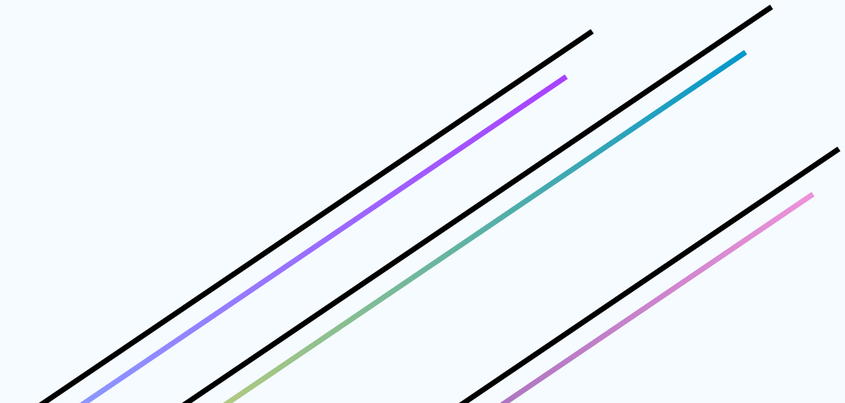
2025 SMARCLE

Reinforcement learning week 2

Contents

1. 이론 복습
2. 주제 발표 세미나
3. 코드 실습
4. 다음 스터디 예고



이론 복습

PPO: 딥레이서 모델을 훈련시키는 핵심 강화학습 알고리즘으로, 특정 행동의 가치 추정보다 정책 자체를 최적화

DQN(가치 반복법) vs PPO(정책 반복법)

- DQN: 각 행동의 예상 보상(Q-value)을 학습해 가장 가치가 높은 행동을 선택하는 간접적 방식. 이산적 행동에 적합
- PPO: 상태별 행동의 확률 분포를 학습&개선하는 직접적 방식. 확률적으로 행동을 선택해 유연성이 높음

PPO가 딥레이서에 적합한 이유

- 조향각도, 속도 같은 연속 제어를 자연스럽게 표현
- 정책이 확률적이므로 별도 ε-greedy(엡실론 그리디) 없이도 탐험&활용 균형

PPO가 해결하는 근본 문제

- 클리핑으로 정책 변경 비율을 신뢰 영역 내로 제한해, 과도한 업데이트로 인한 성능 붕괴 방지하고 안정적 수렴 유도

딥레이서 학습 전체구조

- 상호작용: 정책 네트워크가 시뮬레이터와 상호작용하며 상태, 행동, 보상을 수집
- 데이터저장: 수집 데이터는 경험버퍼에 보관
- 업데이트: PPO가 버퍼에서 샘플링해 정책, 가치 네트워크를 동시 학습

요소1. 행동 공간 정의

- 개념: 에이전트가 선택할 수 있는 행동의 메뉴판 정의
- 주요설정: 최대 조향각도, 조향단계, 최대 속도&속도단계
- Trade-off: 행동 가짓수가 많을수록 정교하지만 탐색공간 폭증 → 학습 난이도 상승 → 초기엔 단순하게 시작

요소2. 환경설정(트랙선택)

- 학습 난이도와 일반화 성능을 좌우함
- 초급트랙: The 2019 DeepRacer Championship Cup -폭 넓고 완만, 기본 정책 학습에 적합
- 고급트랙: Circuit de Barcelona-Catalunya - 급커브&좁은길, 일반화 성능 검증에 유용

보상함수 설계(“어떤 행동이 좋은가?”를 코드로 가르치기)

보상함수의 재료 params 딕셔너리

- distance_from_center, all_wheels_on_track, speed, steering_angle, progress 등센서& 상태 정보 활용

주행 전략별 보상함수 설계

- 전략 1: Follow the Center Line
  - #중앙에 가까울수록 높은 보상. 기본적이고 안정적인 학습
  - if distance_from_center <= 0.1 * track_width: reward = 1.0
- 전략 2: Stay Inside the Borders
  - # 네 바퀴가 트랙 안이면 높은 보상
  - if awll_wheels_on_track and (0.5*track_width - distance_from_center) >= 0.05: reward = 1.0
- 전략 3: Prevent Zig-Zag
  - # 조향 각도 과하면 페널티
  - if abs(steering_angle) > 15.0: reward *= 0.8

테스트 분석

테스트 요약

- #1 4 m/s + CenterLine: 초급 2/5
- #2 2 m/s + CenterLine: 초급 5/5, 고급 0/5
- #3 2m/s + StayInsidetheBorders: 고급 3/5

분석

- 안정성: 속도가 높으면 급코너에서 이탈이 잦음
- 과적합: 중앙선 보상은 단순 트랙에 치우침
- 일반화: StayInsidetheBorders 기반 보상이 고급트랙에서 유리

테스트 환경

- PPO&Agent: 단일 카메라
- 트랙: 2019 ChampinshipCup(초급), Circuit de Barcelona-catalunya(고급)

결론

- 보상함수가 성능을 좌우함 → 보상&페널티 기준을 명확히 설계
- 속도, 행동공간, 하이퍼파라미터를 바꾸며 검증

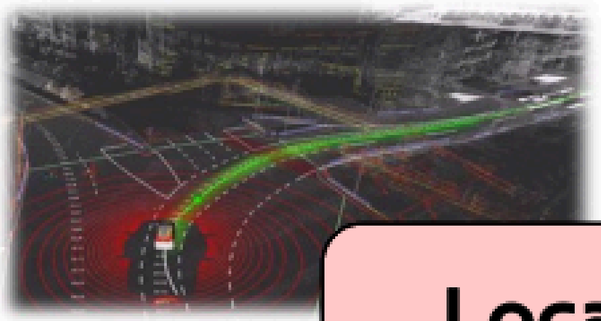
주제 발표 세미나

<자율주행과 강화학습>

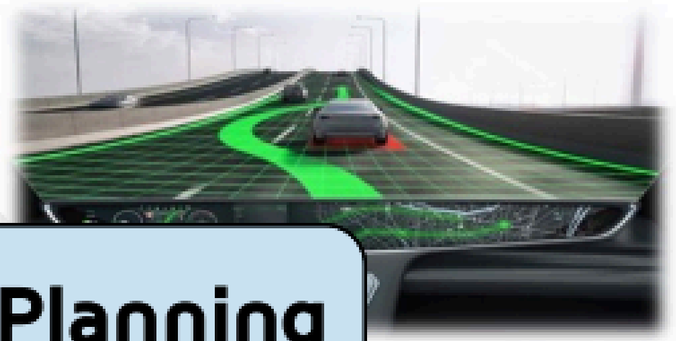
자율주행 시스템 개요

- **인지**
  - 차량의 외부 환경을 센서를 통해 인식하고 의미 있는 정보로 해석하는 과정
  - 객체 검출, 차선 인식, 도로 이해
- **측위**
  - 차량이 현재 위치를 정확하게 파악하는 과정
  - 지도와 센서 정보를 융합하여 위치 추정
- **판단**
  - 인지, 측위 결과를 바탕으로 최적의 경로와 행동을 결정
  - 전역 경로 계획, 행동 계획, 운행 계획
- **제어**
  - 차량의 조향, 가속, 제동을 실제로 수행하며 안정성을 보장
  - PID controller, Stanley controller, MPC 등

자율주행 시스템 개요



Ref.) Autoware



Ref.) Futurama

Localization
"Where are we?"

Motion Planning
"How will we drive?"

Perception
"What are around us?"

Motion Control
"How can we achieve the plan?"

Ref.) GTC Japan 2018 - NVIDIA

자율주행에서 강화학습의 역할

- 목표 : **보상을 최대화하는 방향으로 차량의 행동 정책을 학습**
- 장점
  - **규칙 설계 없이 학습 가능** (기존 rule-based는 모든 경우를 사람이 정의해야 하지만, RL은 보상 설계만으로 다양한 상황 학습 가능)
  - **복잡한 보상 구조 반영** (안전, 편안함, 도착 시간, 연비 등을 하나의 보상 함수로 통합)
  - **시뮬레이션 → 실제 차량 이전** 학습 가능
- 예시
  - 차량 행동 결정 (DQN, PPO)
  - 속도, 가속도 최적화 (Actor-Critic)
  - 차량 협력 주행 (Multi-Agent RL)
  - 비상 상황 대응 (Deep RL)

코드 실습

한국어 문의처 지원 내 계정

aws 에이전틱 AI | AWS 살펴보기 제품 솔루션 요금 리소스 콘솔 로그인 계정 생성

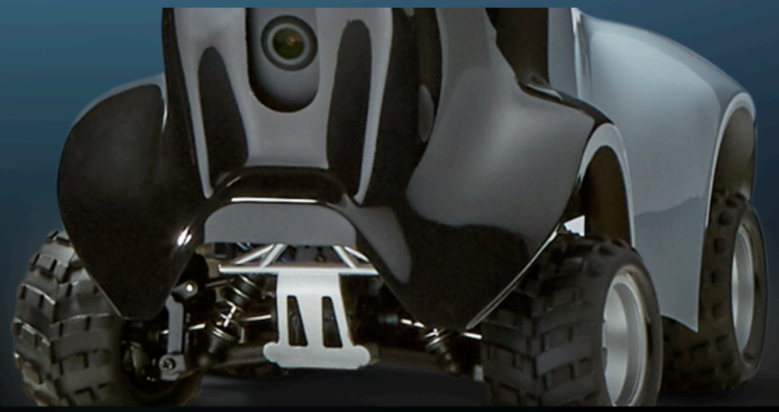
AWS DeepRacer 개요 이벤트 학생 시작하기 요금 FAQ

제품 › 기계 학습 › DeepRacer

시동 걸기

모든 기술 수준의 개발자가 클라우드 기반 3D 레이싱 시뮬레이터, 강화 학습으로 움직이는 1/18 비율의 완전 자율 경주용 자동차를 통해 기계 학습을 경험할 수 있습니다.

모델 구축



AWS에 가입

계정 플랜 선택

무료(6개월)

학습, 실험, 프로토타입 빌드

- ✓ 최대 200 USD의 크레딧 적립
- ✓ 일부 서비스 무료 사용 포함
- ✗ 크레딧 임계값을 넘어 워크로드 규모 조정
- ✗ 모든 AWS 서비스 및 기능 액세스

ⓘ 6개월의 무료 기간이 지나거나 크레딧을 모두 사용한 후에는 유료 플랜으로 업그레이드를 선택할 수 있습니다. 선택하지 않을 경우 계정이 자동으로 해지됩니다.

무료 플랜 선택

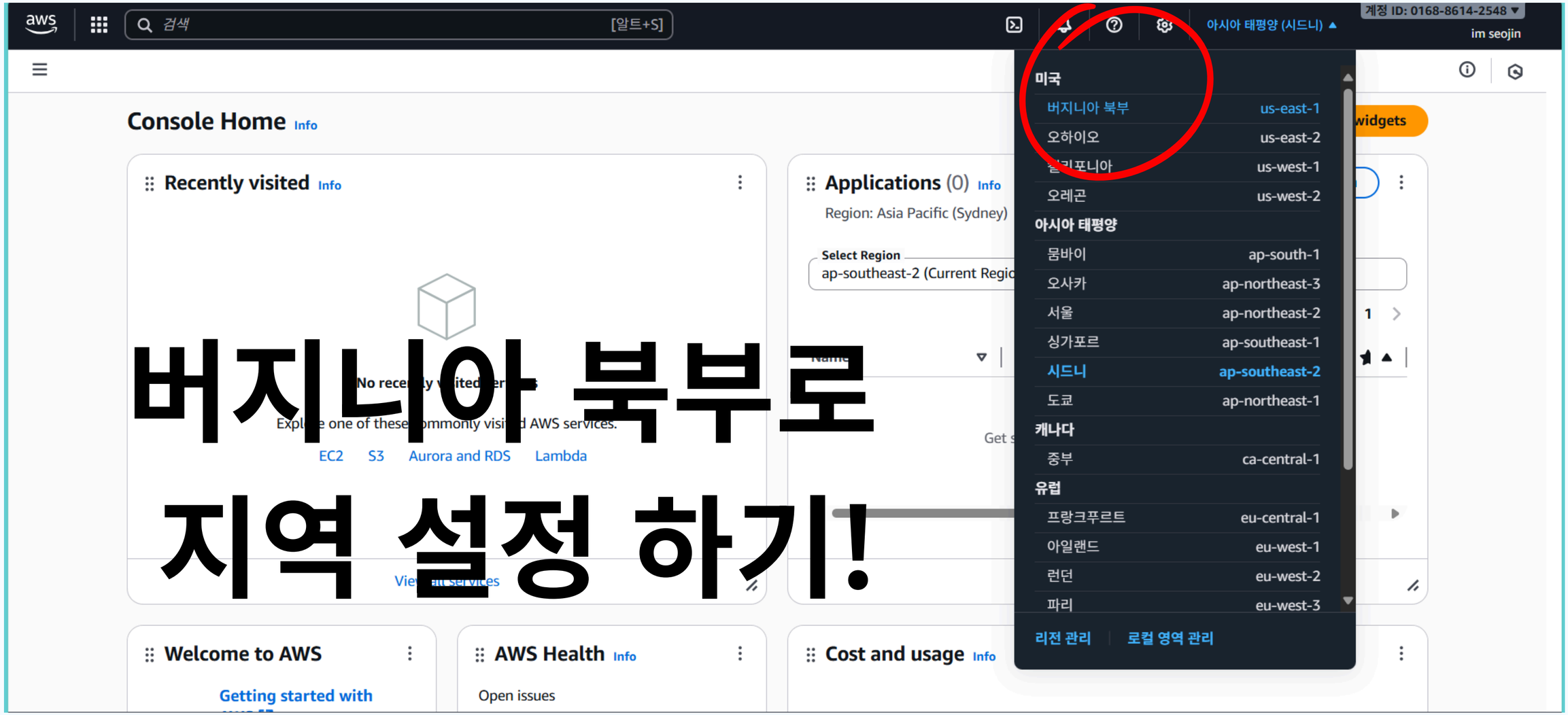
유료

프로덕션 레디 워크로드 개발

- ✓ 최대 200 USD의 크레딧 적립
- ✓ 일부 서비스 무료 사용 포함
- ✓ 크레딧 임계값을 넘어 워크로드 규모 조정
- ✓ 모든 AWS 서비스 및 기능 액세스

ⓘ 크레딧을 모두 사용하고 나면 종량제 요금을 사용해 요금이 부과됩니다.

유료 플랜 선택


aws
검색
[알트+S]
아시아 태평양 (시드니)
계정 ID: 0168-8614-2548
im seojin
Console Home Info
Recently visited Info
Applications (0) Info
Region: Asia Pacific (Sydney)
Select Region
ap-southeast-2 (Current Regio
EC2 S3 Aurora and RDS Lambda
미국
버지니아 북부 us-east-1
오하이오 us-east-2
us-west-1
오레곤 us-west-2
아시아 태평양
뭄바이 ap-south-1
오사카 ap-northeast-3
서울 ap-northeast-2
싱가포르 ap-southeast-1
시드니 ap-southeast-2
도쿄 ap-northeast-1
캐나다
중부 ca-central-1
유럽
프랑크푸르트 eu-central-1
아일랜드 eu-west-1
런던 eu-west-2
파리 eu-west-3
리전 관리
로컬 영역 관리
Welcome to AWS
Getting started with
AWS Health Info
Open issues
Cost and usage Info
widgets
버지니아 북부로
지역 설정 하기!

Console Home

myApplications

All services

AWS Migration Hub
AWS Application Migration Service
Application Discovery Service
Database Migration Service
AWS Transfer Family
AWS Snow Family
DataSync
AWS Transform
AWS Mainframe Modernization
Amazon Elastic VMware Service

네트워킹 및 콘텐츠 전송

VPC
CloudFront
API Gateway
Direct Connect
AWS App Mesh
Global Accelerator
Route 53
AWS 데이터 전송 터미널
AWS Cloud Map
Application Recovery Controller

개발자 도구

CodeCommit
CodeBuild
CodeDeploy
CodePipeline
Cloud9

MediaConnect

Machine Learning

Amazon SageMaker AI
Amazon Augmented AI
Amazon CodeGuru
Amazon DevOps Guru
Amazon Comprehend
Amazon Forecast
Amazon Fraud Detector
Amazon Kendra
Amazon Personalize
Amazon Polly
Amazon Rekognition
Amazon Textract
Amazon Transcribe
Amazon Translate
AWS DeepComposer
AWS DeepRacer
AWS Panorama
Amazon Monitron
AWS HealthLake
Amazon Lookout for Vision
Amazon Lookout for Equipment
Amazon Lookout for Metrics
Amazon Q Business
AWS HealthOmics
Amazon Bedrock
Amazon Bedrock AgentCore
Amazon Q

Simple Notification Ser
Simple Queue Service
SWF
관리형 Apache Airflow
AWS B2B Data Intercha
Amazon EventBridge

비즈니스 애플리케이

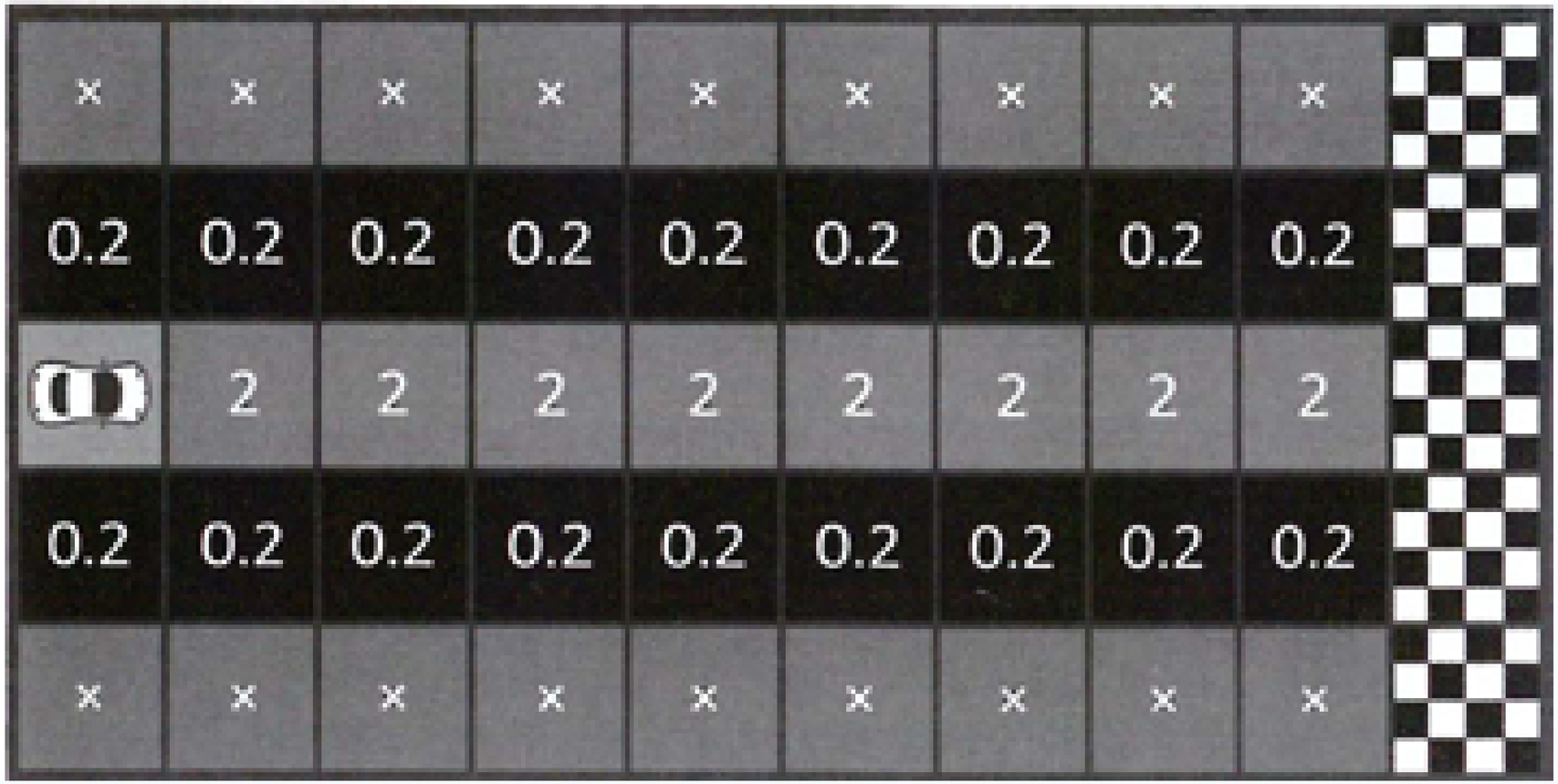
Amazon Connect
Amazon Chime
Amazon Simple Email S
Amazon WorkDocs
Amazon WorkMail
AWS Supply Chain
Amazon Pinpoint
Amazon One Enterprise
AWS Wickr
AWS AppFabric
AWS End User Messagi
Amazon Chime SDK

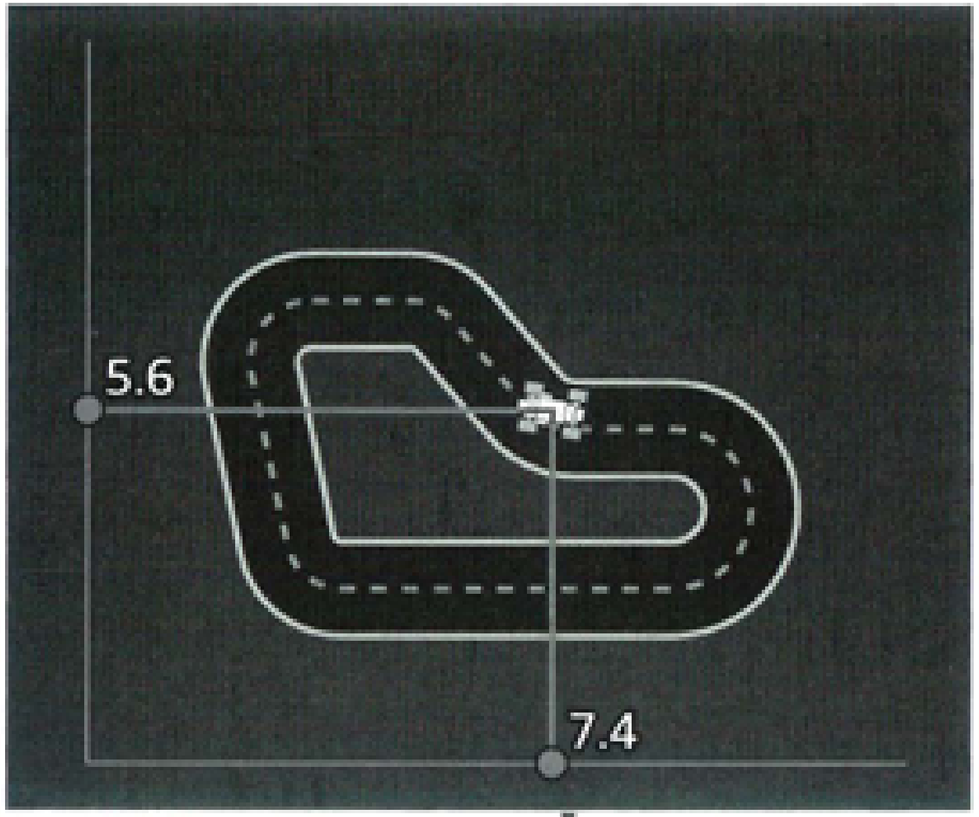
최종 사용자 컴퓨팅

WorkSpaces
AppStream 2.0
WorkSpaces Thin Client
WorkSpaces Secure Bro

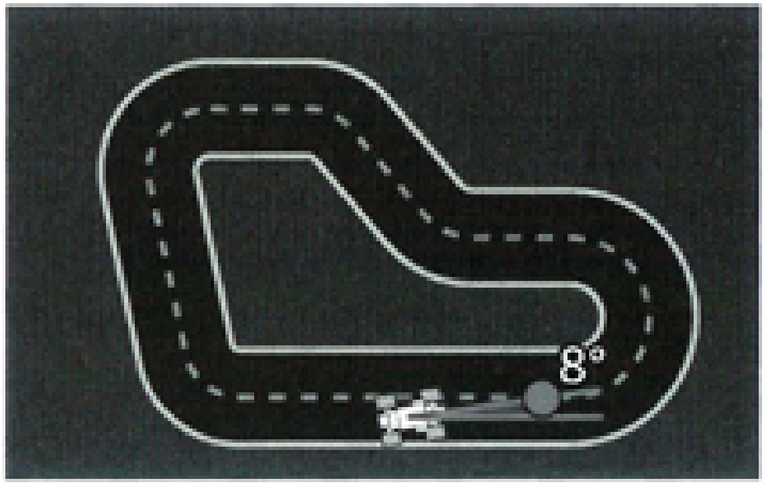
사물 인터넷

IoT Analytics
IoT Device Defender

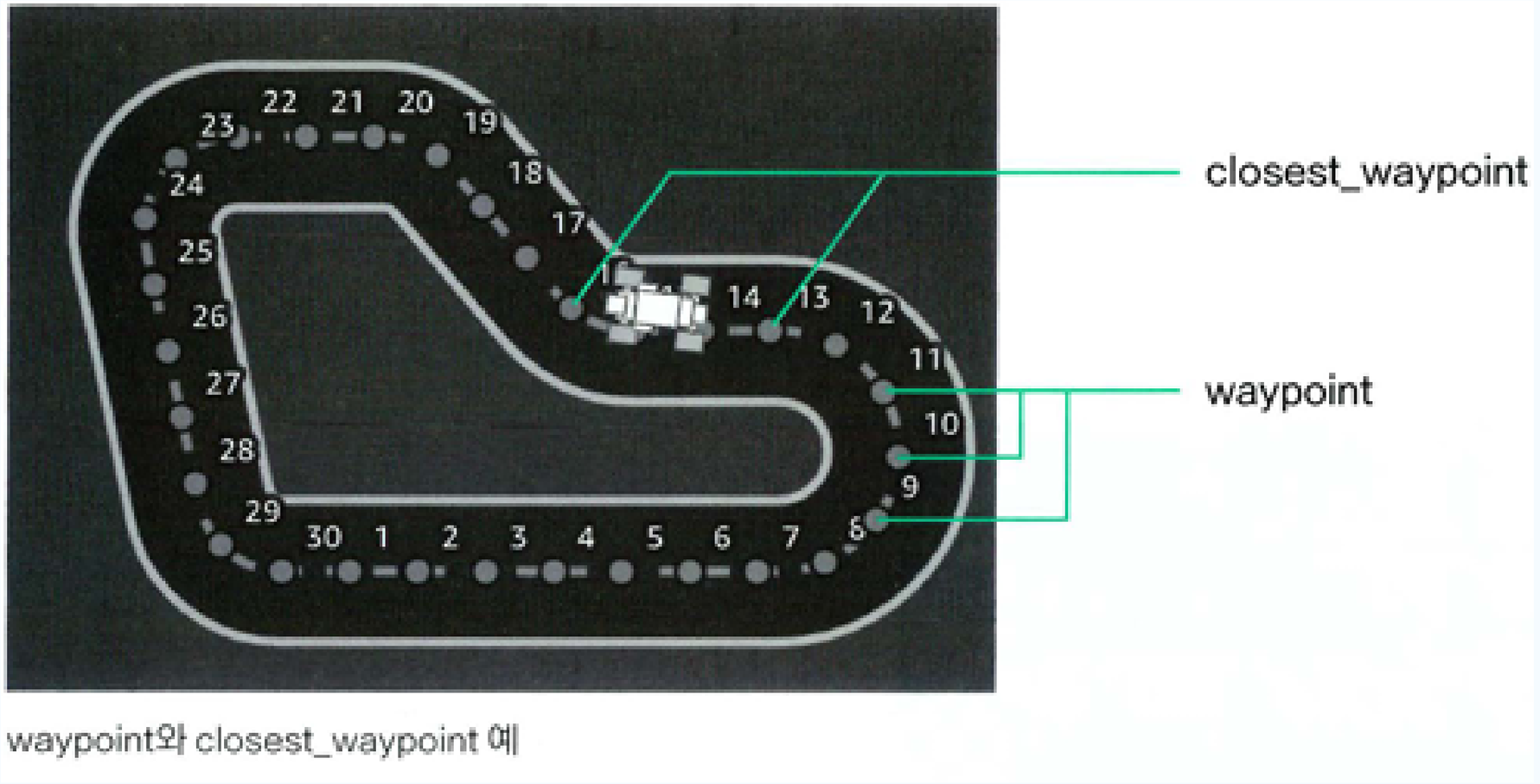

x x x x x x x x x
0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2
2 2 2 2 2 2 2 2
0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2 0.2
x x x x x x x x x



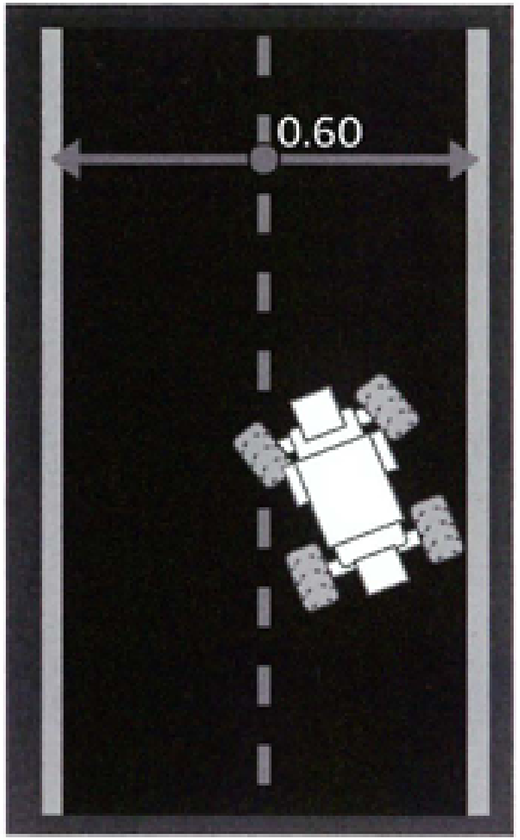
x, y 예



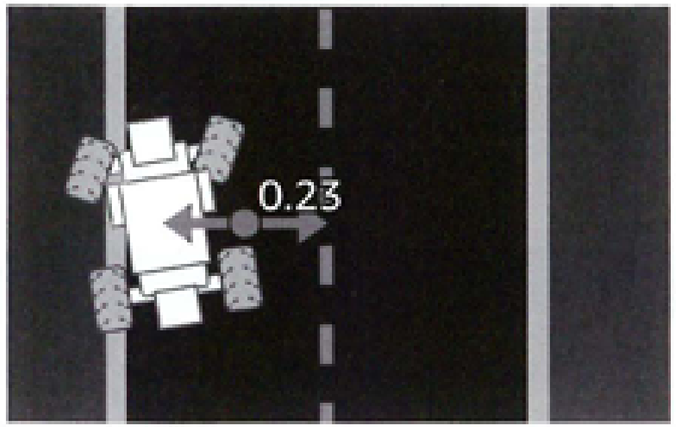
heading 예



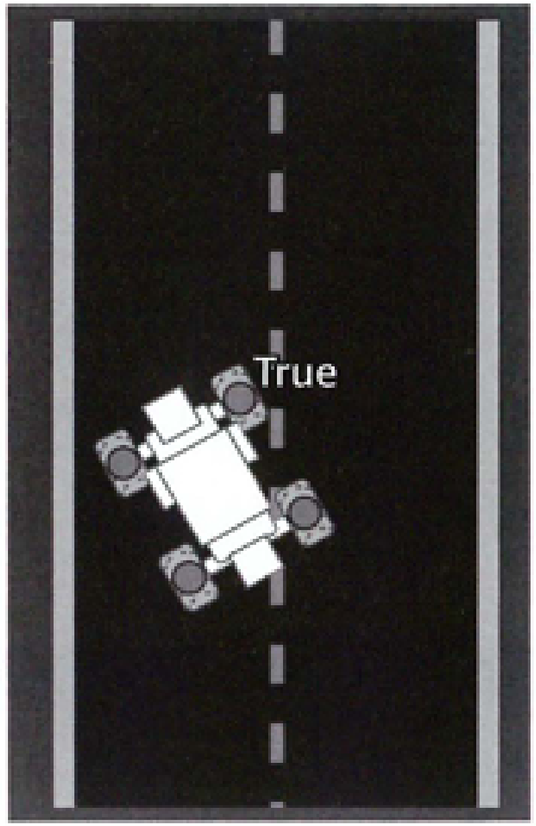
waypoint와 closest_waypoint 예



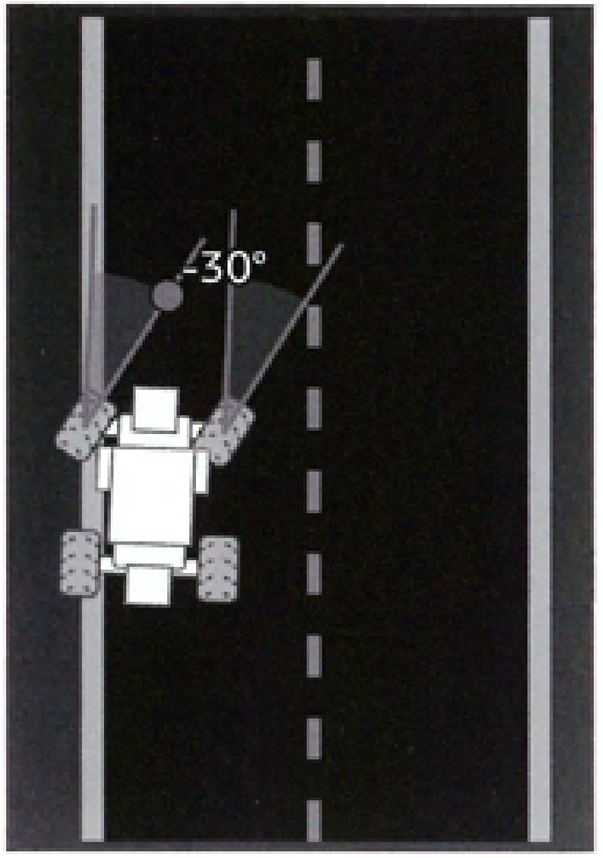
track_width 예



distance_from_center 예



is_left_of_center 예



steering_angle 예

과제안내

과제

✅해당 주차의 이론을 예습 하고 정리 하여 제출

✅과제 형식 : N회차_김마클_과제.md

제출 기한 : 매주 수업전 1시간까지

제출 → github N회차 폴더 > 본인 팀 폴더 > N회차_김마클_과제.md

- 9월 22일 오후 6시 대면 스터디 진행
- 2025 RL Study 카톡방에 과제 공지 예정
- 교재 5장 이론 부분(Do it 실습 제외) 노션으로 정리 후 깃허브 제출
- 과학술제 개인별 구글폼 제출

Thank You