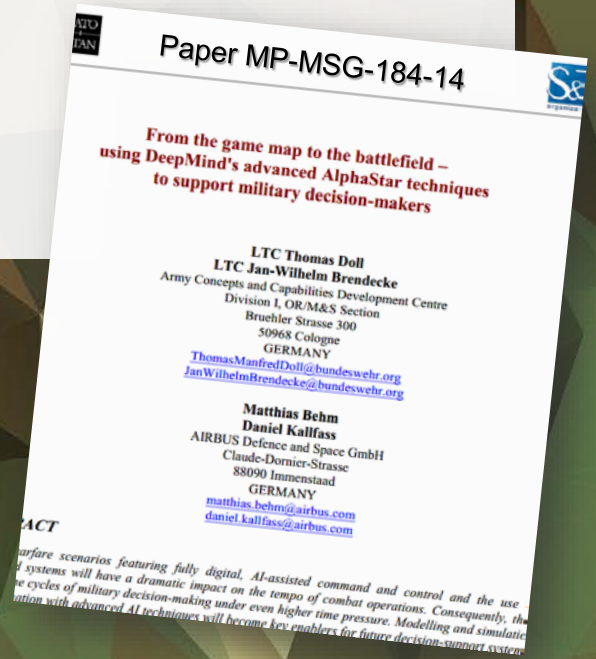
FROM THE GAME MAP TO THE BATTLEFIELD – USING DEEPMIND'S ADVANCED ALPHASTAR TECHNIQUES TO SUPPORT MILITARY DECISION-MAKERS

LTC Thomas Doll & MAJ Stefan Göricke
Army Concepts and Capabilities Development Centre, GERMANY
Daniel Kallfass
Airbus, GERMANY


Paper MP-MSG-184-14

From the game map to the battlefield – using DeepMind's advanced AlphaStar techniques to support military decision-makers

LTC Thomas Doll
LTC Jan-Wilhelm Brendecke
Army Concepts and Capabilities Development Centre
Division I, OR/M&S Section
Bruehler Strasse 300
50968 Cologne
GERMANY
ThomasManfredDoll@bundeswehr.org
JanWilhelmBrendecke@bundeswehr.org

Matthias Behm
Daniel Kallfass
AIRBUS Defence and Space GmbH
Claude-Dornier-Strasse
88090 Immenstaad
GERMANY
matthias.behm@airbus.com
daniel.kallfass@airbus.com

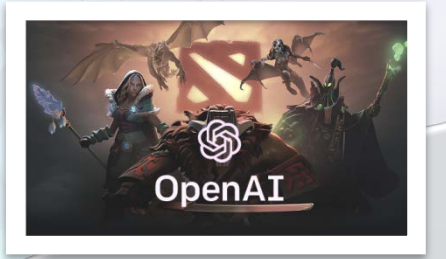
ACT

arfare scenarios featuring fully digital, AI-assisted command and control and the use
systems will have a dramatic impact on the tempo of combat operations. Consequently, th
e cycles of military decision-making under even higher time pressure. Modelling and simulati
ation with advanced AI techniques will become key enablers for future decision-support system

BUNDESWEHR

SUCCESSFUL USE-CASES OF DEEP-RL IN THE LAST DECADE

Source: https://openai.com/projects/five/

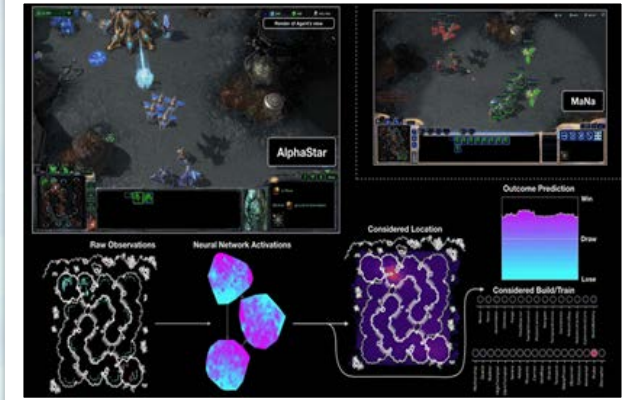


2018 Dota 2 (OpenAI)

2019 AlphaStar (DeepMind)

Source: https://deepmind.com/alphago-china

2016 AlphaGo Zero



Source: https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii

2013 Atari Games

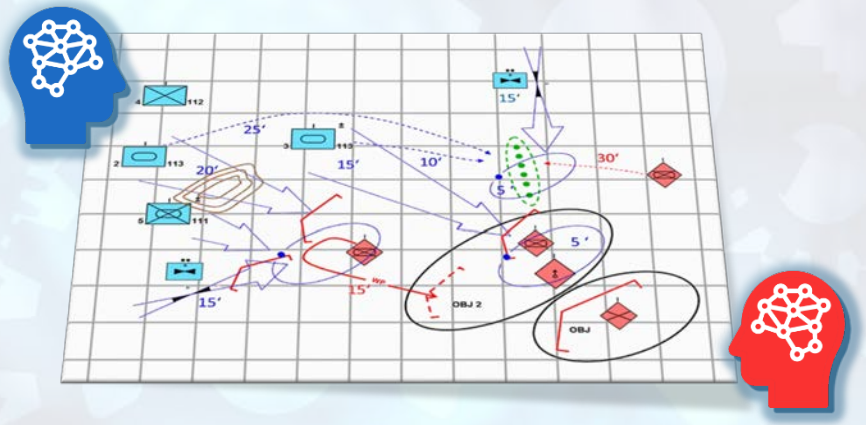
Source: https://www.cs.toronto.edu/~vmnih/docs/dqn.pdf

AI FOR DECISION SUPPORT

'Reinforcement learning for complex battlefield situations (ReLeGs)'

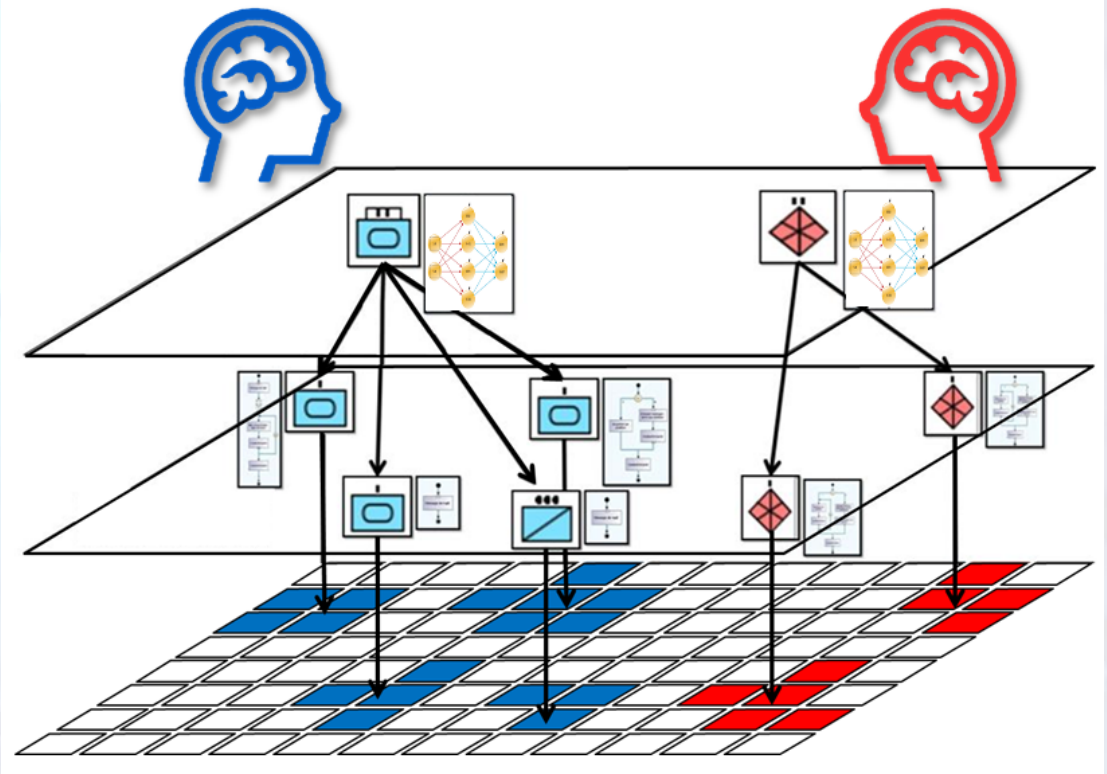
- **Context:** Speed-up the military decision cycle (OODA)
- **Goal:** Implement a simple Decision Support Tool on Bn level
- **Train** a deep neural network capable to:
  - Effectively Command available subordinated elements:
    - Tactical Units (Companies/ Platoons)
    - Combat Support
      (e.g. close air support, artillery, land mines)
  - Cope with uncertainties ('fog of war')
  - Adapt to the given terrain
  - Fulfil a military task

→ **"Super-human strategies"**

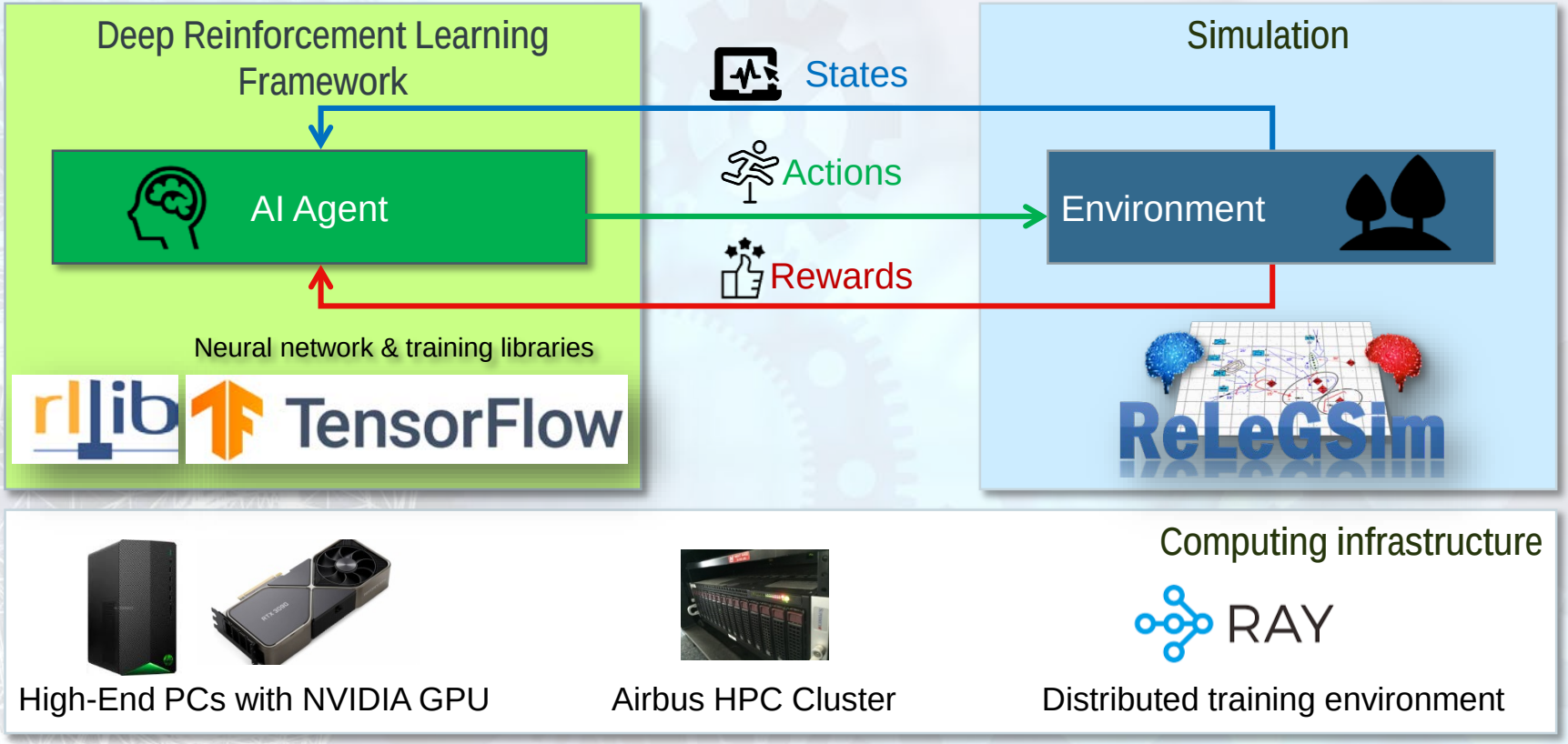


LEVELS OF AUTOMATION

- **AI agent** at the battalion / battle group level
- The AI agent **leads the companies/ platoons,** not the individual platforms
- Companies / platoons are **controlled algorithmically**
- The AI agent has a mission: **attack or defend** (**take / hold a space)**
- The situational picture of the battalion is **generated from the reconnaissance results of its units**

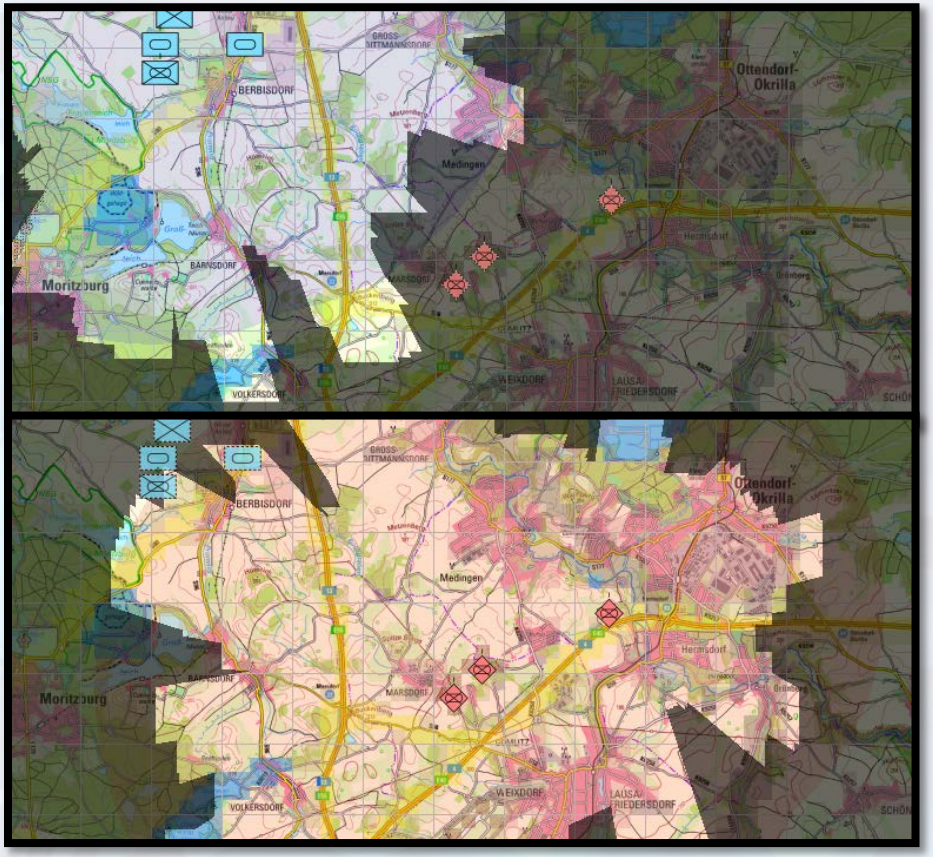


DEEP REINFORCEMENT LEARNING – SYSTEM VIEW

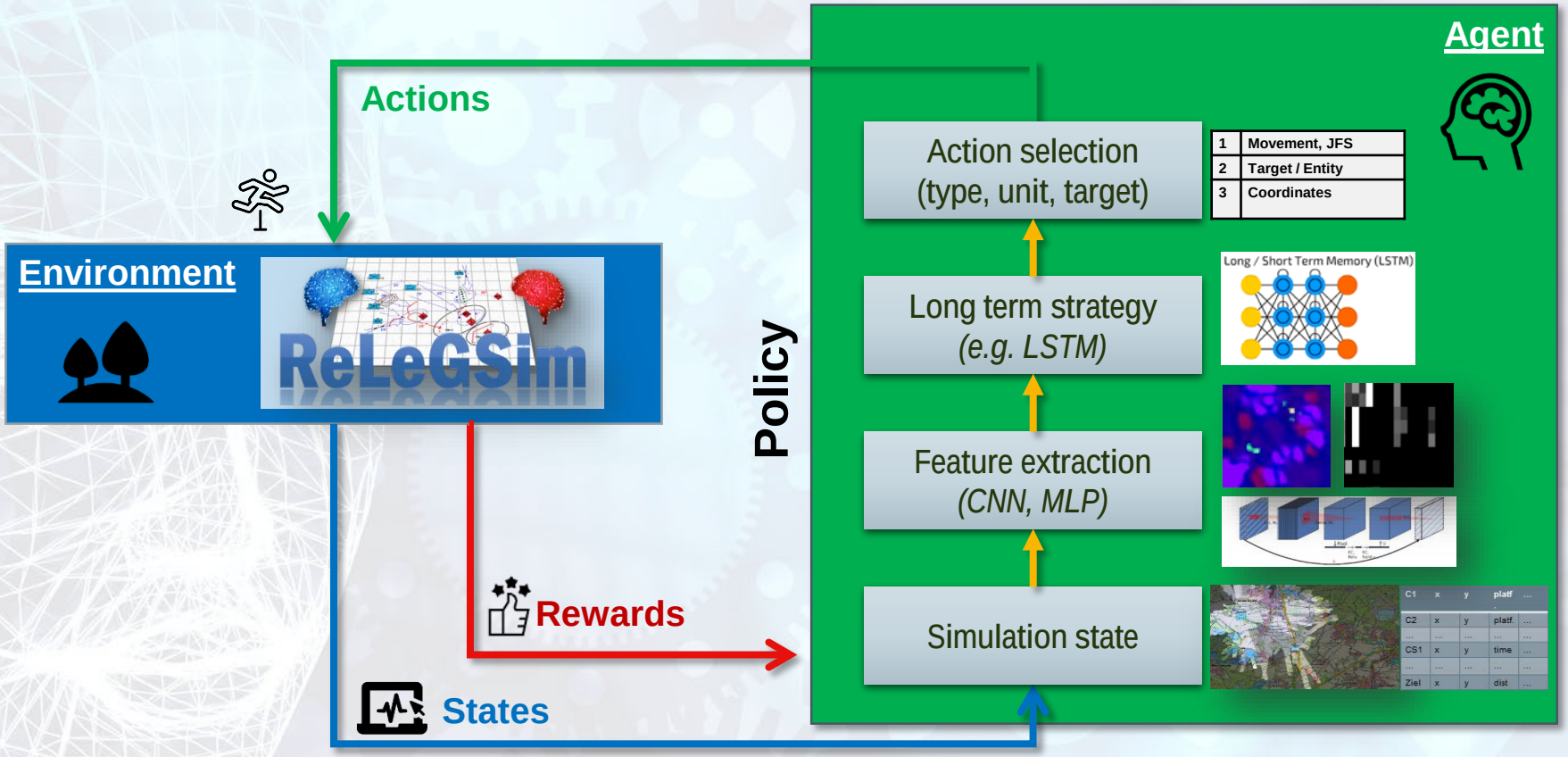


SIMULATION RELEGSIM AS A TRAINING ENVIRONMENT

- ReLeGSim developed as a **dedicated high-performance** training environment (constructive simulation)
- **"as simple as possible, as complex as necessary"**
- **Modelling optimized** for reinforcement learning
- **GUI optimized** for humans
- **Game modes:** computer-computer, human-human, human-computer
- **Randomized** training scenarios
- **Level / League System** for Training (RL)



DEEP REINFORCEMENT LEARNING – OVERVIEW



DEEP REINFORCEMENT LEARNING – NETWORK ARCHITECTURE

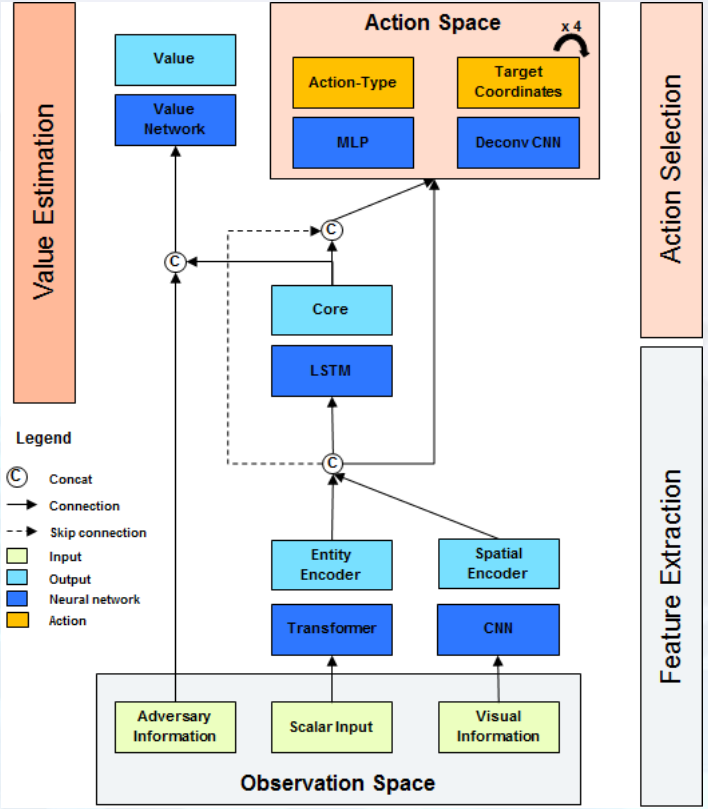
ReLeGs AI Model

based on Deepmind‘s AlphaStar:

- Scalar and Visual encoder
- Memory component
- Action Heads
- Value network

Optimized to tackle the challenges:

- significantly less computing power
- no replay data / no supervised learning



DEEP REINFORCEMENT LEARNING – NETWORK ARCHITECTURE

Observation Space

Scalar Input represents metadata of the environment state

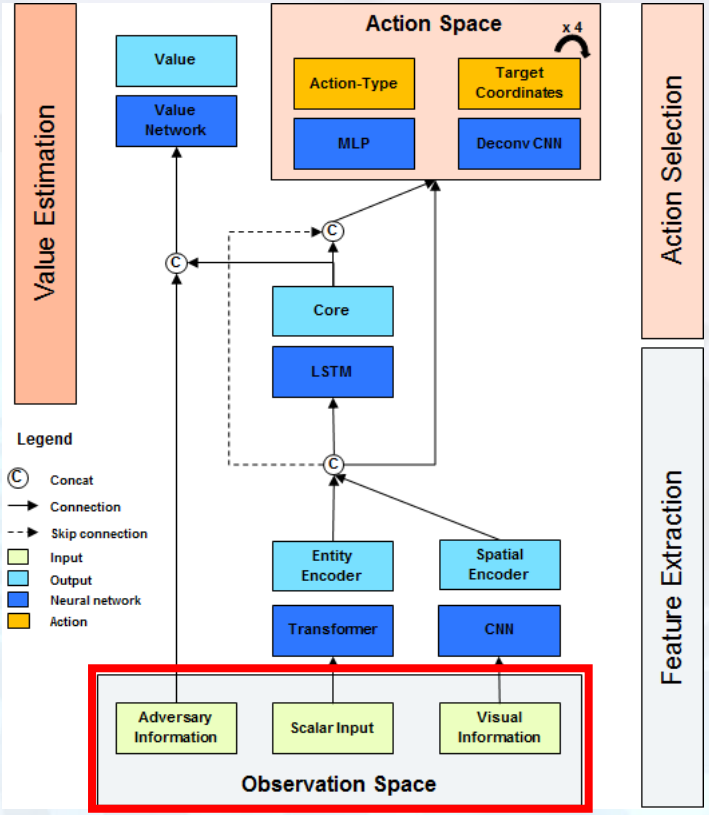
- company and platform parameters
- line-of-sight and effective ranges information
- terrain data (elevation, type etc.)

→ **Entity Encoder**

Visual Input contains the rasterized data

- terrain and elevation data
- units encoded using a defined color scheme

→**Spatial Encoder**

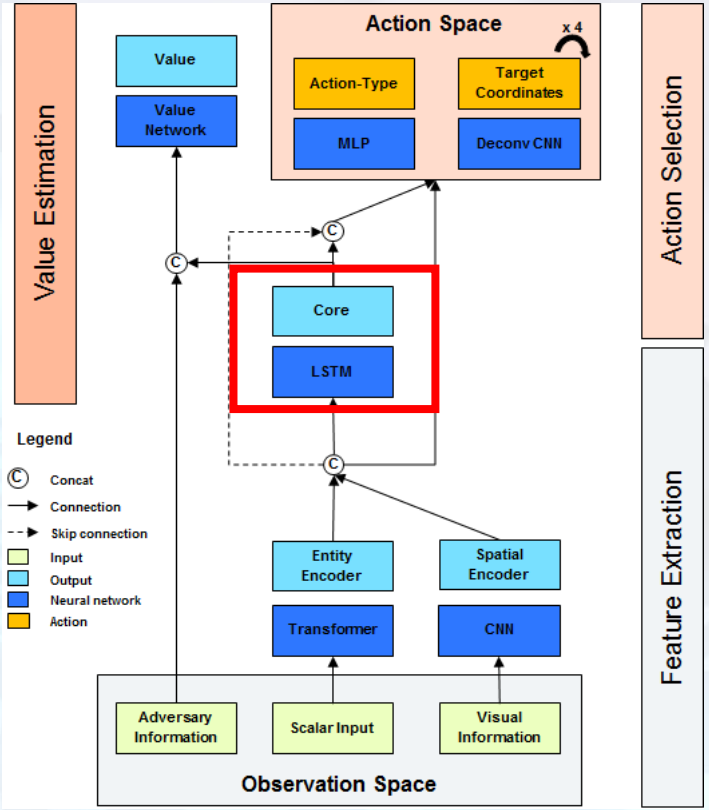


DEEP REINFORCEMENT LEARNING – NETWORK ARCHITECTURE

Core

The core consists of a LSTM component which addresses the long-term planning problem

- tasking / requesting of JFS
- prediction of enemy trajectories even if they went into the fog of war
- long term strategy development



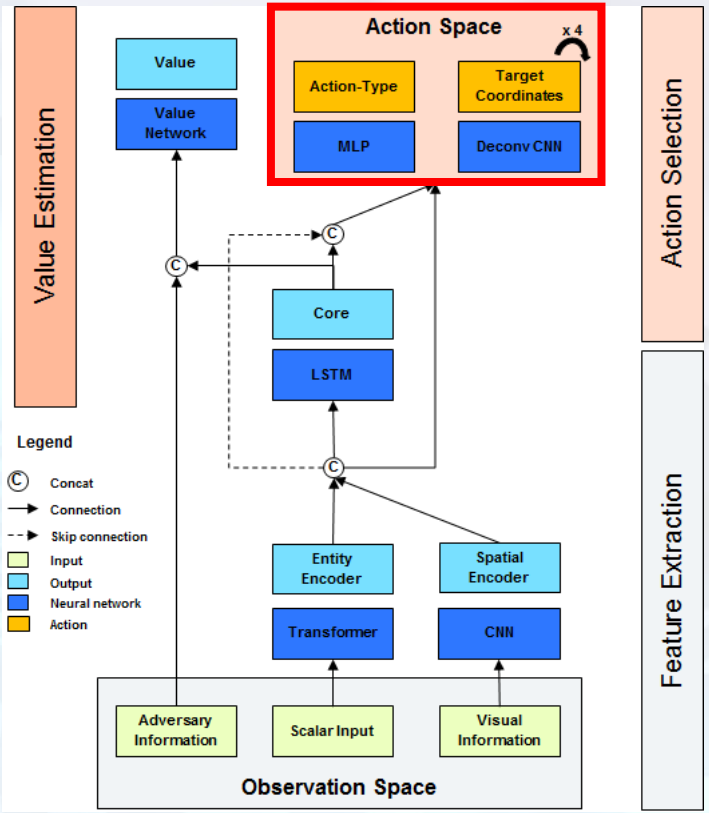
DEEP REINFORCEMENT LEARNING – NETWORK ARCHITECTURE

Action Space

The action heads consist of different MLPs (one per type of entity)

- outputs different actions; “do nothing”, “cancel current action”, “request new action”
- invalid actions are masked (e.g. JFS request with no JFS left)
- in case the requested action requires a position a location head is triggered
- The location head is an upscaling network which is shared by all units irrespective of their type

→ **There is one head for each entity type; regular unit, reconnaissance unit, combat support element, mines**



DEEP REINFORCEMENT LEARNING – REWARD DESIGN

Reward

The performance of the model and the goodness of the reward function are closely related

- the more specific a reward the more influence it has on the training

→can lead to poor and undesired behavior

→generic behavior can be achieved by designing the reward as abstract as possible

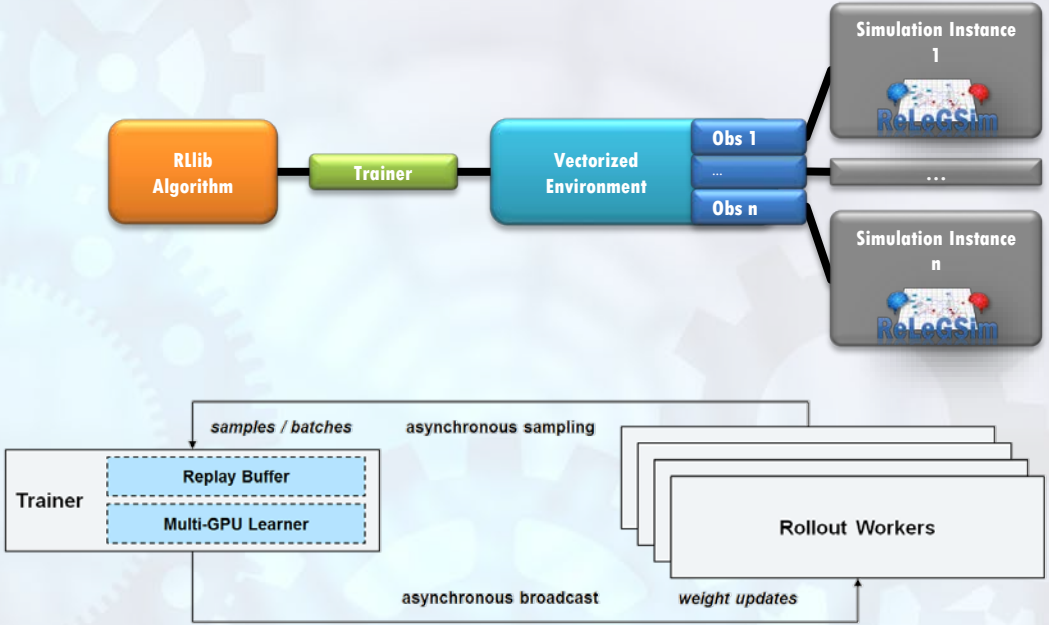
→**the best reward during the study consisted of a combination of rewards: win/loss of combat, reaching the objective, successful engagement / JFS and win/loss of episode**

DEEP REINFORCEMENT LEARNING – EXPERIMENT SETUP

Training Infrastructure

Deep RL requires massive computing power and lots of data as well as distributed computing capabilities

- ReLeGs uses RLlib as RL framework
- ReLeGSim environments run in parallel on a private cluster network
- Experience collection is done asynchronously to optimize CPU / GPU availability
- Rollout worker implementation allowed a high sample throughput of several thousand samples per second



DEEP REINFORCEMENT LEARNING – PRE-TRAINING

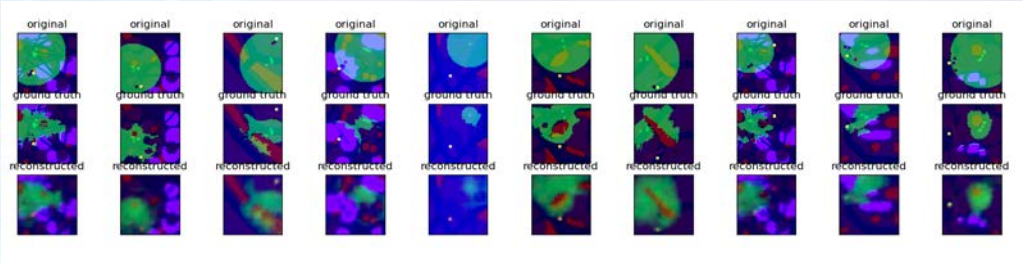
PRE-Training

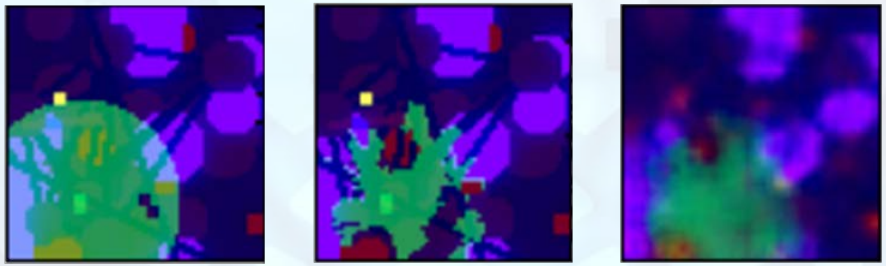
AlphaStar uses Human-Replays for a supervised pre-RL training phase

→ **ReLeGs AI was pre-trained using CNN auto-encoder/decoder approach**

- objective is to train the CNN
- benefits also the optimization of the CNN parameters (design can be optimized)
- train the agent to understand the Fog of War by including elevation and sensor information

→ **no algorithmic calculation necessary (computational expensive)**





COMPARISON ALPHASTAR VS. RELEGS

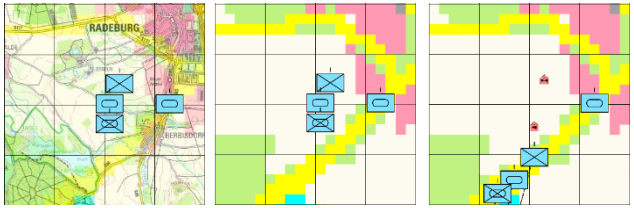
- **Specific ML environment had to be developed**
  - ReLeGSim as a simulation environment
  - Distributed high performance AI training environment
- **Significant reduction of complexity**
  - AlphaStar ~ 30 mio. parameters to be trained
  - ReLeGs ~ 7 mio. parameters to be trained
- **Minimisation of the required computing power**
  - AlphaStar ~ 16 TPUs + more than 1000 CPUs per agent
  - ReLeGs ~ 1 RTX 3090 + 150 CPUs per agent
- **Learning process starts at zero** („Blank Page“)
  - No replays available → no supervised learning phase

VS.

ReLeGSim

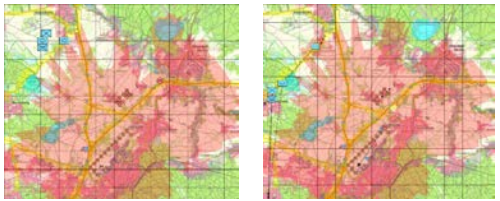
EXAMPLES FOR OBSERVED BEHAVIOR

1.



Early usage of the scout unit to observe and identify hostile units

2.



Take firing distances of hostile troops into account when moving own forces

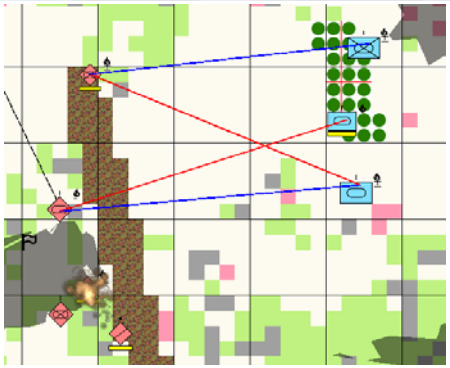
3.

Early request for fire support (before the attack)

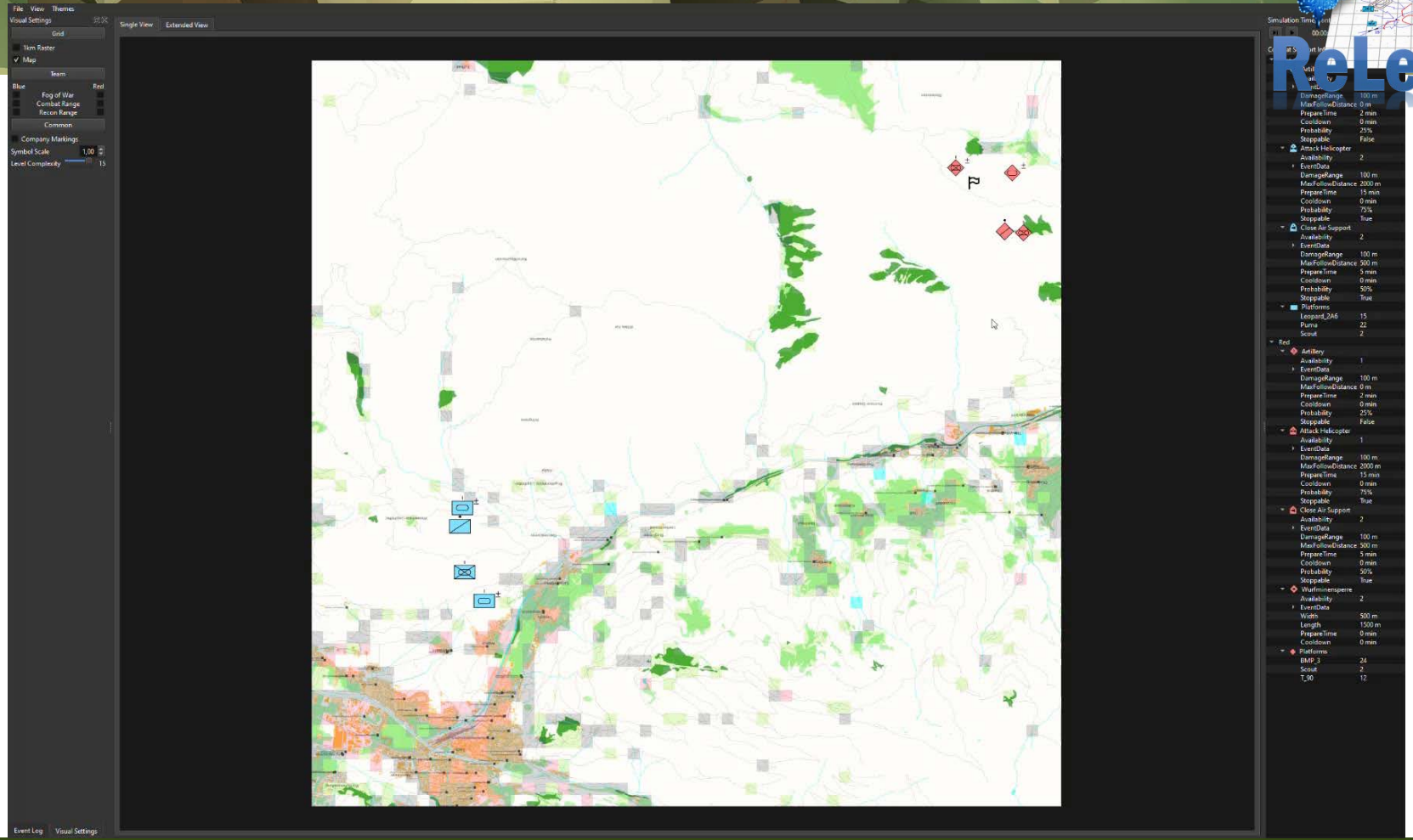
4.

Synchronized attack of ground forces (incl. pincer attack)

5.



Utilization of mines to stop attacker


File View Themes
Visual Settings
Grid
Tkm Raster
Map
Team
Blue
Red
Fog of War
Combat Range
Recon Range
Common
Company Markings
Symbol Scale
Level Complexity
Single View
Extended View
Event Log
Visual Settings
Simulation Time
Attack Helicopter
Close Air Support
Platforms
Leopard_2A6
Puma
Scout
Red
Artillery
Wurfminensperre
BMP_3
T_90

CONCLUSION

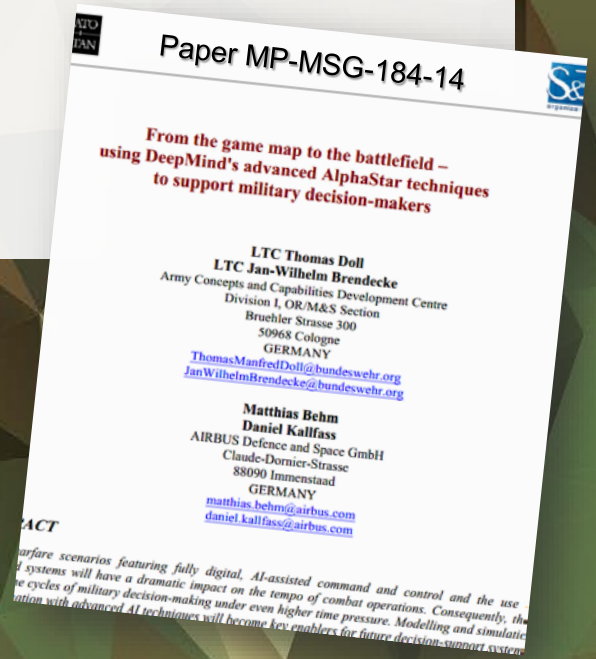
- **Deep reinforcement learning** is a disruptive technology with great potential for automated learning of reasonable tactical behavior
- **Approaches already used within game industry** can successfully be transferred into the military domain
- **ReLeGSim** can be used as **platform** for deep-RL **in extended scenarios** (other level of command, different branches, more capabilities)
- **Different Fields of Application** are possible (Decision Support, Computer Generated Forces for Training and Exercise)
- **HOWEVER:** Know-How of AI in Bundeswehr and Industry needs to be improved and transferred to real-world applications

‚Knowledge Must Become Capability‘
(Clausewitz, *On War***)**

THANK YOU!
ANY QUESTIONS?

FROM THE GAME MAP TO THE BATTLEFIELD – USING DEEPMIND'S ADVANCED ALPHASTAR TECHNIQUES TO SUPPORT MILITARY DECISION-MAKERS

LTC Thomas Doll & MAJ Stefan Göricke
Army Concepts and Capabilities Development Centre, GERMANY
Daniel Kallfass
Airbus, GERMANY


Paper MP-MSG-184-14

From the game map to the battlefield – using DeepMind's advanced AlphaStar techniques to support military decision-makers

LTC Thomas Doll
LTC Jan-Wilhelm Brendecke
Army Concepts and Capabilities Development Centre
Division I, OR/M&S Section
Bruehler Strasse 300
50968 Cologne
GERMANY
ThomasManfredDoll@bundeswehr.org
JanWilhelmBrendecke@bundeswehr.org

Matthias Behm
Daniel Kallfass
AIRBUS Defence and Space GmbH
Claude-Dornier-Strasse
88090 Immenstaad
GERMANY
matthias.behm@airbus.com
daniel.kallfass@airbus.com

ACT

arfare scenarios featuring fully digital, AI-assisted command and control and the use
systems will have a dramatic impact on the tempo of combat operations. Consequently, th
e cycles of military decision-making under even higher time pressure. Modelling and simulati
ation with advanced AI techniques will become key enablers for future decision-support system

BUNDESWEHR