

## POMDP-DEVS를 활용한 전투 개체 모델링

배장원<sup>1</sup> · 이강훈<sup>2</sup> · 김현은<sup>2</sup> · 이준석<sup>1</sup> · 고봉석<sup>2</sup> · 남보원<sup>2</sup> · 문일철<sup>1\*</sup> · 김기웅<sup>2</sup> · 박재현<sup>3</sup>

<sup>1</sup>KAIST 산업 및 시스템 공학과 / <sup>2</sup>KAIST 전산학과 / <sup>3</sup>국방과학연구소

## Modeling Combat Entity with POMDP and DEVS

Jang Won Bae<sup>1</sup> · Kanghoon Lee<sup>2</sup> · Hyeoneun Kim<sup>2</sup> · Jun Seok Lee<sup>1</sup> · Bongseok Goh<sup>2</sup> · Bowon Nam<sup>2</sup>  
Il-Chul Moon<sup>1</sup> · Kee-Eung Kim<sup>2</sup> · Jaehyun Park<sup>3</sup>

<sup>1</sup>Department of Industrial and Systems Engineering, KAIST

<sup>2</sup>Department of Computer Science, KAIST

<sup>3</sup>Agency for Defense Development

Combat Modeling and Simulation (M&S) is significant to decision makers who predict the next direction of wars. Classical methodologies for combat M&S aimed to describe the exact behaviors of combat entities from military doctrines, yet they had a limitation of describing reasonable behaviors of combat entities that did not appear in the doctrines. Hence, this paper proposed a synthesizing modeling methodology for combat entity models considering both 1) the exact behaviors using descriptive modeling and 2) the reasonable behaviors using prescriptive modeling. With the proposed methodology, combat entities can represent a reality for combat actions rather than the classical methodologies. Moreover, the experiment results using the proposed methodology were significantly different from the results using the classical methodologies. Through the analyses of the experiment results, we showed that the reasonable behaviors of combat entities, which are not specified in the doctrines, should be considered in combat M&S.

**Keywords:** Combat Entity Modeling, Descriptive Modeling, Prescriptive Modeling, DEVS, POMDP

### 1. 서론

병력을 통솔하여 전투로 이끄는 지휘관이라면 누구나 향후 발생할 전투의 결과를 예측하고, 그 예측 결과가 최대한 유리하게 도출될 수 있도록 훈련과 전술을 준비해야 한다. 최근에는 국방 M&S의 연구 결과물인 전투모의(simulation) 및 실험(experiment)이 지휘관들의 과업 달성에 도구가 되었다(Kim *et al.*, 2011; Moon *et al.*, 2013). 그러나 전투 모의가 얼마나 자세히 이루어져야 하고, 어떤 방식으로 이루어져야 하는지는 의문의 여지가 있다. 전투 모의는 기본적으로 전투 모의의 목적에 맞는 수준의 추상화를 통해, 전투 모의의 형태와 포함 요소가 개념적으로 결정된다. 여기서 추상화의 대상은 전투에 참여하는

대상 인원, 장비, 전장 환경 등이 포함되는데, 특히 전투 참여 인원은 실질적으로 인간 행동을 추상화하여 모의하게 된다. 그렇다면 인간의 행동을 얼마나 추상화해야 전투 모의의 목적을 잘 달성할 수 있을 것인지 결정해야 한다.

전투 모의를 군에서의 최소 규모인 분대급 전투로 한정할 경우에도 전투 참여 인원(이하, 전투 개체)의 행동 묘사는 인간 행동의 복잡성 때문에 여전히 단순한 문제가 아니다. 전투 개체 모델링의 기초가 되는 분대급 전투의 교범(field manual)은 소대급 전투나 중대급 전투에 비해 매우 단순하다. 그러나 분대급 전투 개체들이 실제 전장에서 보이는 다양한 전투 행동들은 전투 행동 교범에 모두 명시되어 있지도 않고, 명시하기도 어려운 행동들이 많다. 예를 들어, 분대급 전투 시 장애물

본 연구는 방위사업청과 국방과학연구소의 지원으로 수행되었습니다(UD110006MD).

\* 연락저자 : 문일철 교수, 305-701 대전광역시 유성구 대학로 291 한국과학기술원(KAIST) 산업 및 시스템 공학과, Tel : 042-350-3118,

Fax : 042-350-3110, E-mail : icmoon@kaist.ac.kr

2013년 7월 18일 접수; 2013년 9월 13일 수정본 접수; 2013년 11월 27일 게재 확정.

극복에 대해 교범은 “신속히 극복한다”와 같은 추상적 단어로 표현되어 있는 반면, 실질적인 전투 행동은 주변 지형을 관찰하여 인간의 직관으로 장애물을 신속히 극복하는 행동을 보이기 때문이다. 이는 전투 행동 교범만을 전투 모의의 기본 자료로 사용한다면, 실제 전투 개체가 보여야 할 몇몇 행동은 모의가 어려울 수도 있다는 점을 보여준다.

만약 전투 개체가 전투 모의에서 보이는 행동이 전투 교범에 묘사된 행동 이상을 보여야 한다면, 이 추가적인 차원의 행동을 어떻게 모의해야 하며, 이런 추가 모의가 반드시 필요한 것인가를 판단해야 한다. 과거 전투모의 결과물은 대부분 명시적으로 서술된 기술적 모델링(descriptive modeling)이 대부분이었다. 이는 전투 교범에 나타난 전투 개체의 행동 양식을 명시적으로 모델에 기술하여, 전투 모의를 하는 방식을 의미한다(Kim *et al.*, 2012). 여기서 전투 교범에 묘사되지 않은 행동들은 단순한 인간 행동 양식 가정하고 기술하여 모의를 수행하여 왔다. 그러나 인간의 직관 및 행동 양식에 더욱 근접한 행동을 보이기 위해서는, 확률론적으로 추론된 규정적 모델링(prescriptive modeling)이 필요하다. 규정적 모델링은 인간의 이성적 행동 양식을 모델링 하는데 전통적으로 사용되어 왔으며, 명시적으로 기술되기 어려운 인간의 직관을 표현하는데 유용한 방식이다(Lee *et al.*, 2012). 즉, 전투 개체의 행동이 규범에 명시된 행동만이 아니라, 규범에 명시되지 않은 직관적 행동까지 포함해야 한다면, 정확한 모델링을 위해 기술적 모델링과 규정적 모델링을 함께 사용해야 할 필요성이 있다. 만약, 기술적 모델링만을 사용한 전투모의 예측 결과와 기술적 모델링 및 규정적 모델링을 함께 사용한 전투모의 예측 결과가 심각히 다르다면, 지휘관은 어느 모델링이 실제에 더욱 가까운 모델링인지 판단하여, 실제에 가까운 전투 행동을 모의하는 모델링 방법론을 선택하여 활용해야 한다.

본 논문은 최소 부대 단위인 분대급 전투 모의에서 기술적 모델링과 규정적 모델링을 복합하여 전투 모의를 하는 방법 및 사례를 제시하며, 전투 실험을 통하여 기존의 기술적 모델링만을 활용한 예측 결과와 제안된 모델링의 예측 결과가 매우 큰 차이를 보인다는 점을 제시한다. 마지막으로 기술적 및 규정적 모델링을 복합하여 활용하는 경우, 기존의 기술적 모델링만을 활용하는 경우보다 인간의 일반적 행동을 더욱 간단하고도 자세하게 표현할 수 있음을 보여준다. 본 논문이 제안하는 모델링 방법은 기술적 모델링을 이산 사건 모델(discrete event model)에 활용하여 전투 교범을 상위에서 모의하며, 규정적 모델링을 행동 모델(behavior model)에 활용하여 인간의 이성적 전투 행동을 하위에서 모의하는 방식이다. 이와 같은 모델링 양식은 교범에 따라 장병을 지휘하지만, 교범이 허용하는 범위 내에서 인간의 이성적 행동을 인정하는 군의 행동 방식과 궤를 같이 한다.

위 내용을 보이기 위해, 본 논문은 제 2장을 통하여 전통적인 전투 개체 모델링 방법론과 행동 묘사 모델링 방법론의 관련 연구를 소개하며, 기존에 제안된 통합 모델링 방법론을 소

개한다. 제 3장을 통하여, 본 논문이 사례로 제시하는 전투모의 시나리오를 서술한다. 제 4장은 전투 개체 모델 개발의 방법론을 구체적으로 설명한다. 제 5장은 기존 방법과 제안된 방법을 활용하여 전투 실험을 수행하고, 실험 결과의 상이함과 제안된 방법이 실제 전투에 더욱 근접하다는 점을 보여준다.

## 2. 관련 연구

### 2.1 전통적인 전투 개체 모델링 방법론

모델링 및 시뮬레이션(modeling and simulation; M&S)은 반복적으로 재연하기 불가능하고 예측하기 어려운 상황을 다루는 기술로 국방 분야의 무기 체계와 작전 등의 설계, 분석 및 평가를 위한 중요한 기술로 사용되고 있다(Kim *et al.*, 2011; Moon *et al.*, 2013). 위게임(war game)은 국방 M&S의 가장 대표적인 예로, 수많은 전투 개체 모델(combat-entity model)과 다양한 전장 정보를 이용해 실제 전쟁을 모의한다. 일반적으로 전투 개체들은 군 사령부에서 공식적으로 승인된 군사행동의 기본 원칙과 지침을 설명하고 있는 군사 교리(military doctrine)(NSA 2010)라고 하는 부대 운용의 기본 규칙을 따른다. 그렇기 때문에 규칙 기반 모델링(rule-based model)은 전투 개체 모델을 개발하는 가장 직관적이고 직접적인 방법으로 사용되었다(Cil *et al.*, 2010; Pang *et al.*, 2012). 규칙 기반 모델은 모델 내부의 일정한 조건이나 상태에 따라서 다음 행동이나 모델의 상태를 결정한다. 규칙 기반 모델의 상태 결정 과정은 인공 지능 관점(artificial intelligence)과 시스템 공학(system engineering)적인 관점으로 구체화할 수 있다. 인공 지능 관점을 적용한 예로는 Soar(state, operator, and result) 프레임워크가 있다(Laird *et al.*, 1987). Soar 프레임워크는 산출 규칙(production rule) 기반 인공 지능 프레임워크로서, 모델의 상태(state)를 산출 규칙에 적용하여 자율적인 행동을 결정하는 방식을 사용한다. 이러한 방식을 이용해 모델의 행동들 사이의 우선순위나 새로운 규칙을 기억 공간에 저장하며 학습 한다. 또한, 기억 공간을 산출 규칙을 기술하는 장기 기억 공간(long-term memory)과 현재 상태를 기술하는 작업 기억 공간(working memory)으로 분리하여 효율적으로 활용한다. 시스템 공학 관점을 적용한 예로는 DEVS(discrete event system specification) 형식론이 있다. DEVS 형식론은 이산 사건 시스템 모델링을 위한 수학적 틀로서, 집합과 함수를 이용한 모델링을 지원 하고 정형적 모델링(formal modeling)을 통한 모델 검증(model verification)이 용이하다. 또한, DEVS 형식론은 행위 모델을 나타내는 원자(atomic) 모델과 모델의 연결 관계를 나타내는 결합(coupled) 모델을 이용하여, 복잡한 모델을 계층적이고 모듈러하게 개발할 수 있다는 특징이 있다(Ziegler *et al.*, 2000). Soar와 DEVS 같은 규칙 기반 모델링은 대상 시스템을 충실하게 묘사하는 것에 초점을 두고 있는데, 이러한 모델링 방법을 기술적 모델링(descriptive modeling)이라

고 한다. 기술적 모델링은 대상 시스템의 동적 특성을 조건·행동 규칙과 이를 통한 모델 상태 전이를 이용해 대상 시스템을 묘사한다. 최근까지의 국방 M&S는 전투 개체 모델이 전투 교리를 잘 표현하는 것이 중요한 부분이었기 때문에, 기술적 모델링이 자주 사용되어 왔다(Kim *et al.*, 2011; Kim *et al.*, 2010).

그러나 미래의 다양한 무기 체계와 다변하는 전장 환경에 효율적으로 대처하기 위해서는 기술적 모델링을 활용하는 방법으로는 한계가 있다. 왜냐하면 전투 개체가 규칙적으로 움직이기 보다는, 상황에 따른 전투 개체의 행동 결정이 필요하기 때문이다. 기술적 모델링의 바탕이 되는 전투 교리는 전투의 기본원칙과 지침을 설명하고 있지만, 다양한 전장상황에 따른 자세한 행동에 대해서는 기술하지 않고 있다. 미육군성본부(Headquarters, department of the army)에서 제정한 야전 교범(field manual)에도 전투 교리의 내용을 전장의 특별 상황에 대한 적절한 변형(recommended changes) 등이 꼭 필요하다고 설명하고 있다(Headquarters, 1962). 전투 교리의 적절한 변형이 필요한 경우의 예로써 부대 기동을 들 수 있다. 야전 교범은 부대의 다양한 기동 유형의 종류와 특징 등을 설명하고 있지만 다변하는 전장 상황을 모두 고려한 기동 유형을 결정하여 주지는 못하기 때문에, 전투 개체가 상황을 판단하여 교리에 맞는 기동 행동을 하는 것이 필요하다. 또한, 새로운 무기 체계에 대한 운용 교리나 부대 작전을 개발하는 경우에는 전투 개체의 이성적인 행동 선택이 기초 자료가 될 것이다. 즉, 전투 교리의 내용을 반영하는 기술적 모델링으로는 효율적인 전투 개체를 개발하거나 교리가 개발의 기초 자료가 될 수 있는 창발적 행위(emergent behavior)를 기대하기 어렵다. 그러나 이성적인 창발적 행위는 기존 교리의 효율을 높이고 새로운 교리의 개발에 도움을 줄 수 있기 때문에, 미래의 전장에 강인한 전투 개체 개발에 도움을 줄 수 있다.

## 2.2 행동 묘사 모델링 방법론

전투 개체 모델링에 있어서 기술적 모델링은 전투 교리를 충실히 반영할 수 있는 반면에 전투 교리에 기술되지 않은 구체적 행동들에 대한 묘사에는 어려움이 있다. 이러한 어려움은 주로 전장 상황 전개의 불확실성 및 전투 개체들의 전장 상황 인식 제한성으로부터 기인하게 되고, 따라서 전투 개체의 보다 현실적인 묘사를 위해서는 위의 불확실성 및 제한성이 반영된 행동 묘사를 위한 규정적 모델링(prescriptive modeling)이 필요하게 되었다.

부분 관찰 마코프 의사 결정 과정(partially observable Markov decision process; POMDP)(Sondik, 1971)은 의사 결정 이론에 입각한 대표적인 규정적 모델링 기법이다. POMDP는 불확실한 혹은 부분적으로 관찰 가능한 상황을 확률적으로 모델링하여 개체 스스로가 순차적 의사 결정(sequential decision making)을 하도록 하는 방법론으로서 인지과학, 경제학 및 인공지능 분야에서 중요한 모델링 기법이다. 전문가의 지식에 전적으로

의존하여 전투 개체의 행동을 기술하는 규칙 기반 시스템과 달리 POMDP에서는 시간의 흐름에 따라 변화하는 전장 상황을 확률적으로 모델링하고, 동적 프로그래밍(dynamic programming) 및 탐색(search) 알고리즘으로 최적의 행동을 계산함으로써 불확실한 전장 상황에서의 이성적 행동(rational behavior) 정책을 계산할 수 있다는 특징을 지닌다. 즉 전투 개체의 행동 묘사를 위해 모든 상황마다 행동 규칙을 정의하는 것이 아니라 전장 환경의 동적 특성(dynamics)을 확률적 전이 함수로, 전투 개체의 목표를 보상 함수(reward function)로 나타낸다. POMDP 알고리즘은 전투 개체가 행동함으로서 받게 되는 총 보상 값의 기대치를 최대화 할 수 있는 순차적 행동들을 계산함으로써 이성적 행동을 도출하게 된다. POMDP를 이용한 복수 무인기에서의 동적 임무 할당 및 정찰 임무 최적화(Kim *et al.*, 2012) 및 대규모 전투 개체 표현을 위한 POMDP의 확장 모델인 Factored POMDP를 이용한 가상군 자율행위 모델링(Lee *et al.*, 2012) 등의 기존 연구는 POMDP를 활용한 전투 개체 행동 모의의 적용 가능성을 보여준다.

한편, POMDP에서 정확한 최적 행동 정책을 구하는 것은 계산학적으로 힘들다(computationally intractable)고 알려져 있다. 특히 POMDP 개체가 수행 할 수 있는 시간(time horizon)이 유한일 경우에는 POMDP 문제의 복잡도가 PSPACE-complete (Papadimitriou and Tsitsiklis, 1987)로 알려져 있다. 최근 POMDP 연구에서는 point-based value iteration(Pineau *et al.*, 2003) 및 nominal belief-state optimization(Miller *et al.*, 2009) 등의 다항 시간(polynomial time) 근사 알고리즘의 개발을 통해 높은 계산 복잡도 문제를 다루고자 했다. 이러한 연구 성과들에도 불구하고 아직도 POMDP를 이용한 국방 모의에는 한계가 존재한다. 대규모 전투 개체의 행동 모의를 위해 Factored POMDP를 이용한 기존 연구(Lee *et al.*, 2012)가 있었지만, 전투 개체의 구체적 행동 모의가 아닌 전술 행동 모의를 함으로서 문제의 복잡도를 많이 줄이게 되었다. 이처럼 소대 단위 이상의 대규모 전투 개체의 행동 모의는 POMDP 만을 이용해서는 현실적으로 쉽지 않다. 또한 POMDP는 보상 함수로서 전투 개체의 목표를 규정하고 행동에 대한 제약을 두지 않기 때문에, 군사 교리와 같은 강제적인 규칙을 표현하기에 어려움이 있다.

## 2.3 이산사건과 행동 모델의 분리 및 병합 방법론

기술적 모델링(descriptive modeling)은 대상 시스템의 모델을 효율적으로 개발할 수 있지만, 모델에 표현되지 않은 상황에 대한 이성적 행동을 모의하는 것에는 한계점을 지닌다. 반면에, 규정적 모델링(prescriptive modeling)은 의사 결정 이론에 입각하여 환경의 불확실성을 고려한 최적의 행동을 선택할 수 있다는 장점이 있지만, 높은 문제 복잡도로 인하여 대규모 시스템에 적용하기 힘들다는 단점이 있다. 이와 같이, 기술적 모델링과 규정적 모델링의 장단점은 서로 상반되는 경향이 있다. 이러한 특징으로 인해, 두 가지 모델링 방법은 서로 독립적으로

그리고 다른 목적을 위해서 국방 M&S에 적용되었다. 기술적 모델링은 전투 개체 모델에 전투 교리를 자세히 표현하여 기존의 전투 교리를 평가하고 분석하는 곳에 적용되었다(Salas et al., 1995; Kim et al., 2011). 규정적 모델링은 전투 개체의 행동 선택 과정을 표현하여 전장의 불확실성과 의사 결정 이론을 고려하여 새로운 작전을 개발하는데 적용되었다(Schmitt et al., 1999; Bryant et al., 2004).

두 모델링 방법의 장점을 융합할 수 있는 연구로, 최근 협동 모델링(collaborative modeling) 방법론이 제안되었다(Sung et al., 2012). 협동 모델링 방법론은 도메인 지식이 반영된 모델을 개발할 때, 대상 모델을 시스템 공학적 특성이 반영된 이산 사건 시스템 모델(discrete event model; DEM)과 도메인 지식이 반영된 행위 모델(behavioral model; BM)로 분리하여 개발하는 방법이다. 독립적으로 개발된 DEM과 BM은 미리 정의한 인터페이스로 유연하게 합성하고, 궁극적으로 도메인 지식(BM)이 반영된 동시에 시스템 공학적인 방법(DEM)으로 모델을 개발할 수 있다. 이 방법론은 하나의 모델링 방법으로 표현하기 어려운 시스템을 DEM과 BM으로 분리하고 각각을 적합한 모델링 방법을 적용하여 독립적으로 개발함으로써, 효율적으로 대상 시스템의 모델을 개발할 수 있게 해준다는 장점이 있다.

본 연구에서는, 협동 모델링 방법론에 기반하여 전투 개체 모델을 기술적 모델링과 규정적 모델링을 이용한 모델 융합(model integration)으로 개발한다. <Figure 1>은 기술적 모델링과 규정적 모델링을 융합하여 전투 개체 모델을 개발하는 과정을 개념적으로 보여준다. 전투 개체 모델은 역할을 기준으로 기동, 탐지, 교전 및 의사 결정으로 분리하여 개발한다. 또한, 전투 개체의 세부 모델들은 이산 사건 모델(DEM)과 행위 모델(BM)로 분리하여 개발한다. 이산 사건 모델은 야전 교범과 같은 교리를 반영하며, 기술적 모델링을 이용하여 개발하고, 행위 모델은 교리에 표현되지 않은 행동들에 대해서 이성적으로 결정하는 과정을 표현하며, 규정적 모델링을 이용하여 개발한다. 독립적인 방법으로 개발된 이산 사건 모델과 행위 모델은

모델 인터페이스(model interface)를 이용하여 연결한다. 본 연구에서 모델 인터페이스는 DEVS 형식론의 연결 관계(coupling relations)를 이용하여 개발하였다. 이러한 방법을 통하여, 이산 사건 모델은 행위 모델을 이용하여 모델의 이성적 행동을 결정할 수 있고, 행위 모델은 이산 사건 모델을 이용하여 대규모 모델에도 적용할 수 있게 되어, 기술적 모델링과 규정적 모델링의 단점을 보완한 전투 개체 모델의 개발을 가능하게 한다. 이렇게 개발된 전투 개체 모델은 군사 교리를 고려하는 동시에 전투 개체의 기동, 탐지, 교전 등에 관련된 최적의 행동을 결정할 수 있는 능동적이고 전장 환경에 강인한 모델로 개발될 수 있다.

### 3. 전투모의 시나리오

자세한 전투 개체 모델링 내용을 소개하기에 앞서, 모델을 이용하여 전투 실험을 수행할 전투 모의 시나리오를 설명한다. 전투 실험에서 사용할 시나리오는 청군 1개 분대가 경기도 파주시의 봉서산 홍군 경계 진지를 점령하는 임무를 다루고 있다. 목표 지역은 홍군 반개 분대 규모의 경계 팀이 주둔하고 있고, 청군 분대는 두 개의 팀으로 분리되어 군사 교리를 반영한 진지 점령 작전을 수행한다.

<Figure 2>는 봉서산 일대의 분대 수준의 교전 시나리오를 설명한다. 시나리오 초기에는 <Figure 2>와 같이 홍군 경계 진지(security position)에 홍군 반개 규모의 경계 팀이 주둔하고 있고, 청군 분대는 홍군 경계 진지 오른쪽 150미터 지점에 주둔하고 있다. 청군 분대는 분대 1팀(blue team 1)과 분대 2팀(blue team 2)으로 나뉘어서, 홍군 경계 진지 점령 작전을 수행한다. 청군 분대 2팀은 홍군 경계 진지를 우회하여 차단진지(blocking position)로 진군하여 대기하고, 청군 분대1팀은 홍군 경계진지 정면의 급편진지(hastily prepared position)로 기동한다. 특히, 청군 분대 2팀의 차단진지 확보는 적의 기습 사격, 포

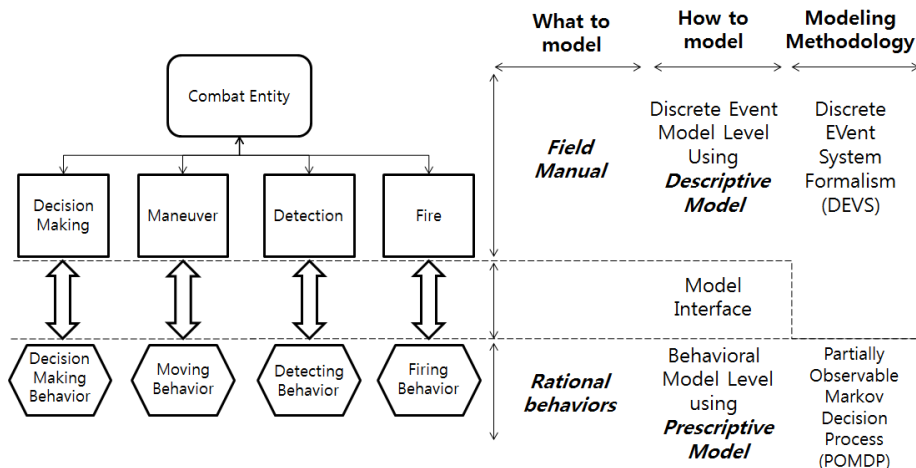


Figure 1. Model integration of descriptive model and prescriptive model

탄 낙하 등과 같은 전술적 상황에 대처하기 위한 상황조치로서, 본 시나리오에서는 홍군 경계팀(red team)의 시야 밖으로 우회하여 차단진지(blocking position)를 점령하고 퇴각하는 적에 대한 매복전투를 수행한다. 한편, 청군 분대 1팀은 홍군과 교전하기 전에, 급편진지를 우선 점령하여 청군 분대 2팀의 기동에 대한 엄호 및 지원사격을 실시한다. 청군 분대 2팀이 차단진지에 도착하면, 분대 1팀은 홍군 경계 진지로 진격하여 홍군 경계팀과 교전한다. 한편, 홍군 경계팀은 접근하는 청군 분대 1팀을 발견하면 즉각 사격을 실시하여 교전을 수행한다. 교전 중에 홍군의 전력이 30% 이하로 떨어지게 되면, 홍군은 차단 진지 방향으로 후퇴를 시작한다. 청군 분대 1팀은 홍군이 퇴각하기 시작하면 홍군 경계 진지로 돌격하여 진지를 탈취하고 시나리오는 종료된다.

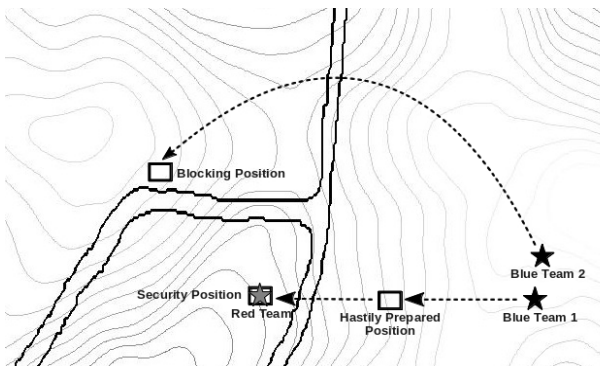


Figure 2. Squad-level Engagement scenario around Mt. BongSeo

본 연구에서는 실제적인 전장 환경 모의를 위해 봉서산 일대 약 300m×400m의 고지를 모의 전장으로 설정하고, 해당 고지의 실제 고도 및 지형 정보를 이용하여 전장 환경을 구성하였다. 전투 개체의 탐지, 교전, 기동 등의 개체 행동은 전장 환경에 모의된 고도 및 지형에 영향을 받게 되고, 이로부터 전장 상황 전개의 불확실성 및 전투 개체들의 전장 상황 인식을 유발하게 된다. 예를 들어, 고도 정보는 전투 개체의 이동 및 시야에 영향을 주게 된다. 전투 개체의 이동 시, 주어진 이동 속도는 고도의 변화에 따라 고도가 높아질수록 속도가 느려지고 고도가 낮아질수록 속도가 높아지도록 모의하였다. 또한 한 전투 개체가 다른 전투 개체를 관찰 시 두 개체 사이의 고도 차이로 인해 가시선이 가려질 경우, 전투 개체 관찰에 제한을 받게 된다. 지형 정보 또한 전투 개체의 이동 속도에 영향을 미치게 된다. 예를 들어, 지형 정보가 도로, 적은 숲, 울창한 숲 등으로 분류되어 있다면, 전투 개체의 실제 이동 속도는 주어진 이동 속도의 100%, 80%, 60%가 되도록 모의하였다.

#### 4. 전투 개체 모델 개발

본 연구의 전투 개체 모델(combat-entity model)은 군사 교리를 반영하는 이산 사건 모델(DEM)과 군사 교리를 고려하여 세부 행동을 결정할 수 있는 행위 모델(BM)로 분리하여 개발하였다. 특히, DEM과 BM은 각각 DEVS 형식론과 POMDP 프레임워크를 이용하여 개발한 후, 독립적으로 개발된 DEM과 BM을

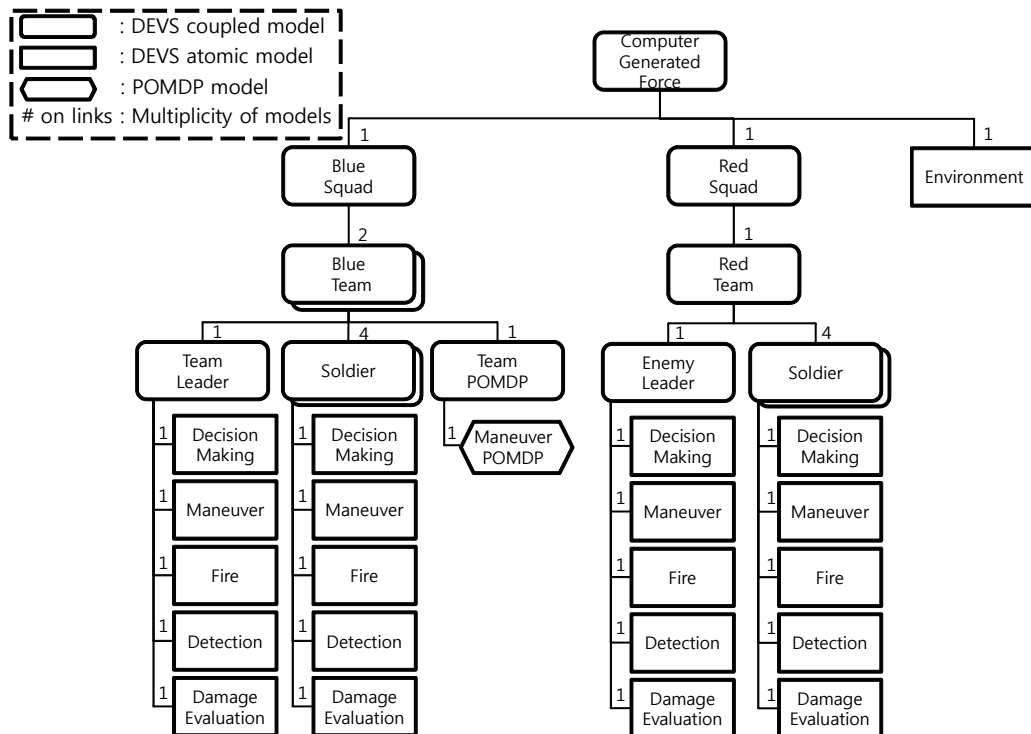


Figure 3. Model structure of the computer generated force model

모델 인터페이스로 연결하여 전투 개체 모델을 표현하였다. 또한, 개발된 전투 개체 모델들을 조합하여 분대 규모의 가상군(computer generated force; CGF) 모델을 개발하였다. 아래 세부 절들에서는 CGF 모델의 전체 구조를 설명하고, CGF를 구성하는 전투 개체 모델과 그 내부의 DEVS 모델, POMDP 모델 그리고 모델 인터페이스에 대해서 자세히 설명한다.

### 4.1 가상군 모델의 구조

본 연구의 가상군(CGF) 모델은 제 3장의 시나리오에 따라 분대 규모로 구성하였다. 가상군 모델은 분대 규모의 청군(아군)과 반개 분대 규모의 홍군(적군)의 분대 전투를 모의 한다. 시나리오에 따라, 청군의 분대는 5인의 팀 두 개로 구성되어 있고 홍군의 분대는 5인의 팀 한 개로 구성되어 있다. 가상군 모델에서 전투 개체 모델은 기동, 탐지, 교전 및 의사 결정을 수행하는 팀원을 나타내며, 교리를 반영하는 DEVS 모델과 행동을 결정할 수 있는 POMDP 모델로 이루어져 있다. 특히, POMDP 모델은 팀의 기동 교리를 고려하여 전투 개체의 기동 방법을 선택하는데 사용하였다.

<Figure 3>은 가상군 모델의 구성을 나무 구조(tree structure)로 표현하였다. <Figure 3>에서 모델과 모델의 계층관계는 모델의 높이(height)와 연결선으로 표시하고, 연결선 위의 숫자는 계층관계에 포함된 모델의 수(multiplicity)를 의미한다. 예를 들어, 가상군 모델은 한 개의 청군 분대(blue Squad) 모델, 한 개의 홍군 분대(red squad) 모델 그리고 한 개의 환경 모델(environment)로 구성된다. 청군 분대 모델은 두 개의 청군 팀 모델(blue team)로 이루어져 있으며, 각 팀은 직진 기동 임무와 우회 기동 임무를 수행한다. 팀 모델은 시나리오를 따라 다섯 개의 전투 개체 모델로 구성되어 있으며, 팀 명령을 전달하는 한 개의 리더 모델(team leader)과 리더의 명령을 따르는 네 개의 사병 모델(soldier)로 이루어져 있다. 리더와 사병을 포함한 전투 개체 모

델은 의사 결정 모델(decision making), 기동 모델(manuever), 교전 모델(fire), 탐지 모델(detection) 그리고 피해 평가 모델(damage evaluation)로 구성된다. 홍군 분대는 한 개의 팀(red team) 모델로 이루어져 있고, 팀 모델의 구성은 청군 분대의 팀 모델과 같다. 환경 모델은 전장의 환경을 모의하는 모델로, 청군과 홍군의 행동 결정에 영향을 미치는 모델이다. POMDP 모델은 청군 분대의 전투 개체 모델들의 기동 모델과 연결되어 팀의 기동 대형을 유지하는 동시에, 효율적인 기동 방법을 선택하는데 사용된다.

<Table 1>은 가상군 모델과 실험에서 사용하는 내부 변수(variable)를 설명한다. 입력 변수(input)는 가상군 모델을 이용한 실험 시 값을 변경하면서 수행할 변수를 의미하며, 출력 변수(output)는 실험의 결과로 출력되는 변수를 의미한다. 시나리오 파라미터는 시뮬레이션 시나리오에 의해 결정되며, 가상군 모델의 규모와 초기 위치 등의 정보를 나타낸다. 전투 개체 모델 변수는 군사 교리나 전문가의 의견을 반영하여 결정되며, 리더 혹은 사병 모델과 같은 전투 개체 모델의 기동, 탐지, 교전 및 의사 결정과 관련된 정보를 의미한다. 예를 들어, <Table 1>의 기동 속도는 군사 교리의 보병 부대의 평균 이동 속력인 4km/h를 반영하였다. 환경 모델 변수는 시나리오 대상 지역의 실제 지형 및 고도 정보를 나타내며, 전투 개체 모델들의 기동, 탐지 및 교전 등의 결과에 영향을 미치는 요소로 작용한다.

### 4.2 가상군 모델의 이산 사건 모델

전투 개체 모델의 이산 사건 모델은 전투 교리를 충실하게 반영하는 것을 목적으로 한다. 아래의 세부 절에서는 이산 사건 모델을 개발하는 데 사용된 DEVS 형식론에 대해서 설명하고, 이산 사건 모델의 결합 모델과 원자 모델에 대해 차례로 설명한다. DEVS(discrete event system specification)는 B.P. Zeigler에 의해 제안된 집합론에 근거한 이산 사건 시스템(discrete

Table 1. Variable list used in the computer generated force model

Variable Type	Variable Name	Initial value(unit)	Implications
Input	Dispersions in the formation	Varied in experiments	Distance between two combat-entity models of maneuver formation
	Angles in the formation		Angles formed where two lines in the maneuver formation cross each other
	Maneuver Behavior Model		Behavior models describing maneuver of combat-entity models
Output	Formation violation penalty	-	Performance measure for evaluating how much team formation is deviated
	Damage ratio of the red		Ratio of military power of the red between at the start and end of simulation
	Damage ratio of the blue		Ratio of military power of the blue between at the start and end of simulation

Scenario Parameters	Number of blue squads	Varied in scenarios	Number of blue squads which were sent to the mission	
	Number of teams in a blue squad		Number of teams in each blue squad	
	Number of soldiers in a blue team		Number of soldiers in each blue team	
	Number of red squads		Number of red squads which were sent to the mission	
	Number of teams in a red squad		Number of teams in each red squad	
	Number of soldiers in a red team		Number of soldiers in each red team	
	Initial points		Initial points of combat-entity models	
	Withdraw point		Withdraw point for a retreat of the red	
	Mission objectives		Mission objectives of the blue	
	Mission maneuver points		Mission points for maneuver missions	
Parameters in the DEM(DEVS)	Detection range	150(m)	Available detecting range of the combat-entity model	
	Shot range	50(m)	Available shot range of the combat-entity model	
	Probability of kill	70(%)	probability of damaging other combat-entity model within the shot range	
	Current point	-	Current point of the combat-entity model	
	Objective point		Mission maneuver point of the combat-entity model	
	Maneuver objectives		Mission objectives of the combat-entity model	
	Maneuver formation		Maneuver formation of the team which the combat-entity model joined to	
	Detected information of the blue		Detected information of the blue in a specific time of the running simulation	
	Detected information of the red		Detected information of the red in a specific time of the running simulation	
	Maneuver direction		Maneuver direction of the combat-entity model to maintain maneuver formation of the team	
	Maximum maneuver speed		3.33(m/s)	Maximum maneuver speed of the combat-entity model
	Current maneuver speed		-	Maneuver speed of the combat-entity model in a specific time of the running simulation
	Elevation		Varied in scenarios	Real elevation information from the scenario regions
Landform	Real landform from the scenario regions			
Noise(mean, variance)	(0.0, 1.0)		Mean and variance value which are used to evaluate noise information affecting detection and engagements	
Parameters in the BM(POMDP)	Weight for reaching goal ( $w_{goal}$ )		1	The weight for reaching goal in the reward function
	Weight for keeping formation( $w_{formation}$ )	1	The weight for keeping formation in the reward function	
	Weight for avoiding red team( $w_{detected}$ )	0 for blue team 1, 1 for blue team 2	The weight for avoiding red team in the reward function	

event system) 형식론이다 (Zeigler, 2000). DEVS 형식론은 원자 모델과 결합 모델을 이용하여 계층적이고 모듈적인 모델링을 지원 한다. 원자 모델과 결합 모델의 수학적 명세는 <Figure 4>와 같다.

DEVS의 수학적 명세는 형식론에 익숙하지 않은 모델 개발자와의 소통을 위해서, DEVS 다이어그램을 이용하여 모델의 내용을 표현한다. DEVS 다이어그램은 FSM(Finite State Machine)의 다이어그램을 확장하여 사용하고 있다. DEVS 원자

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

Figure 4. Formal specification of DEVS atomic and coupled model

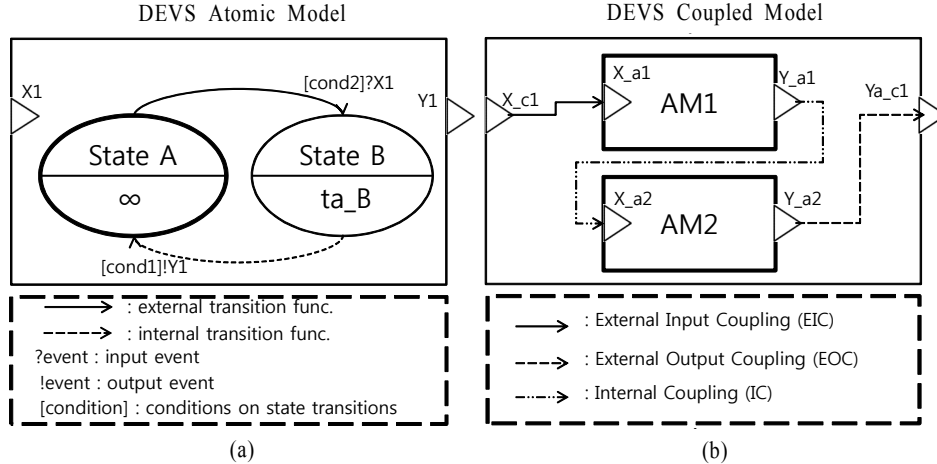


Figure 5. DEVS model diagrams : (a) atomic model (b) coupled model

모델과 결합 모델의 다이어그램은 위의 <Figure 5>와 같이 표현할 수 있다.

<Figure 5>(a)는 DEVS 원자 모델의 다이어그램을 나타낸다. DEVS 원자 모델의 입력(X1)과 출력(Y1) 이벤트는 다이어그램 바깥쪽의 양 옆에 삼각형으로 표시 한다. 원자 모델의 상태는 타원으로 나타내고, 상태의 시간 전진(time advance)값은 밑줄 밑에 실수 값 혹은 변수로 나타낸다. 예를 들어, <Figure 5>(a)의 'State A'에서 입력 이벤트 'X1'을 받고 cond2의 조건을 만족하면 'State B'로 천이하는 내용이 명시되어 있다. <Figure 5>(b)는 DEVS 결합 모델의 다이어그램을 나타낸다. 예를 들어, <Figure 5>(b)의 결합 모델은 AM1과 AM2의 컴포넌트 모델을 포함하고 있다. 결합 모델의 연결 관계(coupling relation)은 결합 모델과 컴포넌트 모델들의 입, 출력 이벤트를 연결하는 실선 화살표로 표시한다.

(1) 가상군 모델의 결합 모델

이번 절에서는 제 4.2.1절에서 설명한 가상군 모델의 이산 사건 모델 중, 모델의 구조를 표현하는 결합 모델에 대해서 설명한다. <Figure 3>과 같이, 가상군 모델의 결합 모델은 가상군 모델(CGF model), 분대 모델(squad model), 팀 모델(team model) 그리고 사병 모델(soldier model)이 있다. 다음의 <Figure 6>은 가상군 모델에서 사용된 결합 모델을 다이어그램으로 표현했다.

<Figure 6>(a)는 가상군 결합 모델을 나타내며, 가상군 결합

모델은 청군, 홍군 분대 모델과 환경 모델로 구성되어 있다. 가상군 모델은 입력으로 시나리오에 기술된 임무를 받고(scenario\_in), 팀의 상태 및 정보를 기록하기 위하여 각 분대의 보고를 외부로 보낸다(report\_out). 청군과 홍군의 분대는 임무의 명령(order\_in)을 바탕으로 부대를 통솔한다. 환경 모델은 청군과 홍군 모델로부터 사망(dead\_out), 기동(mnv\_out) 그리고 탐지 요청(det\_out)정보를 받고, 탐지 결과(env\_out)를 요청 분대에게 전달한다. 또한, 청군과 홍군 분대는 교전 정보(fire\_out)를 교환하며 분대간 교전을 모의한다. <Figure 6>(b)는 청군 분대 모델의 내부 연결 관계를 보여준다. 청군 분대 모델에는 두 개의 팀 모델이 존재하고 각 팀 모델은 팀의 전투 개체의 사망, 기동, 탐지 정보를 분대 모델의 출력 이벤트로 내보낸다. 특히 청군 분대의 두 개의 팀 중, 청군 분대 1팀(blue team1) 모델은 시나리오에 따른 작전 명령을 입력받은 후에, 청군 분대 2팀(blue team2) 모델에게 세부 부대 명령(order\_out)을 전달하여 작전을 수행한다. <Figure 6>(c)는 청군 팀 모델 내부의 연결 관계를 나타낸다. 팀 모델은 한 개의 리더(team leader)와 네 개의 사병(soldier) 모델로 구성된다. 리더와 사병을 포함한 팀원들은 각자의 사망, 기동 그리고 탐지 정보를 외부로 보내고, 상태 정보(report\_in)를 리더에게 전송한다. <Figure 6>(d)는 팀의 구성원 중 하나인 사병의 결합 모델의 다이어그램이다. 사병 모델은 의사 결정(decision making), 기동(manuever), 탐지(detection), 교전(fire) 그리고 피해평가(damage evaluation) 모델로 구



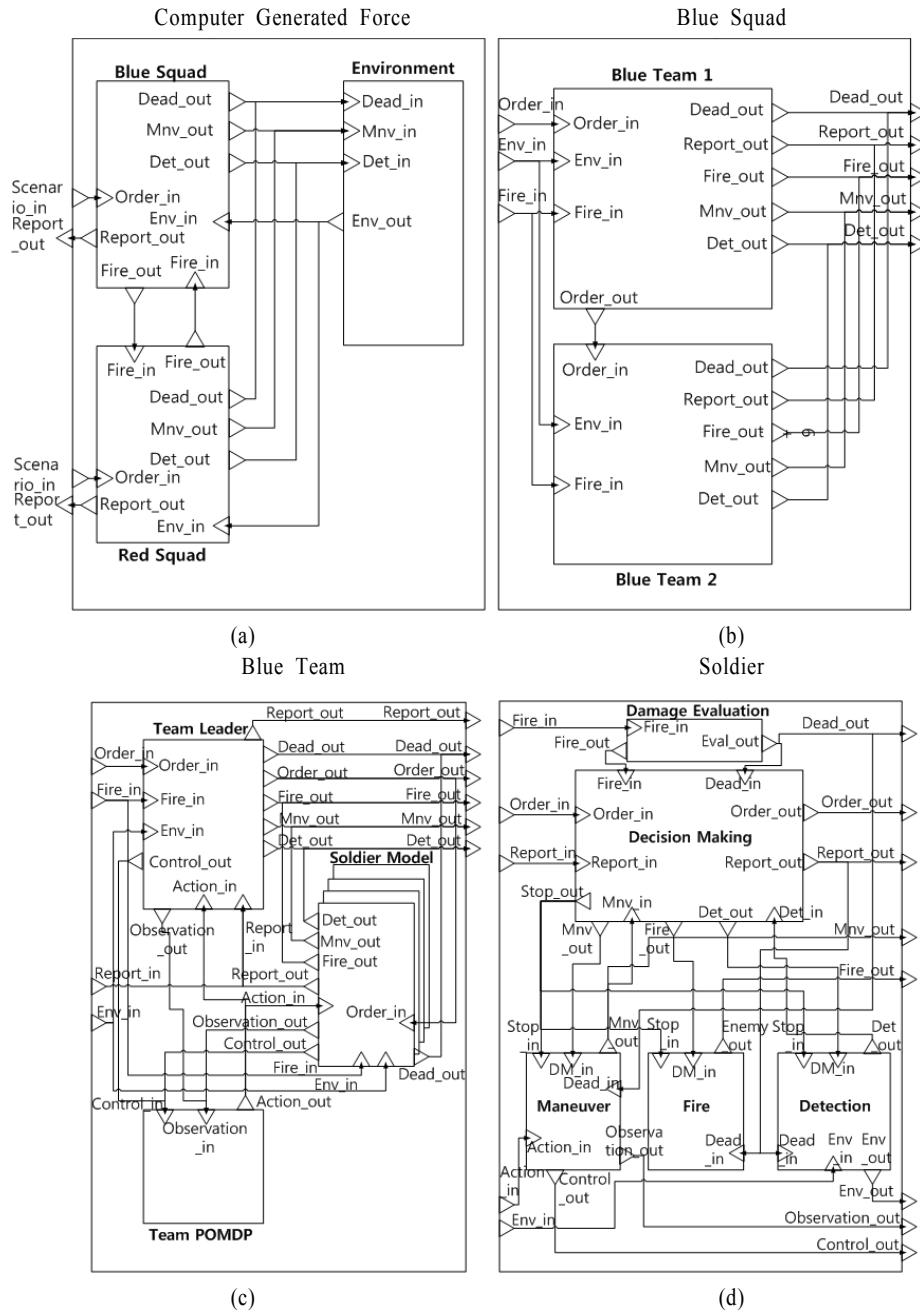


Figure 6. Coupled models in the CGF model : (a) CGF model, (b) Blue squad model, (c) Blue team model, and (d) Soldier model

성된다. 사병 모델의 컴포넌트 모델들은 DEVS 원자 모델이며, 사병 모델은 컴포넌트 모델들이 생성하는 이벤트를 외부로 전달하고, 외부에서 전달된 이벤트를 원자 모델의 입력 이벤트로 전달한다.

(2) 가상군 모델의 원자 모델

이번 절에서는 제 4.2.1절에서 설명한 DEVS 형식론을 적용해서 개발한 가상군 모델의 이산 사건 모델 중, 모델의 행위를 모의하는 원자 모델에 대해서 설명한다. <Figure 3>의 가상군 모델에 포함된 DEVS 원자 모델은 환경(environment) 모델과

전투 개체 모델의 기동(maneuver), 교전(fire), 탐지(detection), 피해 평가(damage evaluation), 의사결정(decision making) 모델이 있다. 각 원자 모델의 행위 내용은 다음 <Table 2>에 간단히 기술한다.

위의 원자 모델 중, 가장 복잡한 구조를 띄고 있는 의사 결정 모델의 다이어그램을 하나의 예제로써 설명 한다. 의사 결정 모델은 초기 상태(ready)에서 외부로부터 작전 명령(order\_in)이 입력 이벤트로 들어오면, 결정 상태(decide)로 천이하여 명령의 내용을 파악하고 전투 개체의 다음 행동(next action)을 결정한다. 입력 받은 작전 명령에는 전투 개체의 행동 방침(course

Table 2. Model behaviors of atomic models in the CGF model

Atomic Model	Behaviors
Environment	<ul style="list-style-type: none"> <li>◦ Describing position of combat-entity models, and elevation and landform in the battlefield</li> <li>◦ Affecting maneuver, detection, and result of engagement of combat-entity models</li> </ul>
Maneuver	<ul style="list-style-type: none"> <li>◦ Updating position of a combat entity model according to the result of POMDP model</li> <li>◦ Deciding new position of a combat-entity model reflecting mission objectives and formation in the POMDP model</li> </ul>
Fire	<ul style="list-style-type: none"> <li>◦ Describing direct firing engagement between combat-entity models</li> </ul>
Detection	<ul style="list-style-type: none"> <li>◦ Detecting combat-entity models with detection range and battlefield</li> <li>◦ Forwarding detecting information to decision making and team POMDP model</li> </ul>
Damage Evaluation	<ul style="list-style-type: none"> <li>◦ Evaluating damage and death of combat-entity model using the result of engagement</li> </ul>
Decision Making	<ul style="list-style-type: none"> <li>◦ Deciding mission of the squad using scenario and detection information</li> <li>◦ Controlling maneuver, engagement, and detection models based on detection information</li> </ul>

of action)이 들어있다. 행동 방침은 전투 개체가 작전 중에 따라야할 일련의 행동들으로써, 군사 교리에 기반을 두고 시나리오를 고려하여 구성된다. 분대 작전 군사 교리와 3장의 시나리오를 바탕으로, 가상군 모델에서는 전략적 기동(tactical move; TC\_Move), 매복대기(ambush wait; A\_Wait), 제압 사격(cover fire; Cvr\_Fire) 그리고 돌격 기동(charge move; Chr\_Move)의 행동들을 이용하여 모든 전투 개체에 대한 행동 방침이 구성되어 있다. 각 전투 개체의 행동 방침은 전투 개체가 편성되어 있는 부대와 전투 개체의 역할에 따라서 달라질 수 있다. 예를 들어, 분대 1팀 리더의 행동 방침은 분대 1팀 사병의 행동 방침과 다르고, 분대 2팀 리더의 행동 방침과도 다르다. 이런 다양한 전투 개체의 행동 방침을 처리하기 위해서, 가상군 모델에서는 <Figure 7>과 같이 행동 방침의 구성에 관계없이 의사 결정을 할 수 있는 일반적인 의사 결정 모델을 이용하였다. 일반적인 의사 결정 모델에서는 작전 명령 이벤트를 통해서 행동 방침을 입력 받고, 입력 받은 행동 방침에 따라 전투 개체의 다음 행동을 결정한다. 결정된 행동의 상태로 천이한 후에, 선택된

행동에 해당하는 세부 작전들을 수행한다. 작전의 세부 행동들이 모두 종료되면, 작전의 결과와 모델의 상태(report\_out)를 외부로 전달하고 다음 행동을 결정하는 상태(decide)로 천이한다. 예를 들어, 입력 받은 행동 방침에 전략적 기동(TC\_Move)의 행동이 들어 있다면, 전투 개체의 다음 상태를 'TC\_Move' 상태로 천이한다. 천이한 후에 차례로 'Stop', 'R\_Fire' 그리고 'Report'의 상태를 거쳐 'Decide' 상태로 천이한다. 'Decide' 상태에서는 다시 행동 방침을 참고하여, 다음의 행동을 선택하고 그에 맞는 상태로 천이하는 과정을 반복한다.

### 4.3 전투 개체의 행동 모델 구조

본 논문에서 전투개체의 행동 모델 중, 교전 및 탐지에 해당되는 행동은 DEVS를 활용한 기술적 모델링을 활용하였다. 이는 국방 M&S에서 전통적으로 활용된 확률 기반의 이산 사건 시스템 모델링을 이론적 기반으로 삼고 있다. 본 논문의 실험으로 POMDP를 전투 개체의 기동 행동에만 적용하였다. 전투

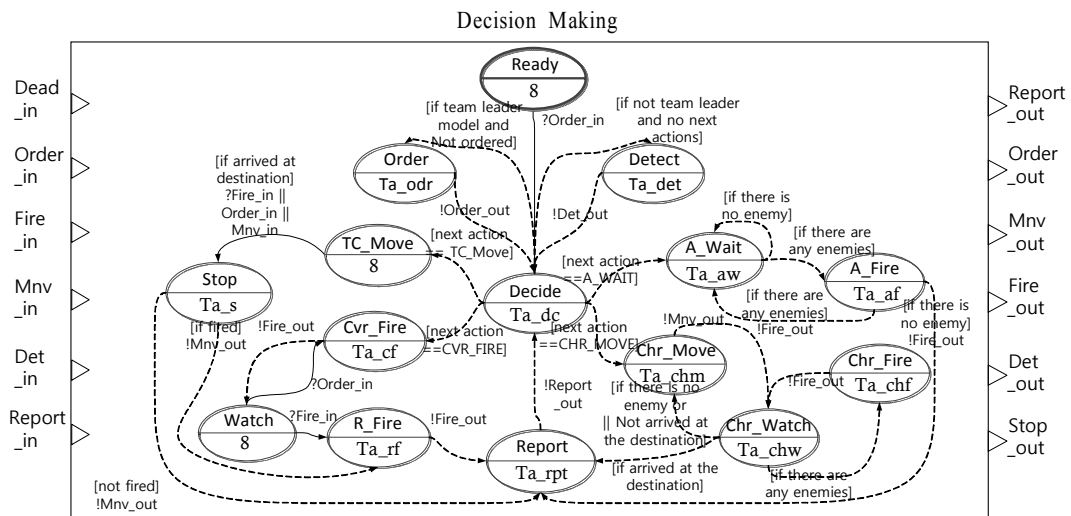


Figure 7. DEVS diagram of decision making model

개체의 탐지행동의 경우 본 시뮬레이션 시나리오 상황에서는 확률적 이산 사건을 기반으로 한 적탐지 이외에는 별다른 행동의 차이점이 발생할 수 없으므로, POMDP의 적용 의미가 크지 않다. 교전의 경우는 교전 대상 및 방법의 선정에서 행동 모델링이 중요하므로 POMDP가 잘 적용될 수 있는 분야이다. 다만 소규모 분대 전투에서 교전 행위가 Reward 기반의 이성적 행동에 속할 것인지, 아니면 교범에 의해 훈련된 반사 신경에서 나오는 행위인지는 논란의 여지가 있다. 만약 POMDP가 교전에 적용될 경우, 기존의 전투 개체 행동과는 상당히 다른 분석 결과가 도출될 것으로 예상되나, 이는 실험을 통해 검증해야 할 사항이며 향후 연구로 미루어둔다.

POMDP는 전장 상황의 불확실한 또는 부분적인 관찰을 바탕으로 전투 개체의 현재 상황을 추론하고 최적의 행동 정책을 계산하기 위한 순차적 의사결정 이론 프레임워크이다. POMDP 모델은 수학적으로  $\langle S, A, Z, T, O, R, b_0 \rangle$ 로 정의된다.  $S$ 는 전장 상황의 가능한 환경 상태(state)들의 집합,  $A$ 는 전투 개체가 수행할 수 있는 가능한 행동(action)들의 집합,  $Z$ 는 주어진 환경 상태에서 전투 개체가 관찰할 수 있는 가능한 관찰 값(observation)들의 집합,  $T$ 는 전투 개체의 환경 상태 전이 함수로서 현재 환경 상태  $s$ 로부터 행동  $a$ 를 수행하여 환경 상태  $s'$ 로 이동할 환경 상태 전이 확률(transition probability)  $T(s, a, s') = \Pr(s'|s, a)$ ,  $O$ 는 전투 개체의 관찰 함수로서 행동  $a$ 를 수행하여 환경 상태  $s'$ 로 이동한 후 관찰 값  $z$ 를 관찰할 확률(observation probability)  $O(s', a, z) = \Pr(z|s', a)$ ,  $R$ 은 전투 개체의 보상 함수(reward function)로서 현재 환경 상태  $s$ 에서 행동  $a$ 를 수행하여 다음 환경 상태  $s'$ 으로 갈 때 얻는 보상 값(reward)  $R(s, a, s')$ ,  $b_0$ 는 전투 개체의 초기 환경 상태 확률 분포(initial belief state)를 나타낸다. POMDP에서는 전장 상황을 불확실한 혹은 부분적으로만 관찰 가능하기 때문에 현재 환경 상태를 직접적으로 알 수 없지만, 초기 환경 상태 확률 분포로부터 수행한 행동 및 관찰 값들의 이력(history)을 바탕으로 시간  $t$ 에서의 환경 상태 확률 분포(belief state)  $b_t$ 를 추론하게 된다. POMDP에서는 다음 환경 상태는 오직 현재 환경 상태와 현재의 행동에만 영향을 받게 되는 마코프 속성(Markov property)를 가정한다.

POMDP에서의 행동 정책(policy)이란 각 belief state에서 수행할 행동에 대한 함수로서  $\pi: \Delta(S) \rightarrow A$ 로 표현된다(여기서  $\Delta(S)$ 란  $|S|$ 차원의 확률 심플렉스(probability simplex)로서 belief state가 존재하는 공간을 나타낸다). 행동 정책의 좋고 나쁨을 평가하기 위한 기준으로 가치 함수(value function)  $V$ 가 있다. POMDP를 푸는 것은 위의 value function을 최대화하는 최적 행동 정책(optimal policy)  $\pi^*$ 를 찾는 것이다. 그리고 이에 해당하는 value  $V$ 를 optimal value  $V^*$ 라고 부른다.

#### (1) 전투 개체의 자율 기동을 위한 기동 POMDP 모델 정의

본 연구에서는 청군 분대에서 시나리오 상 직진 기동을 하는 분대 1팀(분대장조, blue team 1)과 우회 기동을 하는 분대 2

팀(부분대장조, blue team 2)의 기동 모델을 위해 POMDP 모델을 사용한다. 각 팀은 하나의 POMDP 모델로 정의되어, 작전 및 기동 대형 등을 감안하는 5인 병사의 움직임을 결정짓는다. 두 팀의 구성 및 제원은 같으므로 POMDP 모델에서 환경 상태, 행동, 관찰 값들의 집합( $S, A, Z$ ) 및 환경 상태 전이 확률( $T$ ), 관찰 확률( $Z$ )은 두 팀이 동일하다. 하지만 시나리오 상 직진 기동 후 홍군(red team) 섬멸을 목적으로 하는 분대 1팀과 우회 기동 후 홍군의 퇴로 차단을 목적으로 하는 분대 2팀은 서로 다른 목표를 가지게 되므로, 서로 다른 보상 함수를 통해 이를 표현해야 한다.

- **환경 상태**: 각 팀의 환경 상태  $s$ 는  $\langle \vec{\alpha}, \vec{\beta}, \vec{g} \rangle$ 로 정의된다.  $\vec{\alpha}$ 는 청군 각 팀의 분대원들의 위치 좌표들이다. 즉  $\vec{\alpha} = \langle \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \rangle$ 로 표현할 수 있고, 각  $\alpha_i$ 는  $i$ 번째 분대원의  $\langle x, y \rangle$  위치 좌표로서  $\alpha_i = \langle \alpha_{i,x}, \alpha_{i,y} \rangle$ 로 표현된다. 편의상 1번째 분대원을 팀의 리더(분대장 혹은 부분대장)로 가정한다. 그리고  $\vec{\beta} = \langle \beta_x, \beta_y \rangle$ 는 관찰 값으로부터 추정되는 홍군 병사 실제 위치들의 중앙점(centroid),  $\vec{g} = \langle g_x, g_y \rangle$ 는 도달하고자 하는 목표 지점의 위치를 표현한다.
- **행동**: 각 팀의 행동  $a$ 는 청군 각 팀의 분대원들의 이동 속도 및 방향  $\langle \vec{r}, \vec{\phi} \rangle$ 로 정의된다.  $\vec{r} = \langle r_1, r_2, r_3, r_4, r_5 \rangle$ 은 분대원 각각의 이동 속도로서  $r_i$ 는 0에서 주어진 분대원의 최고 이동 속도  $r_{MAX}$  이하의 연속적인 실수값을 가진다.  $\vec{\phi} = \langle \phi_1, \phi_2, \phi_3, \phi_4, \phi_5 \rangle$ 는 분대원 각각의 이동 방향을 라디안(radian)으로 나타내고,  $\phi_i$ 는 0에서  $2\pi$  사이의 연속적인 실수값을 갖는다.
- **관찰 값 및 관찰 확률**: 각 팀의 관찰 값  $z$ 는 환경 상태와 동일하게  $\langle \vec{\alpha}, \vec{\beta}, \vec{g} \rangle$ 로 정의된다. 단, 본 시나리오에서는 청군 자신의 위치  $\vec{\alpha}$ 는 온전하게 알려진다고 가정하고, 목표 위치  $\vec{g}$ 도 DEVS에서 지정되기 때문에  $\vec{\alpha}, \vec{g}$  값은 청군 각 팀의 분대원들의 환경 상태의 실제 위치 및 실제 목표 지점의 값이 들어온다. 하지만 홍군의 위치  $\vec{\beta}$ 는 청군 각 팀의 위치로부터 홍군의 위치까지의 가시선 차단 유무(line of sight; LOS)에 따라 관찰 불가능하다고 가정한다. LOS 값이 참(true)일 경우 청군 각 팀은 홍군의 실제 위치를 관찰 값으로 받게 되고, 거짓(false)일 경우 청군 각 팀은 홍군의 위치에 대한 관찰 값을 얻지 못하게(null) 된다.
- **환경 상태 전이 확률**: 환경 상태  $s$ 에서 행동  $a$ 를 수행하고 나면 청군 각 팀의 분대원은 현재 위치  $\vec{\alpha}$ 로부터 행동  $a$ 가 표현하는 방향과 거리를 이동한다. 즉,  $i$ 번째 분대원은 현재 위치  $(\alpha_{i,x}, \alpha_{i,y})$ 로부터  $(\alpha_{i,x} + r_i \cos \phi_i, \alpha_{i,y} + r_i \sin \phi_i)$ 로 이동하려고 시도한다. 하지만 전장의 지형 속성 및 경사도는 분대원들의 이동에 영향을 미치게 된다. 또한, 전장 상황에 따른 분대원들의 이동 결과의 불확실성을 반영하기 위해 확률적 부가 오차를 사용한다. 홍군은 현재 위치에서

확률적 부가 오차만큼의 이동이 있다고 가정한다. 본 연구에서는 확률적 부가 오차로서 평균이 0인 Gaussian noise를 사용하였다. 정리하면, 청군  $i$ 번째 분대원은 아래 식 (1)

$$(\alpha_{i,x}, \alpha_{i,y}) \rightarrow (\alpha_{i,x} + \rho r_i \cos \phi_i + N(0, 1), \alpha_{i,y} + \rho r_i \sin \phi_i + N(0, 1)) \quad (1)$$

where  $\rho$ : 지형 속성 및 경사도에 따른 이동거리 감소율,  
 $N(0, 1)$ : Gaussian noise

에 의해 이동하고 홍군은 아래 식 (2)

$$(\beta_x, \beta_y) \rightarrow (\beta_x + N(0, 1), \beta_y + N(0, 1)) \quad (2)$$

에 의해 이동한다고 정의하였다.

• **보상 함수.** 각 팀의 기동에 있어서 전투 개체가 달성해야 할 목표는 다음의 3가지가 있다.

1. 목표 지점까지의 이동  $r_{goal}(s, a, s')$
2. 기동 시 대형 유지  $r_{formation}(s, a, s')$
3. 홍군의 탐지를 피함으로써 공격 회피  $r_{detected}(s, a, s')$

보상 함수  $R(s, a, s')$ 은 위의 3가지를 조합하여 다음과 식 (3)과 같이 정의한다.

$$R(s, a, s') = w_{goal} r_{goal}(s, a, s') + w_{formation} r_{formation}(s, a, s') + w_{detected} r_{detected}(s, a, s') \quad (3)$$

이 때, 분대 1팀의 목표는 홍군의 탐지 여부와 상관없이 대형을 유지하여 목표 지점까지 빠르게 이동하는 직진 기동이다. 따라서 분대 1팀에서는 3가지 목표 중 앞의 두 가지만을 만족시켜야 하고,  $w_{goal} = 1, w_{formation} = 1, w_{detected} = 0$ 으로 정의할 수 있다. 분대 2팀의 목표는 홍군의 탐지를 피함으로써 적군으로부터 공격받지 않고 우회 기동하는 것이다. 따라서 분대 2팀에서는 3가지 목표를 모두 만족시켜야 하고,  $w_{goal} = 1, w_{formation} = 1, w_{detected} = 1$ 으로 정의할 수 있다.

$r_{goal}(s, a, s')$ 은 목표 지점까지의 이동을 표현하기 위한 보상 함수이다.  $r_{goal}(s, a, s')$ 을 정의하기 위해서 환경 상태  $s$ 에서 팀의 리더와 목표 지점 사이의 거리로서 정의되는 포텐셜 (potential) 함수  $\Phi_{goal}(s)$ 를 정의한다. 그러면  $r_{goal}(s, a, s')$ 은 현재 환경 상태  $s$ 에서 행동  $a$ 를 취한 후 천이될 다음 환경 상태  $s'$ 에서 팀의 리더와 목표 지점 사이의 거리  $\Phi_{goal}(s')$ 과 행동  $a$ 을 하기 전 현재 환경 상태  $s$ 에서 팀의 리더와 목표 지점 사이의 거리  $\Phi_{goal}(s)$ 의 차이  $r_{goal}(s, a, s') = -(\Phi_{goal}(s') - \Phi_{goal}(s))$ 로 정의할 수 있다.

$r_{formation}(s, a, s')$ 은 기동시 분대원들 사이의 거리 및 각도를 유지하기 위한 보상 함수이다.  $r_{formation}(s, a, s')$ 은 분대원간 거리 유지를 위한 보상 함수  $r_{formation,d}(s, a, s')$ 과 분대

원간 각도 유지를 위한 보상 함수  $r_{formation,\theta}(s, a, s')$ 의 합  $r_{formation}(s, a, s') = r_{formation,d}(s, a, s') + r_{formation,\theta}(s, a, s')$ 으로 정의한다.

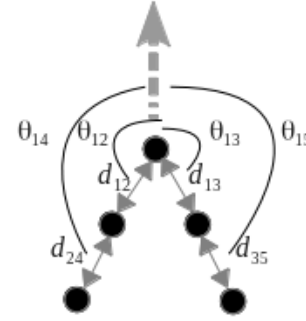


Figure 8. Reward function with maneuver formations

<Figure 8>은 분대원들이 맞추고 있는 대형을 나타낸다.  $d_{ij}$ 는 현재 환경 상태  $s$ 에서  $i$ 번째 분대원과  $j$ 번째 분대원 사이의 거리를,  $\theta_{1i}$ 는 현재 환경 상태  $s$ 에서  $i$ 번째 분대원과 팀의 리더가 향하는 방향 사이의 각도를 나타낸다. 분대원간 유지해야 하는 거리가  $d$ 라면 분대원들이 이 거리를 유지할 때 받는 보상값  $r_{formation,d}(s, a, s') = -\exp((d_{12}-d)^2 + (d_{24}-d)^2 + (d_{13}-d)^2 + (d_{35}-d)^2)$ 이 된다. 분대원간 유지해야 하는 각도가  $\theta$ 라면 분대원들이 이 각도를 유지할 때 받는 보상값은  $r_{formation,\theta}(s, a, s') = -\exp((\theta_{12}-\theta)^2 + (\theta_{14}-\theta)^2 + (\theta_{13}-\theta)^2 + (\theta_{15}-\theta)^2)$ 이 된다.

$r_{detected}(s, a, s')$ 는 홍군의 탐지를 피하기 위해 홍군 시야 밖으로 우회 기동을 하기 위한 보상 함수이다.  $r_{detected}(s, a, s')$ 를 정의하기 위해서 환경 상태  $s$ 에서 팀의 리더와 홍군 사이의 거리로 정의되는 포텐셜 함수  $\Phi_{detected}(s)$ 를 정의한다. 그러면  $r_{detected}(s, a, s')$ 은 홍군의 가시선이 청군에 닿을 경우에, 현재 환경 상태  $s$ 에서 행동  $a$ 를 취한 후 천이될 다음 환경 상태  $s'$ 에서 팀의 리더와 홍군 사이의 거리  $\Phi_{detected}(s')$ 과 행동  $a$ 을 하기 전 현재 환경 상태  $s$ 에서 팀의 리더와 홍군 사이의 거리  $\Phi_{detected}(s)$ 의 차이  $r_{detected}(s, a, s') = \Phi_{detected}(s') - \Phi_{detected}(s)$ 로 정의할 수 있다. 홍군의 가시선이 청군에 닿지 않을 경우에는  $r_{detected}(s, a, s') = 0$ 으로 정의한다. <Table 3>에서는 POMDP 모델에 대한 용어 및 가상군 모델링에서의 적용 방법을 정리하였다.

(2) 최적 경로 탐색을 위한 POMDP 행동 정책 계산 알고리즘  
 전투 개체의 기동 POMDP 모델에서는 환경 상태, 행동 및 관찰 값을 연속(continuous) 변수로 정의하였다. 따라서 유한개의 환경 상태, 행동, 관찰 값을 가지는 모델을 가정하는 일반적인 POMDP 행동 정책 계산 알고리즘은 본 모델의 행동 정책 계산에 적합하지 않다. 기존의 연속 환경의 무한 차원(infinite dimension) POMDP의 알고리즘은 유한 차원(finite dimension)

Table 3. Terminology in POMDP model

POMDP Model		
Notation	Implementation	Description
$S$	$\langle \vec{\alpha}, \vec{\beta}, \vec{g} \rangle$	a set of states
$A$	$\langle \vec{r}, \vec{\phi} \rangle$	a set of actions
$Z$	$\langle \vec{\alpha}, \vec{\beta}, \vec{g} \rangle$	a set of observations
$T$	$(\alpha_{i,x}, \alpha_{i,y}) \rightarrow (\alpha_{i,x} + \rho r_i \cos \phi_i + N(0, 1), \alpha_{i,y} + \rho r_i \sin \phi_i + N(0, 1))$ $(\beta_x, \beta_y) \rightarrow (\beta_x + N(0, 1), \beta_y + N(0, 1))$ $\vec{g}$ is not changed	transition probability
$O$	$\langle \vec{\alpha}, \vec{\beta}, \vec{g} \rangle$ if $LOS = true$ $\langle \vec{\alpha}, NULL, \vec{g} \rangle$ if $LOS = false$	observation probability
$R$	$w_{goal} r_{goal}(s, a, s') + w_{formation} r_{formation}(s, a, s') + w_{detected} r_{detected}(s, a, s')$	reward function

POMDP 문제로 근사(approximation)하기 위해 문제의 특징적인 feature를 이용하는 모수적(parametric) 표현 (Bertsekas and Tsitsiklis, 1996) 혹은 Monte-Carlo 방법 (Bertsekas and Castanon, 1999)(Chong *et al.*, 2000), particle filter 등의 비모수적(non-parametric) 표현(Doshi-Velez, 2009)을 사용해왔다. 위의 접근들은 무한 차원 POMDP 문제들에 대한 해법을 제시했지만, 실제 전투 개체 모델링에 적용하기에는 시간 복잡도(time complexity)면에서 여전히 한계를 보인다. 대신에 경로 탐색이라는 특수한 목적을 위한 무한 차원 POMDP 문제들에 대한 근사 알고리즘은 좀 더 현실적인 해법을 제시해 줄 수 있다. 본 연구에서는 온라인 최적 경로 탐색에 특화된 근사 알고리즘 중 하나인 nominal belief-state optimization(NBO)(Miller *et al.*, 2009) 알고리즘을 사용하여 POMDP의 행동 정책 알고리즘을 개발하고자 한다.

NBO는 아래 식 (4)와 같이 표현할 수 있다.

$$V^*(b) \approx \max_{(a_0, \dots, a_{H-1})} \sum_{t=0}^{H-1} R(\hat{b}_t, a_t, \hat{b}_{t+1}) \quad (4)$$

where  $H$ : planning horizon,  $\hat{b}_0 = b$

위의 식에서 보듯이 주어진 belief state  $b$ 에서 주어진 행동 리스트  $(a_0, \dots, a_{H-1})$ 를 취함으로써 얻을 수 있는 다음 belief state들을 하나씩만 샘플링(sampling)하고 그 때의 belief들을 기반으로 value function을 구하게 된다. 이 때 현재 belief state로

부터 샘플링 된 하나의 다음 belief state를 nominal belief state라 부른다. nominal belief state의 계산은 앞 장의 POMDP 환경 상태 천이 확률 및 관찰 확률 정의 시 주었던 랜덤 에러를 0으로 가정하고 가장 확률이 높은 관찰 값에 대한 다음 time step에서의 belief state로 계산된다.

NBO 알고리즘을 사용하여 최적 경로 탐색 과정은 아래의 <Table 4>와 같다.

위의 의사 코드(pseudo-code)에서 보듯이 NBO 알고리즘은 nominal belief-state approximation을 바탕으로 (2)에서 value function을  $(a_0, \dots, a_{H-1})$ 에 대한 함수로 도출하고 (3)에서와 이 함수를 최적화하는 optimization 과정을 거쳐 최적의 행동 시퀀스를 구한다. 본 연구에서는 optimization 문제를 풀기 위해 MATLAB의 fmincon 함수를 사용하였다.

#### 4.4 이산 사건 모델과 행동 모델의 인터페이스

본 연구의 전투 개체 모델은 전투 개체를 이산 사건 모델(DEM)과 행동 모델(BM)로 구분하여 독립적으로 개발한 후, 인터페이스를 통해서 두 모델을 연결한다. 이번 절에서는 이산 사건 모델과 BM의 연결 위해 사용한 인터페이스에 대해서 자세히 설명한다. <Figure 9>는 팀 모델(<Figure 6>(c))의 다이어그램 중, 이산 사건 모델(team leader와 soldier model)과 행동 모델(POMDP model)의 연결 관계를 강조하여 표현하였다.

Table 4. Pseudo-code for NBO algorithm

- Define nominal belief-state sequence  $(\hat{b}_0, \dots, \hat{b}_{H-1})$   
( $H$ : given planning horizon,  $\hat{b}_0 = b$ )

$$\hat{b}_{t+1} = \Phi(\hat{b}_t, a_t)$$

where  $\Phi$ : nominal belief state update assuming that the random noise is 0

- Define approximate value function

$$V^*(b) \approx \sum_{t=0}^{H-1} R(\hat{b}_t, a_t, \hat{b}_{t+1})$$

- Optimize over action sequence  $(a_0, \dots, a_{H-1})$

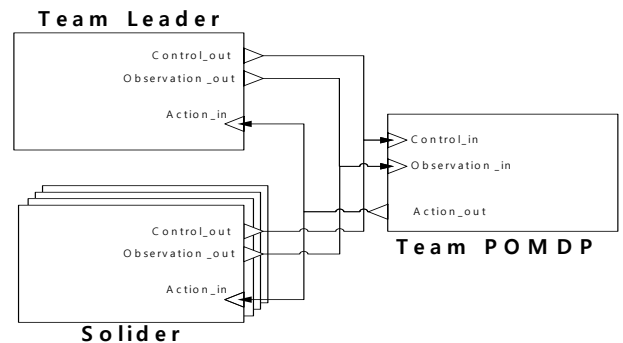


Figure 9. Coupling information between DEVS models(team Leader and soldiers) and POMDP model(Team POMDP)

<Figure 9>는 전투 개체 모델의 행동을 결정하기 위해 이산 사건 모델(team leader and soldier; DEVS model)과 행동 모델(team POMDP; POMDP model) 사이에서 이루어지는 메시지 교환 내용을 나타낸다. 이산 사건 모델은 행동 모델에게 행동 결정의 제한 사항(control\_out)과 관찰 값(observation\_out)의 내용을 보내 팀의 임무 내용과 팀 내 전투 개체 모델들의 상태를 전달하고, 행동 모델은 결정된 행동(action\_out)을 이산 사건 모델에게 전달한다. 본 연구에서는 전투 개체 모델의 행동 중, 청군 팀 모델의 기동 행동을 결정하는데 <Figure 9>와 같은 방법을 적용하였다. 팀의 기동은 작전 내용과 군사 교리에 기동 대

형이 결정되고, 팀 구성원들은 결정된 기동 대형을 따라 기동하게 된다. 또한, 팀 기동 중에도, 팀 모델 내부의 전투 개체들은 현재 전장 상황(작전, 군사 교리, 탐지된 청군과 홍군의 위치 등)을 고려하여 기동 방법(기동 방향 및 기동 속도)을 결정한다. 본 연구의 전투 개체 모델의 이산 사건 모델과 행동 모델이 교환하는 메시지의 내용은 <Table 5>의 내용과 같다. <Table 5>는 이산 사건 모델과 행위 모델 사이에서 교환되는 메시지의 내용이 교환되는 DEVS 모델의 이벤트와 교환되는 메시지가 영향을 주는 POMDP 모델의 상태 변수의 매핑 관계로 설명하였다.

**Table 5.** Relations between variables and model specifications in DEVS and POMDP

Variables	Message	Inputs or Outputs in DEVS	Components in POMDP	Implications
Current point	Control	<i>Mnv_out</i> in Maneuver	A component of state, $\vec{\alpha}$	Current point of the combat-entity model
Maximum maneuver speed		<i>Control_out</i> in Maneuver	Involved in policy computation step	Maximum maneuver speed of the combat-entity model
Objective point		<i>Mnv_out</i> in Decision Making	A component of state, $\vec{g}$	Mission maneuver point of the combat-entity model
Maneuver objectives		<i>Mnv_out</i> in Decision Making	Involved in reward function	Mission objectives of the combat-entity model
Maneuver formation		<i>Mnv_out</i> in Decision Making	Involved in reward function	Maneuver formation of the team which the combat-entity model joined to
Detected information of the blue		<i>Det_out</i> in Detection	Involved in reward function	Detected information of the blue in a specific time of the running simulation
Detected information of the red		<i>Det_out</i> in Detection	A component of state, $\vec{\beta}$	Detected information of the red in a specific time of the running simulation
Detected information of the blue	Observation	<i>Det_out</i> in Detection	Not Used	Current point of the combat-entity model
Detected information of the red		<i>Det_out</i> in Detection	A component of observation, $\vec{\beta}$	Detected information of the blue in a specific time of the running simulation
Maximum maneuver speed		<i>Control_out</i> in Maneuver	Involved in policy computation step	Maximum maneuver speed of the combat-entity model
Current point	Action	<i>Mnv_out</i> in Maneuver	A component of observation, $\vec{\alpha}$	Current point of the combat-entity model
Maneuver direction		<i>Action_in</i> in Maneuver	A component of action, $\vec{\phi}$	Maneuver direction of the combat-entity model to maintain maneuver formation of the team
Maximum maneuver speed			A component of action, $\vec{r}$	Maximum maneuver speed of the combat-entity model

이를 모의하기 위해, 팀 모델의 기동은 다음과 같은 과정을 통해서 이루어진다. 시나리오를 따라 분대와 팀의 작전이 정해지면, 작전의 행동 방침이 리더 모델에게 전달된다. 리더 모델은 기동의 목적 및 대형, 기동 속도, 리더의 위치를 포함한 시야 내 청군 및 홍군의 위치(control)를, 그리고 팀 내의 사병 모델들은 기동 속도, 자신의 위치를 포함한 시야 내 청군 및 홍군의 위치(observation)를 POMDP 모델에게 전달한다. 입력 받은 정보를 이용하여, POMDP 모델은 전투 개체들이 대형을 맞추어 작전 기동을 할 수 있도록 각 전투 개체들의 기동 방향과 속도(action)를 결정하여 전달한다. 전투 개체들은 결정된 기동 정보를 통해서 팀 구성원들이 이동하고, 주기적으로 observation 정보를 POMDP 모델에게 전달한다. POMDP 모델은 입력 받은 observation 정보를 이용하여 현재 전장 상황에 맞는 팀 구성원의 기동 방향과 속도를 결정하여 해당 모델들에게 보내주는 과정이 반복된다. 이러한 과정은 리더 및 일반 사병이 작전 목적 위치에 도달하여 리더 모델이 control 메시지를 통해 POMDP의 행동 결정을 중지할 때까지 계속된다.

## 5. 전투 실험

전투 개체 모델을 이산 사건 모델과 행위 모델로 분리한 모델 구조가 전투 효과도(combat effectiveness)에 미치는 영향을 파악하기 위해 아래의 전투 실험을 수행했다. 특히 본 연구에서는 POMDP를 사용한 자율 기동 모델이 기존의 DEVS만을 사용한 규칙 기반 모델에 비하여 보다 쉽게 복합적인 목표가 고려될 수 있음을 보이고자 한다. 여기서 복합적인 목표란,

1. 주어진 목표 지점까지 이동
2. 이동시 대형 유지
3. 우회 기동 시, 적군에게 발각을 최소화
4. 임무 수행의 효율성 증가

를 동시에 만족시키는 것이다. 이러한 목표들을 POMDP 모델의 보상 함수(reward function)로 표현하고 최적화 알고리즘을 사용하여 각 전투 개체의 효율적인 기동을 결정한다. 따라서 본 전투 실험에서는 기술적 모델링의 DEVS 모델과 규정적 모델링의 POMDP 모델을 조합하여 개발한 전투 개체 모델(POMDP-DEVS 모델)이 DEVS 모델로 개발한 전투 개체 모델(DEVS 모델)에 비하여, 주어진 복합적인 목표를 더 잘 달성할 수 있음을 실험적으로 보이고자 한다.

### 5.1 전투 실험 계획

본 연구에서 수행하는 전투 실험은 제 4장의 가상군 모델을 이용하여 수행한다. 특히, <Figure 3>의 구조를 갖는 POMDP-DEVS 모델과 <Figure 3>의 구조에서 POMDP 모델이 존재하

지 않는 DEVS 모델의 두 가지 가상군 모델을 실험에 사용한다. 가상군 모델 내부의 전투 개체 모델(리더 혹은 사병)의 기동은 전투 교리에 의해서 다수(부대 혹은 팀)가 일정한 대형을 이루며 움직여야 한다. 두 모델이 복합적인 목표를 얼마나 충실히 수행하는지를 수치적으로 확인하기 위해 두 가지 성능 지수를 정의한다. 복합적인 목표 중, 1~3 목표의 달성도를 측정하기 위해서, 대형 위반 불이익(formation violation penalty; fvp)이라는 성능 지수를 정의하였다. 전투 개체 모델이 기동하는 지역의 지형, 고도 등과 같은 전장 환경의 요인은 다수의 전투 개체가 대형을 유지하며 기동하기 어렵게 한다. 따라서 본 전투 실험에서는 먼저 두 가상군 모델의 효율적인 기동의 정도를 파악하기 위해서 대형 위반 불이익을 아래 식 (5)와 같이 정의하였다.

$$\text{대형 위반 불이익} = E[fvp] \quad (5)$$

where  $fvp = fvp_d + fvp_\theta$ ,

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

$$fvp_\theta = (\theta_{12} - \theta)^2 + (\theta_{14} - \theta)^2 + (\theta_{13} - \theta)^2 + (\theta_{15} - \theta)^2$$

대형 위반 불이익을 측정할 때 쓰이는 변수들  $d_{ij}, \theta_{ij}$  등은 제 4.2.3절의 <Figure 8>에서 정의되었다. 대형 분산  $d$ 와 대형 각도  $\theta$ 는 대형의 모양을 결정짓는 값으로  $d$ 는 대형을 이루는 사병 사이의 거리를 의미하고,  $d$  값이 클수록 팀의 기동 대형은 큰 대형을 유지하게 된다.  $\theta$ 는 대형을 이루는 팀의 리더와 사병들이 대형에 맞추어 기동할 때 주위의 팀원들과 벌어진 각도를 의미하며, 대형 각도의 값이 클수록 가로 대형에 가깝고 작을수록 세로 대형에 가깝게 된다.

대형 위반 불이익은 팀 모델의 기동 대형 유지 정도를 나타내고 있지만, 기동 대형을 유지하여 기동하는 것이 임무의 달성이나 전투 효과도(combat effectiveness)에 미치는 영향력을 분석하는 것은 중요한 일이다. 즉, 복합적인 목표 중 4번에 대한 결과를 측정하고 분석하는 것이 필요하다. 따라서 본 연구에서는 전투 효과도(combat effectiveness) 지표로 자주 사용되는 피해율(damage ratio)을 성능 지수로 사용하였다(식 (6)).

$$\text{피해율} = \frac{\text{초기분대 전력} - \text{종료 후분대 전력}}{\text{초기분대 전력}} \quad (6)$$

청군의 피해율은 청군의 임무 목표 지점을 탈취하는 과정에서 목표 지점을 선점하고 있는 홍군과 발생하는 교전의 결과를 반영하여, 교전 중에 입은 청군의 피해를 초기 전력과 비교하여 비율로 나타낸다. 청군의 피해율을 통해 규정적 모델링의 행동 결정이 실제 전투 효과도에 어떻게 반영되는지 파악할 수 있다. 위의 두 가지 성능 지수를 측정하기 위한 전투 실험을 수행하기 위해, <Table 6>과 같은 전투 실험 계획을 수립하였다. 전투 실험은 청군의 분대 1팀과 분대 2팀이 목표 지점까지 기동을 한 후, 분대 1팀이 홍군의 경계진지로 돌격하여

Table 6. Experimental design of the battle experiments

Variables(unit)	Variation Range	Implications
Dispersions in the formation, $d(m)$	1.1, 1.65, 2.2(3 cases)	Distance between two combat-entity models of the maneuver formation
Angles in the formation, $\theta(\text{degree})$	120, 135, 150(3 cases)	Angles between team members in the maneuver formation
Maneuver Behavior Model	DEVS model, or DEVS-POMDP model(2 case)	Behavior model describing maneuver of combat-entity models
total cases	$3 \times 3 \times 2 = 18 \text{ cells}$	Each cell is replicated for 30 times

탈취할 때까지 진행되었다. 전투 실험에서 청군 분대 팀들의 기동 대형은 삼각대 대형( $\Delta$ )으로 기동하고, 대형의 각도( $d$ )와 분산( $\theta$ )을 <Table 6>의 변경 범위에 따라 변화시키면서 실험을 수행하였다. 변수의 변화를 POMDP-DEVS 모델과 DEVS 모델에 각각 적용하여, 총 18개의 실험 케이스에 대한 실험을 진행하였다. 각 실험 케이스는 30회씩 반복하여 실험 결과를 통계적으로 분석하였다.

위 전투 실험의 경우 POMDP-DEVS의 경우 180회 시뮬레이션 실행이 필요하며, DEVS만 활용하였을 경우 역시 180회의 실행이 필요하다. POMDP의 경우 기동 행동 설정을 위한 연산 작업이 필요하므로 DEVS보다 시간이 더 많이 걸리게 된다. 실제 실험 결과 현실상의 시간(Wall Time Clock)으로는 POMDP-DEVS의 경우 1회 실행 당 평균 2분 40초, DEVS의 경우 평균 25초의 시간이 소요된다.

또한 DEVS만을 활용하여 전투 실험 결과를 도출하는 과정에서, DEVS의 기동 모델은 대형을 생각하지 않은 모델이다. 현재 실험에서 DEVS 기동 모델은 특정 좌표에서 전 전투 개체가 특정한 다음 좌표까지 직선 이동을 하면서 다른 개체와의 평균 거리가 일정거리 이상 벌어질 경우, 잠시 기동을 멈추는 형태로 모델링 되어 있다. 이는 군집으로서 전투개체의 이동은 모델링이 가능하나, 자세한 대형유지 수준의 모델링은 아니다. 현재 실험이 전투 대형의 유지가 얼마나 전투 실험 결과에 영향을 미치며, POMDP-DEVS가 어떻게 전투 대형 유지를 실현할 수 있는지가 목적이므로, DEVS를 활용한 자세한 기동 모델은 생략하였다.

5.2 전투 실험 결과 및 분석

이번 절에서는 제 5.1절의 실험계획을 따라 수행한 실험 결과에 대해서 설명한다. 제 5.1절에서 정의한 성능 지수를 분석하기 전에, 가상군 DEVS 모델과 POMDP-DEVS 모델의 기동 대형 유지 정도를 전투 실험의 화면 캡처를 이용하여 확인한다. <Figure 10>의 화면 캡처들은 전투 실험에서 발생한 전투 개체의 위치를 로그(log) 형태로 출력하여, KAIST 산업 및 시스템 공학과의 SESLab에서 개발한 가시화 도구를 이용하여 재현한 내용의 일부를 캡처 하였다.

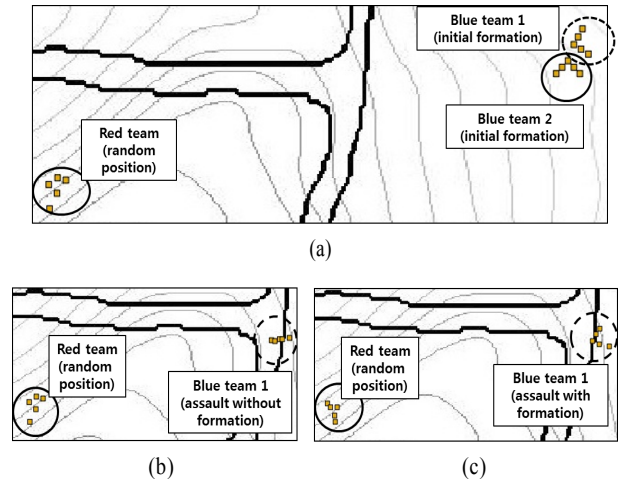


Figure 10. Screen captures of the battle experiments ( $d : 2.2m, \Theta = 135$ ) : (a) Initial position of red and blue teams, assault of blue team 1 in (b) DEVS model and (c) POMDP-DEVS model

<Figure 10>(a)는 시뮬레이션이 시작 전의 전투 개체들의 초기 대형을 보여 준다. 홍군 경계팀(red team)은 화면 왼쪽 아래에 랜덤하게 배치되고, 청군의 분대 1팀(blue team1)과 분대 2팀(blue team2)은 기동할 방향을 바라보며 삼각대 대형으로 배치되어 있다. <Figure 10>(b)와 <Figure 10>(c)는 시뮬레이션이 진행되어 홍군 경계팀에게 접근하는 청군 분대 1팀의 기동 대형을 보여준다. <Figure 10>(b)는 가상군 DEVS 모델의 경우로써, 시뮬레이션이 진행됨에 따라서 삼각대 대형이 무너지는 것을 확인할 수 있다. 반면에 <Figure 10>(c)는 가상군 POMDP-DEVS 모델의 경우로써, 삼각대 대형을 유지하면서 기동 중인 것을 확인할 수 있다. <Figure 10>을 통해서, POMDP-DEVS 모델이 팀 모델의 기동 대형을 유지하는데 더 유리하게 동작하고 있음을 확인할 수 있다.

본 시뮬레이션이 모델링 요건을 만족하였는지 검증(Verification)하기 위하여, 시뮬레이션 시각화를 통한 실행 과정의 검증을 진행하였다. 본 검증 방법은 시뮬레이션의 상태 천이 수준의 검증이 아닌 구현된 전투 개체가 모델링 목표를 정확히 표현하고 있는지 시뮬레이션 전문가 및 분야 전문가가 관측하



고 검증하는 방식이다. <Figure 10>에서 볼 수 있듯이, 대열을 상세히 모델링한 POMDP-DEVS의 경우는 V자형 대형을 잘 유지하여 이동하는 것을 볼 수 있다. 또한 반복 전투 실험 결과 역시 분야 전문가의 납득이 불가능한 수준의 결과가 도출되지는 않았다.

### (1) 대형 위반 불이익에 대한 실험 결과 분석

전투 실험에서 가상군 모델 내부의 청군 분대의 팀들이 기동을 유지하며 기동하는 정도를 수치적으로 확인하기 위해서, 각 실험 케이스에 대한 대형 위반 불이익의 결과를 <Table 7>에 정리하였다. 대형 위반 불이익은 홍군 경계 진지로 직진 기동을 하는 분대 1팀(blue team 1)과 차단 진지로 우회 기동하는 분대 2팀(blue team 2)에 대해서 독립적으로 측정하였다.

**Table 7.** Experimental results for formation violation penalty

$(d, \theta)$	Blue Team #	Avg. Formation Violation Penalty	
		DEVS Model (95% conf. interval)	POMDP-DEVS Model(95% conf. interval)
(1.1m, 120°)	1	21.13(±1.12)	2.04(±0.29)
	2	22.66(±2.40)	1.54(±0.11)
(1.1m, 135°)	1	21.11(±1.11)	2.71(±0.33)
	2	16.50(±1.93)	2.73(±0.33)
(1.1m, 150°)	1	19.13(±1.12)	3.53(±0.38)
	2	13.55(±1.84)	3.51(±0.45)
(1.65m, 120°)	1	21.41(±1.27)	1.94(±0.21)
	2	16.79(±2.67)	2.00(±0.14)
(1.65m, 135°)	1	19.87(±1.56)	2.39(±0.23)
	2	15.79(±1.80)	3.12(±0.25)
(1.65m, 150°)	1	18.60(±1.27)	3.48(±0.31)
	2	11.67(±1.48)	4.49(±0.28)
(2.2m, 120°)	1	22.13(±1.17)	1.77(±0.15)
	2	15.28(±2.91)	2.06(±0.16)
(2.2m, 135°)	1	19.98(±1.54)	2.38(±0.22)
	2	11.95(±1.84)	3.26(±0.14)
(2.2m, 150°)	1	18.90(±1.29)	3.27(±0.35)
	2	12.29(±1.16)	4.70(±0.32)

<Table 7>에서 가상군 DEVS 모델의 경우 대형 위반 불이익 값이 가상군 POMDP-DEVS 모델에 비해 평균적으로 3~15배 이상 더 크게 나옴을 볼 수 있다. 이는 가상군 DEVS 모델은 지형 속성 및 고도 차이 등의 주변 환경의 기동 불확실성에 크게 영향을 받음으로서 대형 유지에 어려움을 겪는 반면에 가상군 POMDP-DEVS 모델은 주변 환경의 기동 불확실성에도 강인하게 스스로 대형을 맞추어 감을 나타낸다. 또한 이러한 수치는 95% 신뢰구간을 고려하였을 때 통계적으로 유의미함을 볼 수

있다. 가상군 DEVS 모델의 경우 분대 1팀의 대형 위반 불이익 값이 분대 2팀의 대형 위반 불이익 값보다 대부분 더 큼을 볼 수 있다. 이는 분대 1팀의 기동 수행 시간이 분대 2팀의 기동 수행 시간보다 길어짐으로 인해 시간에 따라 대형 위반 불이익 값이 커지기 때문이다. 반면에 가상군 POMDP-DEVS 모델에서는 분대 1팀의 대형 위반 불이익 값과 분대 2팀의 대형 위반 불이익 값이 크게 차이가 없음을 알 수 있는데, 이는 기동 수행 시간이 길어짐에도 스스로 대형을 유지하기 때문에 기동 중에 대형 위반 불이익 값이 일정하게 유지되기 때문으로 볼 수 있다.

### (2) 전투 효과도에 대한 실험 결과 분석

청군의 피해는 청군의 분대 1팀이 홍군 경계 진지를 탈취하는 과정에서 홍군 경계팀과의 교전에서 발생한다. 특히, 청군 분대 1팀이 홍군 경계팀의 전투 개체들을 셋 이상 사살하면(즉, 전력이 30% 이하로 떨어지면), 홍군 경계팀은 정해진 후퇴 지점으로 퇴각을 시작한다. 퇴각이 시작되면, 청군 분대 1팀은 적군을 쫓지 않고 경계 진지를 탈취 한다. 전투 실험의 결과로 나온 청군과 홍군의 피해율에 대한 대형 분산( $d$ )과 각도( $\theta$ )의 영향을 파악하기 위해, 각 변수에 대해서 주변화(marginalization)하여 분석하였다. <Figure 11>은 청군(<Figure 11>(a))과 홍군(<Figure 11>(b))의 피해율을 대형 분산에 대해서 주변화한 그래프이고, <Figure 12>는 청군(<Figure 12>(a))과 홍군(<Figure 12>(b))의 피해율을 대형 각도에 대해서 주변화한 그래프이다. <Figure 11>과 <Figure 12>의 결과 POMDP-DEVS 모델을 사용해서 팀 모델의 대형을 유지한 경우와 DEVS 모델만 사용하여 대형을 무시한 경우에 비해서 청군과 홍군의 피해율이 3~5% 정도 차이 나는 것을 확인하였다. 위의 결과를 통해, 전투 실험 시 전투 개체 모델의 대형 유지 유무가 전투 효과도에도 영향을 미칠 수 있음을 확인할 수 있다.

### (3) 메타 모델을 통한 실험 결과 분석

대형 유지에 따른 피해율의 변화를 대형 분산과 대형 각도에 대해서 자세히 분석하기 위해, 선형 회귀 분석(linear regression analysis)을 수행하였다. 선형 회귀 분석은 POMDP-DEVS 모델과 DEVS 모델에 대해서 각각 수행하였으며, 성능 지수인 청군과 홍군의 피해율에 대해서 대형 각도와 분산이 미치는 영향의 정도와 유의성을 파악해 보기 위해 수행하였다. 선형 회귀 분석 결과의 결과로 얻은 표준화 계수(standardized coefficient)를 <Table 8>에 정리하였다.

<Table 8>의 POMDP-DEVS 모델과 DEVS 모델에 대한 표준화 계수 결과를 통해, 다음의 분석 결과를 얻을 수 있었다. 첫째로, 두 모델의 경우에서 청군의 피해율에 대한 표준화 계수의 유의 확률(p-value)이 홍군의 피해율보다는 상대적으로 낮은 것을 확인할 수 있다. 즉, 청군의 피해율에 대한 대형 분산과 각도의 표준화 계수가 좀 더 의미 있는 것으로 해석할 수 있고, 이는 홍군의 피해율 보다는 청군의 피해율이 대형 분산과

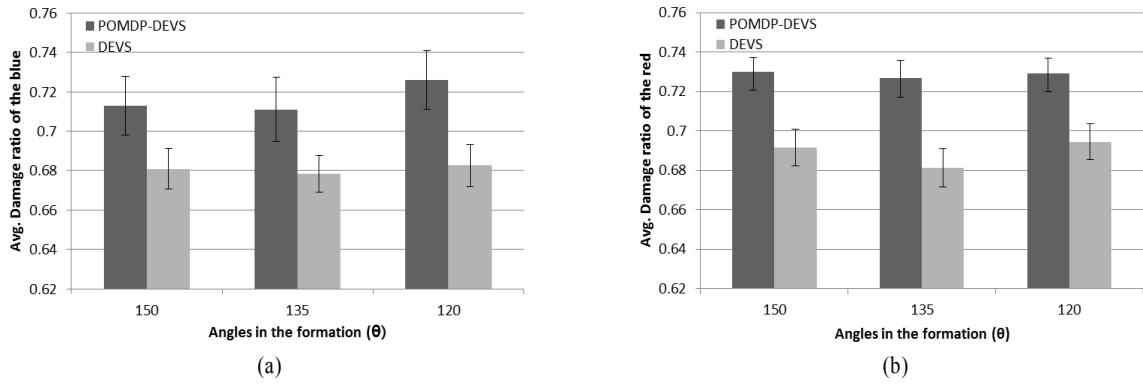


Figure 11. Average damage ratio marginalized over dispersions : (a) blue squad (b) red squad

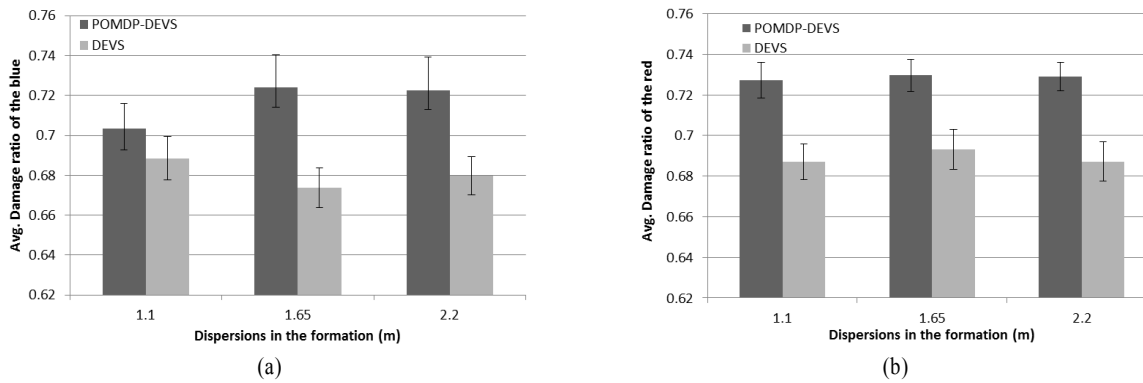


Figure 12. Average damage ratio marginalized over angles : (a) blue squad (b) red squad

Table 8. Standardized coefficients from the linear regression analysis

Behavior model	POMDP-DEVS model		DEVS model	
	Damage ratio of the blue(p-value)	Damage ratio of the red(p-value)	Damage ratio of the blue(p-value)	Damage ratio of the red(p-value)
Dispersions in the formation(d)	0.107(0.080)	0.17(0.777)	-0.74(0.226)	0.01(0.984)
Angles in the formation( $\Theta$ )	0.072(0.235)	-0.11(0.859)	0.15(0.806)	0.027(0.664)
Adj. R-square	0.17	0.00	0.06	0.01

각도에 대해 상대적으로 영향을 받는 것으로 해석할 수 있다. 반면에, 홍군의 피해율에 대한 표준화 계수의 유의 확률 값이 크기 때문에, 홍군의 피해율에 대한 표준화 계수는 해석의 의미가 크지 않다는 것을 의미한다. 따라서 대형 분산과 각도의 차이는 홍군의 피해율과 크게 상관이 없다는 것으로 해석할 수 있다. 둘째로, POMDP-DEVS 모델의 청군의 피해율에 대한 대형 분산의 표준화 계수는 다른 표준화 계수 보다 낮은 유의 확률(0.080)을 갖는 것을 확인할 수 있다. 이는 청군의 피해율에 대형 분산이 미치는 영향이 다른 변수 보다 좀 더 의미가 있다고 해석할 수 있고, 표준화 계수의 값(0.107)으로 대형 분산의 크기가 커질수록 청군의 피해율이 증가하는 것을 파악할 수 있다. 이 분석 결과를 통해, 대형 분산의 크기가 커질수록, 전투 개체 간의 거리가 넓어져 직사화기를 이용한 전투에서는 청군에게 불리하다는 것을 파악할 수 있다.

## 6. 결론

기존의 전투 모의는 대부분 기술적 모델링에 입각하여 전투 행동을 교범에 맞추어 기술하는데 초점이 맞추어져 있었다. 이와 같은 모델링의 장점은 교범에 정의된 행동들을 명시적으로 모델링하여, 누구나 개별 모델들의 행동을 쉽게 예측할 수 있다는 장점이 있었다. 그러나 기술적 모델링만을 활용할 경우에는 규범에 명확히 드러나지 않은 인간의 이성적 행동, 혹은 규범에 나타나 있어도 명시적으로 모델링하기 힘들기 때문에 매우 추상화하여 모델링하였다. 이런 추상화된 결과가 실제 전투 상황시의 전투 개체가 보이는 행동과 전투 모의에서의 행동 결과에 차이를 가져오게 되며, 결과적으로 전투 모의를 활용하는 주된 목표인 전투 효과도 예측 등에 부정확성을 가져오게 된다. 그러므로 본 논문은 교범을 활용한 기술적 모델

링의 하위 모델링으로 인간의 이성적 행동을 모델링하는 규정적 모델링을 추가하여 확장하는 전투 모의 사례를 소개하였다. 전투 모의를 활용한 실험 결과에 따르면, 전투 행동 중, 인간의 이성적 행동을 기술적 모델링만으로 모의하는 방식과 규정적 모델링을 추가적으로 적용하는 방식의 실험 결과는 유의미한 차이가 있다는 점을 확인할 수 있었다. 이는 행동 모델링을 더욱 현실감 있게 모델링 할 수 있는 규정적 모델링을 기술적 모델링과 함께 사용해야만, 더욱 현실감 있는 전투 모의와 전투 실험 결과를 획득할 수 있을 것이라는 가설을 보장해준다. 결론적으로, 기존의 기술적 모델링만을 활용하는 전투 모의에서 벗어나 규정적 모델링을 복합하여 사용하는 차세대 전투 개체 기술이 필요하다는 것을 강조하며 본 논문을 맺고자 한다.

## 참고문헌

- Bertsekas, D. P. and Castanon, D. A. (1999), Rollout algorithms for stochastic scheduling problems, *Journal of Heuristics*, **5**(1), 89-108.
- Bertsekas, D. P. and Tsitsiklis, J. N. (1996), *Neuro-Dynamic Programming*, Athena Scientific, Belmont, Mass, USA.
- Bryant, D. J., Lichacz, F. M. J., Hollands, J. G., and Baranski, J. V. (2004), Chapter 6. Modeling situation awareness in an organizational context : military command and control, *A Cognitive Approach To Situation Awareness : Theory And Application*, Ashgate Publishing, 104-116.
- Chong, E. K. P., Givan, R. L., and Chang, H. S. (2000), A framework for simulation-based network control via hindsight optimization, *In Proceedings of the 39th IEEE Conference on Decision and Control (CDC)*, **2**, 1433-1438.
- Cil, I. and Mala, M. (2010), A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare, *Expert Systems with Applications*, **37**(2), 1331-1343.
- Doshi-Velez, F. (2009), The Infinite Partially Observable Markov Decision Process, In Proceedings of NIPS.
- Headquarters, department of the army (1962), *Department of the army field manual : Rifle company*, infantry and airborne battle groups, department of the army.
- Kim, J., Choi, C. B., Moon, I. C., and Kim, T. G. (2010), DEVS-based doctrine validation of fleet anti-air defense, Proceedings of the 2010 Spring Simulation Multiconference 28 : 1-28 : 8, San Diego, CA, USA : Society for Computer Simulation International.
- Kim, D., Lee, J., Choi, J., and Kim, K. (2012), A POMDP Framework for Dynamic Task Allocation and Reconnaissance of Multiple Unmanned Aerial Vehicles, *Journal of KIISE : Software and Applications*, **39**(6).
- Kim, J., Moon, I. C., and Kim, T. G. (2011), New insight into doctrine via simulation interoperation of heterogeneous levels of models in battle experimentation, *SIMULATION : Transactions of The Society for Modeling and Simulation International*, **88**(6), 649-667.
- Kim, T. G. and Moon, I. C. (2012), Combat Modeling using the DEVS Formalism, In Andreas Tolk(Eds.), *Engineering Principles of Combat Modeling and Distributed Simulation*, Wiley, 479-510.
- Kim, T. G., Sung, C. H., Hong, S. Y., Hong, J. H., Choi, C. B., Kim, J. H., Seo, K. M., and Bae, J. W. (2011), DEVSim++ Toolset for Defense Modeling and Simulation and Interoperation, *The Journal of Defense Modeling and Simulation : Applications, Methodology, Technology*, **8**(3), 129-142.
- Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987), SOAR : An architecture for general intelligence, *Artificial Intelligence*, **33**(1), 1-64. doi:http://dx.doi.org/10.1016/0004-3702(87)90050-6.
- Lee, K., Lim, H., and Kim, K. (2012), A Case Study of Modeling Computer Generated Forces based on Factored POMDPs, *In Proc. of Korea Computer Congress(KCC)*, **39**(1B).
- Miller, S. A., Harris, Z. A., and Chong, E. K. P. (2009), A POMDP Framework for Coordinated Guidance of Autonomous UAVs for Multitarget Tracking, *EURASIP Journal on Advances in Signal Processing*, Volume 2009, Article ID 724597, 17 pages.
- Moon, I. C., Carley, K. M., and Kim, T. G. (2013), Chapter 3. Modeling and Simulation Command and Control for Naval Air Defense Operation. *Modeling and Simulating Command and Control*, 35-66, SpringerBriefs in Computer Science.
- NATO Standardization Agency (2010), *NATO Glossary of Terms and Definitions*, NSA.
- Pang, C. K., Le, C. V., Gan, O. P., Hudus, G., Middleton, M. B., and Lewis, F. L. (2012), Discrete Event Command and Control of Multiple Military Missions in Network Centric Warfare. the 8th International Conference on Intelligent Unmanned Systems(ICIUS) 74-79, SIM university, Singapore.
- Papadimitriou, C. and Tsitsiklis, J. N. (1987), The Complexity of Markov Decision Processes, *Mathematics of Operations Research*, **12**(3), 441-450.
- Pineau, J. Gordon, G., and Thrun, S. (2003), Point-based Value Iteration : An Anytime Algorithm for POMDPs, In Proceedings of IJCAI.
- Salas, E., Bowers, C. A., and Cannon-Bowers, J. A. (1995), Military Team Research : 10 Years of Progress, *Military Psychology*, **7**(2), 55-75. doi:10.1207/s15327876mp0702\_2.
- Schmitt, J. and Klein, G. (1999), A Recognitional Planning Model, 1999 Command and Control Research and Technology Symposium, Newport, Rhode Island : Naval War College.
- Sondik, E. J. (1971), The Optimal Control of Partially Observable Markov Process, Ph.D. thesis, Stanford University.
- Sung, C. and Kim, T. G. (2012), Collaborative Modeling Process for Development of Domain-Specific Discrete Event Simulation Systems, *IEEE Transactions on Systems, Man, and Cybernetics-Part C : Applications and Reviews*, **42**(4), 532-546.
- Zeigler, B. P., Praehofer, H., and Kim, T. G. (2000), *Theory of Modeling and Simulation*, *Simulation* (2nd ed.), Academic Press, doi: 10.1159/000074301.