

# 지능형 선박 항해 에이전트 개발을 위한 강화학습 환경 분석

† 박세길 · 오재용\* · 김혜진\*\*

\*,\*\*,† 선박해양플랜트연구소

## The Analysis of Reinforcement Learning Environment for Intelligent Ship Navigation Agents

† Se-Kil Park · Jae-Yong Oh\* · Hye-Jin Kim\*\*

\*,\*\*,† Korea Research Institute of Ships & Ocean Engineering (KRISO)

**요 약** : 본 논문에서는 복잡한 해상교통 환경 하에서도 해양 안전을 도모할 수 있는 강화학습 기반 지능형 선박 항해 에이전트 개발의 사전 단계로서 기존의 강화학습 환경을 분석하였다. 강화학습 기반 접근법은 선박 항해 에이전트 스스로가 복잡하고 동적인 해상교통 환경을 이해하고 주어진 목표를 달성할 수 있도록 도와주는데, 이를 위해서는 에이전트 자신을 제외한 모든 사항들이 정의되는 환경을 보다 정확하고 효과적으로 개발하는 것이 매우 중요하다. 실제 해상교통 환경은 학습 환경으로의 모델링 및 에이전트 학습의 난이도가 매우 높은 환경으로 학습 환경이 가질 수 있는 여러 속성들을 적절히 설정하여 선박 항해 에이전트의 활용 목적에 맞는 가성비 높은 환경을 구축하는 것이 바람직하다.

**핵심용어** : 선박항해, 에이전트, 강화학습, 인공지능, 자율운항선

### 1. 서 론

해상교통 환경의 복잡화와 자율운항선박 등 신개념 선박의 등장은 해양 안전을 위한 새로운 해상교통 분석, 운용 기술을 요구하고 있다. 특히 복잡한 해상교통 환경 하에서 안전을 위해 고려해야할 다양한 변수들을 구조적으로 정리하고 이를 상황에 맞춰 적용함으로써 해양 안전을 강화할 수 있는 방법이 필요하다. 그러나 기존의 논리 기반 접근 방법만으로는 해상교통 환경을 충분히 이해하고 안전을 저해하는 요인을 극복해 나가기에 한계가 존재한다. 따라서 객체 레벨에서 모델링 및 상호작용이 가능하도록 만드는 방법, 즉, 객체 스스로가 해상교통 환경과 주어진 목표를 이해하도록 유도하는 형태의 접근 방법이 필요하다. 이러한 관점에서 지능형 선박 항해 에이전트는 동적으로 변화하는 해양환경 하에서 자율적으로 주어진 목표를 달성할 수 있는 시스템을 의미하며, 복잡한 해상교통 환경 하에서 선박의 안전을 향상 시킬 수 있는 효과적인 방안으로 판단된다.

지능형 에이전트는 활용 분야에 따라 다양하게 정의될 수 있으나 일반적으로 자율적으로 환경과 상호작용하며 주어진 목표를 달성하도록 고안된 가상 또는 실제의 개체를 의미한다. 여기서 환경과의 상호작용은 센서를 통해 환경을 관측하거나 역으로 환경을 변화시키는 행동을 취하는 것을 포함하며, 환경은 동적인 상태를 가질 수 있다. Fig. 1은 지능형 에이전트 개발을 위한 여러 가지 방법들 중 강화학습을 통한 에이전트

학습 개요를 보여준다. 강화학습은 주어진 환경 하에서 수많은 시행착오를 거치면서 보상을 최대화하는 행동 혹은 행동 순서를 배우는 방법으로 보상을 포함하는 환경을 어떻게 정의하고 에이전트를 학습시키는지 에이전트의 성능과 실제세계의 적용 가능성에 큰 영향을 준다.

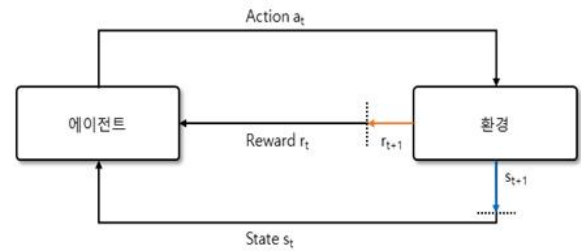


Fig. 1 강화학습을 통한 에이전트 학습 개요

본 논문에서는 강화학습 기반의 지능형 선박 항해 에이전트 개발을 위해 기존의 강화학습 환경들을 분석하고 이로부터 선박 항해 에이전트를 위한 강화학습 환경의 개발 방향에 대해 논의하고자 한다.

### 2. 강화학습 환경 분석

강화학습 환경은 환경이 제시하는 목표를 에이전트가 학습을 통해 풀어나간다는 관점에서 일종의 문제 집합이라고 할

수 있다. 이러한 문제들은 실세계에서 풀고자 하는 문제를 모델링한 것은 물론 실세계와 무관한 문제들도 존재한다. 그리고 쉬운 문제부터 어려운 문제까지 다양한 난이도를 가진다. Fig. 2는 OpenAI, DeepMind 등에서 제공하는 다양한 형태의 강화학습 환경 예시들을 보여준다(Minh, 2013, Brockman, 2016, Shah, 2018). 기존의 게임이나 시뮬레이션 등도 직접 코드를 수정하거나 OpenAI Universe 등을 활용하여 강화학습 환경으로 이용하고 있으며, 점차 적용 분야별 특성을 정확하게 반영하기 위해 고유 목적의 학습 환경을 제작하여 활용하는 사례가 많아지고 있다.

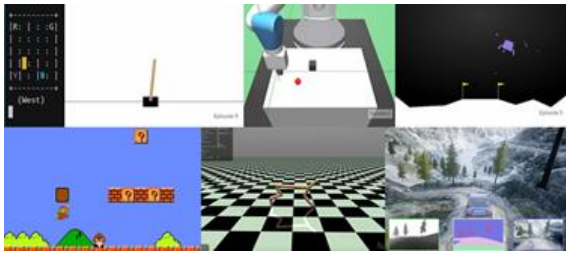


Fig. 2 강화학습 환경 예시

기존의 강화학습 환경들은 에이전트의 학습 내용에 따라 시각적 사실성이 중요한 환경, 물리적 사실성이 중요한 환경, 그리고 이러한 사실적인 표현보다는 인지적 표현이 보다 중요한 환경 등으로 분류 가능하며, 기술의 적용 범위가 넓어짐에 따라 점차 복합적 특성을 가지는 환경이 요구되고 있다.

강화학습 환경은 또한 환경이 가지는 여러 속성에 따라 다음의 여러 형태로도 분류할 수 있다.

Table 1 강화학습 환경의 분류

환경의 형태		분류 기준
확률론적 환경	결정론적 환경	결과의 불확실성 여부
완벽하게 관측 가능한 환경	부분 관측 가능한 환경	모든 시간에서의 시스템 상태 파악 여부
이산적인 환경	연속적인 환경	상태 변화 행동의 유한성 여부
에피소드/비순차 환경	비에피소드/순차 환경	현재 행동과 미래 행동 사이의 독립 여부
단일 에이전트 환경	다중 에이전트 환경	에이전트의 수가 다중인지 여부
정적 환경	동적 환경	시간에 따른 환경의 변화 여부

### 3. 해상교통 학습 환경

해상교통 학습 환경은 선박 항해 에이전트의 학습을 위해 해역, 선박, 항로표지 등 여러 객체들에 대한 형상은 물론 각 객체들이 가지는 다양한 속성들도 사실적으로 재현할 필요가 있

다. 또한 바람, 파도, 날씨, 가지거리, 시간 등 여러 환경 요소들의 영향과 GPS, 카메라, 레이더, 라이다 등 센서들의 결과역시 재현할 필요가 있다. 그러나 실제 해상교통 환경이 가지는 모든 요소를 모두 사실적으로 재현하는 것은 많은 시간과 비용이 소요되고 매우 비효율적인 방법이다. 따라서 해상교통 학습 환경을 구축하기 전에 선박 항해 에이전트의 핵심 기능이 무엇이며 이 핵심 기능을 달성하기 위해 선박 항해 에이전트를 어떠한 해상교통 학습 환경에서 어떠한 보상 체계와 상호 작용을 통해 학습을 시키는 것이 효과적일지에 대한 검토가 매우 중요하다. 더불어 선박 항해 전문가의 궤적을 가장 잘 설명하는 보상을 찾아내기 위해 역 강화학습을 적용하는 등 학습 환경 개발 후 이를 효과적으로 활용하는 방안에 대한 고민도 함께 필요하다.

## 4. 결 론

에이전트 관점에서 실제 해상교통 환경은 불확실성이 크고, 정보가 제한되며, 연속적이고 동적인 환경에서 서로 다른 운동 특성을 가지는 다중 객체가 활동하는 난이도 높은 환경이다. 따라서 지능형 선박 항해 에이전트를 위한 해상교통 학습 환경은 먼저 에이전트가 선박 항해 에이전트로서 갖춰야 할 요소들을 구체화하고 이를 바탕으로 학습 환경이 가질 수 있는 여러 속성들을 적절히 설정함으로써 목적에 맞는 가성비 높은 환경을 구축하는 것이 바람직하다.

## 사 사

본 논문은 선박해양플랜트연구소의 주요사업인 “해상교통 분석을 위한 에이전트 모델링 및 연동 기술 개발(1/5)”에 의해 수행되었습니다(PES3120).

## 참 고 문 헌

- [1] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602
- [2] Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J. and Zaremba, W. (2016). Openai gym. arXiv preprint arXiv:1606.01540
- [3] Shah, S., Dey, D., Lovett, C. and Kapoor, A. (2018). Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In Field and service robotics (pp. 621-635). Springer, Cham.