different (even longer than that from more W_A or No Maneuver MDP cases).

This result represents the effect of applying the proposed model compared to the descriptive-only model (i.e., No Maneuver MDP). As Table VII indicated, the MDP parts generate different performances through reward weight settings (i.e., better or even worse than no MDP case), which means human reasoning process in the military units significantly affects to the combat efficiency. We note that the maneuvering route changes in Fig. 10 are results from the reasoning process considering both the field manual and the environmental factors, so they could be utilized as a backup data for interpreting the different

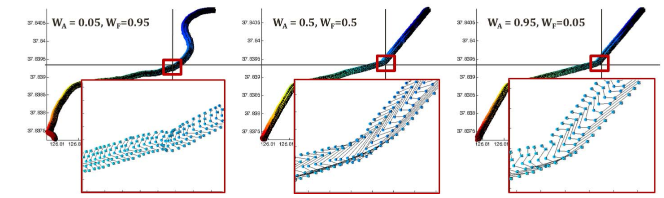


Fig. 11. Illustration of the formation maintenance of the third team of the second Platoon in Blueforce as the maneuver coefficients (W_A and W_F) changes.

TABLE IX
MEAN, STANDARD DEVIATION, AND STANDARD ERROR OF FVPS FROM THE MANEUVER LOGS OF THE THIRD TEAM OF THE SECOND PLATOON (IN FIG. 11)

	Formation Violation Penalty (FVP)		
(W_F, W_A)	(0.05, 0.95)	(0.5, 0.5)	(0.95, 0.05)
Mean	4.322	26.664	55.757
std. Dev	7.494	11.259	30.000
std. Error	2.370	3.560	9.487

performances and eventually used for improving the current military systems and designing the future military systems.

Fig. 11, on the other hand, illustrates the formation maintenance as the maneuver coefficients vary. The subfigures, in particular, emphasize keeping the team formation during simulation execution when the team passes the same point in the battlefield. In these figures, we can see that the formation is better maintained when W_F is larger, and the differences of the three cases are numerically compared by FVP (see Table IX). From these results figures, we argues that Maneuver MDP was operated as we expected, and, furthermore, the result analysis in the above section also become more convincing.

VI. DISCUSSION

This section deliberates about the benefits and the losses of the proposed method. The layered approach was suggested to complement the weak points of descriptive and prescriptive methods with the characteristics of each other. The descriptive-only model behaves according to the explicit action rule (e.g., military doctrines), meaning a minor change in the battlefield or in the objective would lead to the redescription of the whole agents' behavior. Meanwhile, prescriptive-only models show robust responses to changes in a combat situation, but their higher time complexity did not allow them to be applied to practical DM&S. In this sense, the proposed layered modeling adopts the descriptive features to address the practicality issues, and also adopts the prescriptive features to show robust behavior in response to changes occurring in the battlefield.

The provided case study illustrates other advantages of the proposed layered modeling: the proposed approach empirically sheds light on the need of human reasoning process in DM&S, which has been raised by military experts and researchers [58], [59], however, without any solid foundations. Having said that, the provided case study showed that the combination of descriptive and prescriptive modeling can statistically outperform descriptive-only model. These results imply that prescriptive modeling would contain key factors for improving combat efficiency, concurring with the need of the reasoning process.

The layered modeling also enables to explore how system behavior would be differed by changing individual reasoning process via employing additional parameters from the prescriptive parts. For example, the case study results (Fig. 10) show how team maneuvers were differed by the values of the weights in the maneuver MDP (W_F and W_A). Such illustration is certainly not obtainable from descriptive-only models, and it could provide an insight of enhancing military systems at the individual level.

Generalization is a key aspect in the modeling approach. Currently, the proposed layered model is implemented as a specialized case of integrating the descriptive DEVS and the prescriptive MDP models for specific purposes, such as maneuver and engagement. Having said that, we note that this paper presents that the concept of the proposed method (e.g., modularized components and hierarchical structure) can be rather considered as a theoretical and technical foundation. Nevertheless, we also agree to need more cases to prove the general value of applying the layered modeling to various domains. On the other hand, for this specific case, our layered modeling provided a generality from the scenario perspective. Often, the scenario imposes a certain behavior guideline that a simulator regenerates, but such guidelines are also often incomplete. Therefore, we need a generality in our model to handle such missing detailed behavior rules, and the prescriptive part provides a generalized answer under the rational behavior assumption.

Let us elaborate on the applicability of the proposed method. Modeling and simulation is applicable to man-made systems and nature systems. If one simulates a man-made system, the modeler could identify every behavior rules and implement them into simulation models. However, if a system is a natural system consisting of an individual and a collection of individuals, the behavior rule modeling becomes incomplete at a certain resolution level. From this aspect, the descriptive approach on modeling behavior rules is limited, and we require a general framework to generate such rule given outside stimuli. However, at the same time, if such a natural system is not so natural with some artificial components, i.e., social norms in the collective of individuals, we still can receive benefits from the rule description on such the artificial components.

Given this discussion, we would argue that the proposed layered model could be generally applicable to the domains of overlapping artificial rules and natural behaviors. One clear example of such overlapping domain is the defense community that models and simulates the artificial operations and doctrines as well as the human instincts on the combat field. This case can be found in diverse social simulation areas, e.g., political modeling with laws and ideologies, population modeling with welfare policy and individual family planning, etc. Also, the suggested interface enables to communicate DEVS and MDP modules, and the protocols used in the interface are built on their semantics. It means any DEVS and MDP modules can be interacted with the interface, and this generality, moreover, not only increases its applicability to other domain problems, but also suggests a general interface structure encompassing multiple methods with different modeling paradigms.

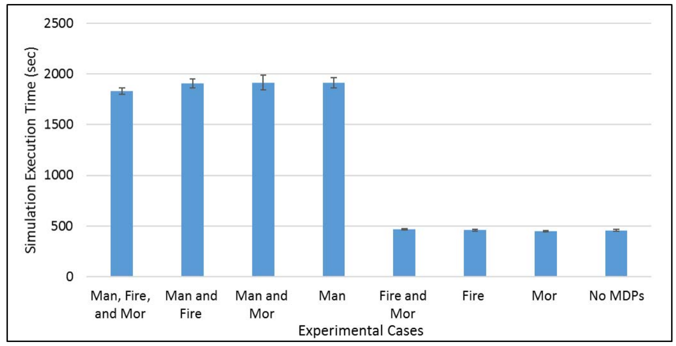


Fig. 12. Average of simulation execution time for experimental cases with 10 replications (Man: Maneuver MDP, Fire: Fire MDP, Mor: Mortar MDP; other parameters are fixed as W_F and W_A as 0.5, and fire damage as 60).

While the layered modeling is expected to provide the above benefits, it also holds negative sides. First, it requires much execution time by the addition of the reasoning process and interfacing two methods. Fig. 12 shows simulation execution time with respect to the MDP combinations in the case study (i.e., maneuver, fire, and mortar MDP). This figure represents that the simulation execution time of the proposed method takes about four times more than the descriptive-only model do (i.e., 1833 s from the maneuver, fire, and mortar case; 455.2 s from no MDPs case). We also note that, in the case study, the computation cost in the maneuver MDP is much dominant in the overall execution time, which means that the maneuver MDP could be improved or abstracted for any practical usages. Second, through the layered modeling, researchers would have more opportunities to investigate the effects of the reasoning process, but it also might cause the hardship on the interpretation and generalization of the observations. Hence, it is required that researchers consider this tradeoff relationship between modeling details and practical interpretations.

VII. Conclusion

Evaluating combat efficiency conducted by a myriad of humans is still a challenging problem, and DM&S has been successfully applied to resolve that problem. However, their practices still remained in the use of descriptive-only modeling, which is lacking of the considerations of human reasoning processes in military systems. Such descriptive-only modeling would overlook the influences from the unnoticed behaviors, so it could lead to missing opportunities of enhancing combat effectiveness or complementing unseen faults in the current military systems.

This paper proposes the layered behavior modeling of combat entities where the combat behaviors are modeled by considering both military doctrines and human reasoning processes. Specifically, the proposed behavior modeling consists of descriptive (i.e., DEVS) and prescriptive (i.e., MDP) parts, and their interactions are realized via the interface considering their semantics. We note that the proposed method enables to incorporate human reasoning process into military doctrines, which resolves the above concerns. The provided case study illustrates that the proposed method not only sheds light on the human reasoning process in the systems, but also enhances the overall combat effectiveness. Moreover, by exploring parameters of the prescriptive part, the proposed method would provide various behavioral patterns that may not be generated from the descriptive-only method.

This paper offers discussions on the proposed method from various perspectives. The generalization of the proposed method needs more evidences, yet we note that the structure of the proposed method (i.e., modularized components connected via interfaces based on their semantics) is designed to easily adapt to various domain problems with modeling approaches. In this sense, the provided case study shows how the proposed method is realized for a certain problem. It means the proposed method could be implemented by a pair of modeling methods, if there exists the similarity between them and extended to multilevel layers. In addition, considering its characteristics, this paper also notes that the proposed method is rather proper to the domains of overlapping artificial rules and natural behaviors. Meanwhile, its potential negatives are also presented, such as long execution time and difficulty on result analysis. Having said those features, we note that researchers who adapts the proposed method to their domain problems should consider such tradeoff relationships, and nevertheless expect that the proposed method would be used to create new insights for analyzing many complex systems including DM&S.

REFERENCES

[1] J. Poropudas and K. Virtanen, "Game-theoretic validation and analysis of air combat simulation models," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 40, no. 5, pp. 1057–1070, Sep. 2010.

[2] J. W. Bae, J. H. Kim, I.-C. Moon, and T. G. Kim, "Accelerated simulation of hierarchical military operations with tabulation technique," *J. Simulat.*, vol. 10, no. 1, pp. 36–49, 2016.

[3] K. Schwenn, J. Colombi, T. Wu, K. Oyama, and A. Johnson, "Toward agent-based modeling of the U.S. department of defense acquisition system," *Procedia Comput. Sci.*, vol. 44, pp. 383–392, Dec. 2015.

[4] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical, Multilevel, Systems*, vol. 68. New York, NY, USA: Academic Press, 2000.

[5] A. Tolk, *Engineering Principles of Combat Modeling and Distributed Simulation*. Hoboken, NJ, USA: Wiley, 2012.

[6] R. Mitchell and I.-R. Chen, "Adaptive intrusion detection of malicious unmanned air vehicles using behavior rule specifications," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 5, pp. 593–604, May 2014.

[7] T. G. Kim *et al.*, "DEVSim++ toolset for defense modeling and simulation and interoperation," *J. Defense Model. Simulat. Appl. Methodol. Technol.*, vol. 8 no. 3, pp. 129–142, 2010.

[8] W.-S. Yun, I.-C. Moon, and T.-E. Lee, "Agent-based simulation of time to decide: Military commands and time delays," *J. Artif. Soc. Soc. Simulat.*, vol. 18, no. 4, pp. 1–10, 2015. [Online]. Available: http://jasss.soc.surrey.ac.uk/18/4/10.html

[9] B. G. Silverman, M. Johns, J. Cornwell, and K. O'Brien, "Human behavior models for agents in simulators and games: Part I: Enabling science with PMFserv," *Presence Teleoper. Virtual Environ.*, vol. 15, no. 2, pp. 139–162, 2006.

[10] S. R. Goerger, "Validating human behavioral models for combat simulations using techniques for the evaluation of human performance," DTIC, Fort Belvoir, VA, USA, Rep., 2004. [Online]. Available: http://www.dtic.mil/docs/citations/ADA495197

[11] E. Hosseini and R. Berangi, "Optimal SMDP-based connection admission control mechanism in cognitive radio sensor networks," *ETRI J.*, vol. 39, no. 3, pp. 345–352, 2017, doi: 10.4218/etrij.17.0116.0520.

[12] B. Kim and S.-S. Ko, "Optimal admission control and state space reduction in two-class preemptive loss systems," *ETRI J.*, vol. 37, no. 5, pp. 917–921, 2015, doi: 10.4218/etrij.15.0114.0348.

[13] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, vol. 1. Cambridge, MA, USA: MIT Press, 1998.

[14] C. Liu, X. Xu, and D. Hu, "Multiobjective reinforcement learning: A comprehensive overview," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 3, pp. 385–398, Mar. 2015.

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

[16] J.-C. Castella, S. P. Kam, D. D. Quang, P. H. Verburg, and C. T. 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.

[17] 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*. London, U.K.: Springer, 2013, pp. 35–66.

[18] W. Yu *et al.*, "An UHV grid security and stability defense system: Considering the risk of power system communication," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 491–500, Jan. 2016.

[19] P. Gorniak and I. Davis, "SquadSmart: Hierarchical planning and coordinated plan execution for squads of characters," in *Proc. AIIDE*, 2007, pp. 14–19.

[20] J. Stavash *et al.*, "Agent based models for logistics in wargaming," in *Proc. Fall SISO Simulat. Interoperability Workshop*, Orlando, FL, USA, 2003. [Online]. Available: https://www.sisostds.org/DigitalLibrary.aspx?EntryId=24708

[21] F. Moradi and J. Schubert, "Simulation-based defense planning," in *Proc. NATO Symp. Integr. Model. Simulat. Defence Acquisition Lifecycle Military Train. Curriculum*, 2014, pp. 1–16.

[22] S. Bisht, A. Malhotra, and S. B. Taneja, "Modelling and simulation of tactical team behaviour," *Defence Sci. J.*, vol. 57, no. 6, pp. 853–864, 2007.

[23] S. C. Jethi and R. K. Jain, "Simulation of naval wargames," *Defence Sci. J.*, vol. 54, no. 3, pp. 407–415, 2004.

[24] U. Dompke, "Computer generated forces-background, definition and basic technologies," DTIC, Fort Belvoir, VA, USA, Rep., 2003. [Online]. Available: http://www.dtic.mil/docs/citations/ADA485518

[25] I. Cil and M. Mala, "A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1331–1343, 2010.

[26] C. K. Pang *et al.*, "Discrete event command and control of multiple military missions in network centric warfare," in *Proc. 8th Int. Conf. Intell. Unmanned Syst.*, 2012, pp. 74–79.

[27] J. W. Bae and I.-C. Moon, "LDEF formalism for agent-based model development," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 6, pp. 793–808, Jun. 2016.

[28] J. Kim, C. Choi, I.-C. Moon, and T. Kim, "DEVS-based doctrine validation of fleet anti-air defense," in *Proc. Spring Simulat. Multiconf.*, Orlando, FL, USA, 2010, p. 28.

[29] E. Salas, C. A. Bowers, and J. A. Cannon-Bowers, "Military team research: 10 years of progress," *Military Psychol.*, vol. 7, no. 2, pp. 55–75, 1995.

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

[31] K. Carley, J. Kjaer-Hansen, A. Newell, and M. Prietula, "Plural-soar: A prolegomenon to artificial agents and organizational behavior," in *Artificial Intelligence in Organization and Management Theory: Models of Distributed Activity*. Amsterdam, The Netherlands: Elsevier, 1992, pp. 87–118.

[32] B. P. Zeigler, H. Praehofer, and T. G. Kim, *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. San Diego, CA, USA: Academic Press, 2000.

[33] B. P. Zeigler, *Multifacetted Modelling and Discrete Event Simulation*. London, U.K.: Academic Press, 1984.

[34] K. H. Lee, J. H. Hong, and T. G. Kim, "System of systems approach to formal modeling of CPS for simulation-based analysis," *ETRI J.*, vol. 37, no. 1, pp. 175–185, 2015, doi: 10.4218/etrij.15.0114.0863.

[35] N. R. Council *et al.*, *Modeling Human and Organizational Behavior: Application to Military Simulations*. Washington, DC, USA: Nat. Acad. Press, 1998.

[36] R. E. Wray and J. E. Laird, "Variability in human behavior modeling for military simulations," in *Proc. Behav. Represent. Model. Simulat. Conf. (BRIMS)*, Scottsdale, AZ, USA, 2003, pp. 176–185.

[37] D. Davis, D. Brutzman, C. Blais, and R. McGhee, "Ethical mission definition and execution for maritime robotic vehicles: A practical approach," in *Proc. IEEE OCEANS MTS/IEEE Monterey*, Monterey, CA, USA, 2016, pp. 1–10.

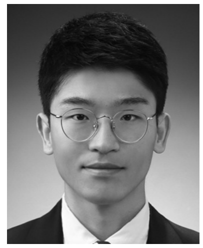
[38] T. Wenqian, L. Zongyuan, Q. Xiangju, W. Yanyang, and S. Liguo, "Aircraft-pilot system modeling and pilot control behavior research for airdrop task," *IFAC-PapersOnLine*, vol. 49, no. 32, pp. 177–182, 2016.
[39] R. H. Kewley and M. J. Embrechts, "Computational military tactical planning system," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 32, no. 2, pp. 161–171, May 2002.
[40] D. Aberdeen, S. Thiébaux, and L. Zhang, "Decision-theoretic military operations planning," in *Proc. ICAPS*, 2004, pp. 402–412.
[41] C. Gonzalez, N. Ben-Asher, and D. Morrison, "Dynamics of decision making in cyber defense: Using multi-agent cognitive modeling to understand cyberwar," in *Theory and Models for Cyber Situation Awareness*. Cham, Switzerland: Springer, 2017, pp. 113–127.
[42] S. Bisht, H. S. Bharati, S. B. Taneja, and P. Bedi, "Command agent belief architecture to support commander decision making in military simulation," *Defence Sci. J.*, vol. 68, no. 1, pp. 46–53, 2018.
[43] M. Mishra *et al.*, "Context-aware decision support for anti-submarine warfare mission planning within a dynamic environment," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.
[44] R. Bellman, "A Markovian decision process," *Indiana Univ. Math. J.*, vol. 6, no. 5, pp. 679–684, 1957.
[45] X. Chen, "Police patrol optimization with security level functions," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 5, pp. 1042–1051, Sep. 2013.
[46] S. Jia, X. Wang, and L. Shen, "A continuous-time Markov decision process-based method with application in a pursuit-evasion example," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 9, pp. 1215–1225, Sep. 2016.
[47] C. Heinze *et al.*, "Interchanging agents and humans in military simulation," *AI Mag.*, vol. 23, no. 2, p. 37, 2002.
[48] T. M. Cioppa, T. W. Lucas, and S. M. Sanchez, "Military applications of agent-based simulations," in *Proc. 36th Conf. Winter Simulat. (WSC)*, Washington, DC, USA, 2004, pp. 171–180.
[49] I. Cil and M. Mala, "A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1331–1343, 2010.
[50] M. F. Hocaoglu, "AdSiF: Agent driven simulation framework," in *Proc. Hunstv. Simulat. Conf. HSC*, 2005, pp. 1–12.
[51] A. Collins, J. Sokolowski, and C. Banks, "Applying reinforcement learning to an insurgency agent-based simulation," *J. Defense Model. Simulat.*, vol. 11, no. 4, pp. 353–364, 2014.
[52] S. Elkosantini, "Toward a new generic behavior model for human centered system simulation," *Simulat. Model. Pract. Theory*, vol. 52, pp. 108–122, Mar. 2015.
[53] Y. Fu, X. Yu, and Y. Zhang, "Sense and collision avoidance of unmanned aerial vehicles using Markov decision process and flatness approach," in *Proc. IEEE Int. Conf. Inf. Autom.*, Lijiang, China, 2015, pp. 714–719.
[54] S. Chen, F. Wu, L. Shen, J. Chen, and S. D. Ramchurn, "Decentralized patrolling under constraints in dynamic environments," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3364–3376, Dec. 2016.
[55] S. Sanner, *Relational Dynamic Influence Diagram Language (RDDL): Language Description*, Aust. Nat. Univ., Canberra, ACT, Australia, 2010.
[56] M. Ghallab *et al.*, "PDDL—The planning domain definition language," Yale Center Comput. Vis. Control, New Haven, CT, USA, Rep. CVC TR-98-003/DCS TR-1165, 1998.
[57] A. Weinstein and M. L. Littman, "Bandit-based planning and learning in continuous-action Markov decision processes," in *Proc. ICAPS*, Atibaia, Brazil, 2012, pp. 306–314.
[58] C. Heinze, B. Smith, and M. Cross, "Thinking quickly: Agents for modeling air warfare," in *Advanced Topics in Artificial Intelligence*. Berlin, Germany: Springer, 1998, pp. 47–58.
[59] G. Tidhar *et al.*, "Using intelligent agents in military simulation or 'using agents intelligently,"' in *Proc. AAAI/IAAI*, Orlando, FL, USA, 1999, pp. 829–836.

Jang Won Bae received the M.S. degree in electrical engineering and the Ph.D. degree in industrial and systems engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2007 and 2015, respectively.

He is a Senior Researcher with the Electronics and Telecommunications Research Institute, Daejeon. His current research interests include agent-based modeling and simulation, simulation-based optimization and evaluation, interconnecting simulation, and AI techniques.

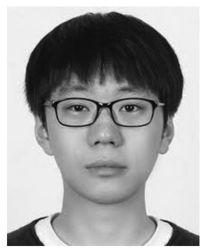
Junseok Lee received the M.S. degree in industrial and systems engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2015.

He is an Engineer with Samsung Electronics, Suwon, South Korea. His current research interests include agent-based modeling and simulation with AI. He is currently developing mobile browser services and a simulation testbed for browsing performance.



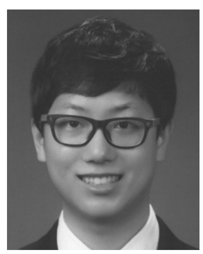
Do-Hyung Kim received the M.S. degree in industrial and systems engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2018.

He is currently a Data Analytics CoE with ICT Research and Development Center, SK Telecom, Seoul, South Korea.



Kanghoon Lee received the dual B.S. degrees in computer science and mathematics from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2008, where he is currently pursuing the Ph.D. degree in computer science.

In 2016, he also joined T-Brain, SK Telecom, Seoul, South Korea, as a Research Scientist. His current research interests include theory and applications of reinforcement learning, Markov decision process, Bayesian learning, and deep learning.



Jongmin Lee received the B.S. degree in computer science and engineering from Seoul National University, Seoul, South Korea, in 2014, and the M.S. degree in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2017, where he is currently pursuing the Ph.D. degree in computer science.

His current research interests include machine learning, sequential decision making problems in artificial intelligence, and reinforcement learning.

Kee-Eung Kim (M'10) received the B.S. degree in computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 1995, and the M.Sc. and Ph.D. degrees in computer science from Brown University, Providence, RI, USA, in 1998 and 2001, respectively.

From 2001 to 2006, he was a Senior Software Engineer with Samsung SDS, Seoul, South Korea, and a Senior Research Staff Member with Samsung Advanced Institute of Technology, Suwon, South Korea. In 2006, he joined the faculty of Computer Science Department, KAIST, where he is currently an Associate Professor. His current research interests include representations and algorithms for sequential decision making problems in artificial intelligence and machine learning, Markov decision processes, and reinforcement learning.

Il-Chul Moon received the Ph.D. degree from the School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, in 2008.

He is an Assistant Professor with the Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. His current research interests include computer science, management, sociology, operations research, military command and control analysis, counterterrorism analysis, intelligence analysis, and disaster management.