

THE AWKWARD REAL-TIME ADJUSTMENT OF REACTIVE PLANS

Leila Methnani, Andreas Antoniadis and
Andreas Theodorou

leila.methnani@cs.umu.se



UMEÅ UNIVERSITY

AGENDA

- Hybrid architecture: AWKWARD
- Implementation in DOTA2
- Results
- On-going extensions



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

LONG TERM MOTIVATION

- Behaviour adjustment to satisfy social obligation in multi-agent setting.
- Multi-agent interactions can be difficult to interpret.
- Human control; leverage human knowledge about the problem.



Can we achieve real-time adjustment of reactive plans with the added benefit of transparency into social interactions between agents?



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

Can we achieve real-time adjustment of reactive plans with the added benefit of transparency into social interactions between agents?



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

SYMBOLIC MEETS BEHAVIOUR-BASED AI

- **OperA** framework for developing MAS.
- Offers transparency by modelling the agent interactions.
- Deontic logic norm validation.
- Real-time adjustment.
- **Behaviour Oriented Design** cognitive architecture.
- Handles uncertainty in the environment.
- Modularity and agile methodology.
- Used in embodied agents.

Dignum, V. (2004). *A model for organizational interaction: based on agents, founded in logic*. PhD thesis.

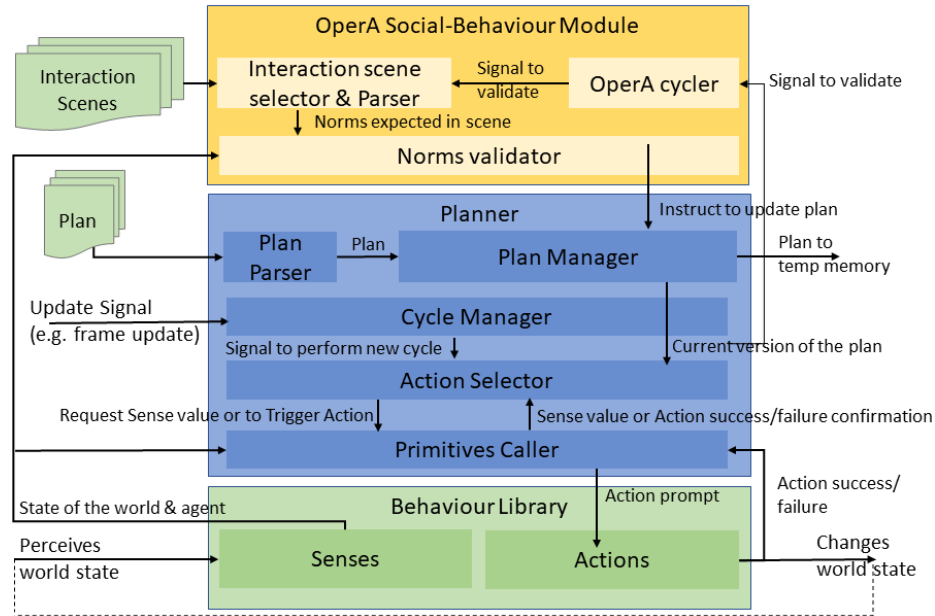
Bryson, J. J. (2001). *Intelligence by design: Principles of modularity and coordination for engineering complex adaptive agents*. PhD thesis,



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

AGENTS WITH KNOWLEDGE ABOUT REAL-TIME DUTIES



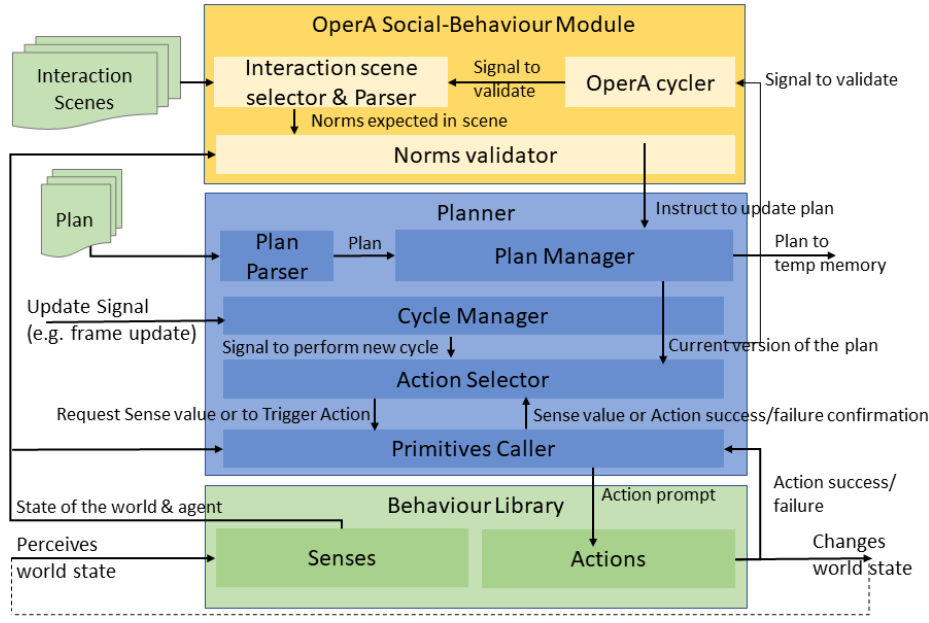
Methnani, L., Antoniadou, A., & Theodorou, A. (2022).
Embracing AWKWARD! Real-time Adjustment of Reactive
Planning Using Social Norms. *arXiv preprint*
arXiv:2204.10740.



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

AWKWARD



OperA module corresponds to Kahneman's 'slow' system 2.

Planner module corresponds to Kahneman's 'fast' system 1.

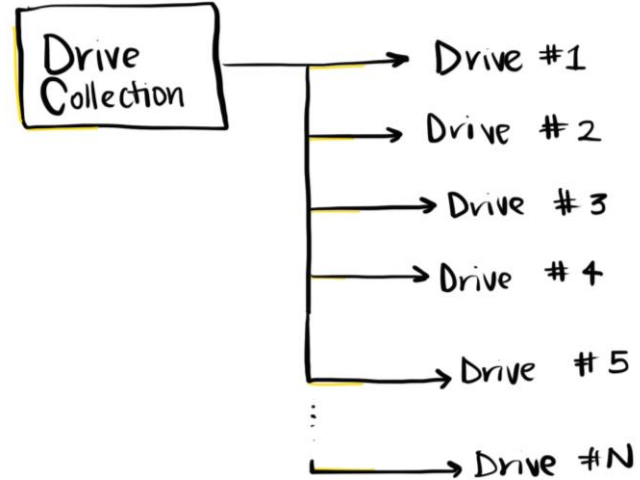
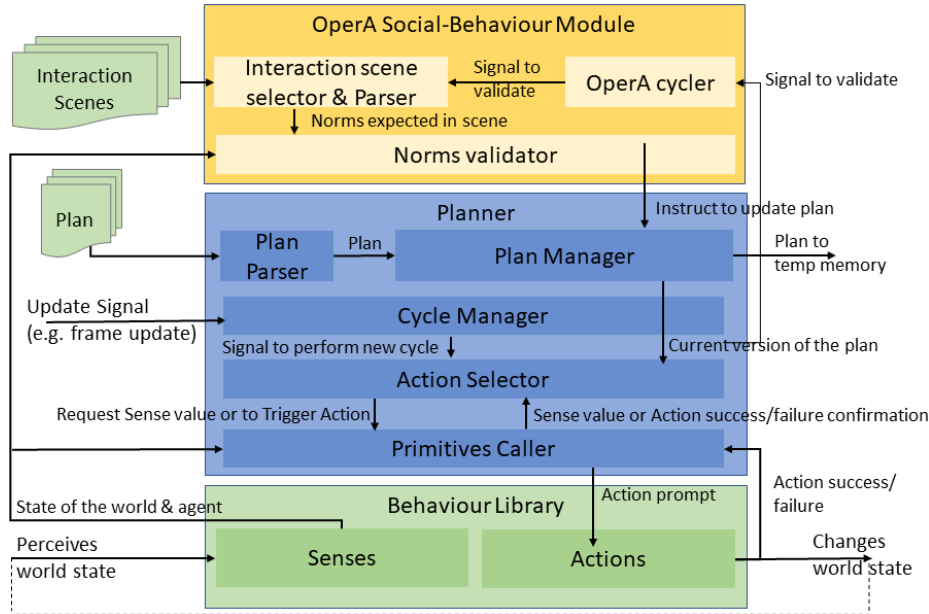
Shared access embodiment.

Methnani, L., Antoniadou, A., & Theodorou, A. (2022). Embracing AWKWARD! Real-time Adjustment of Reactive Planning Using Social Norms. *arXiv preprint arXiv:2204.10740*.



UMEÅ UNIVERSITY
LEILA.METHNANI@CS.UMU.SE

AWKWARD

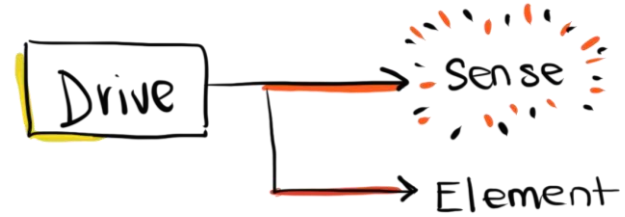
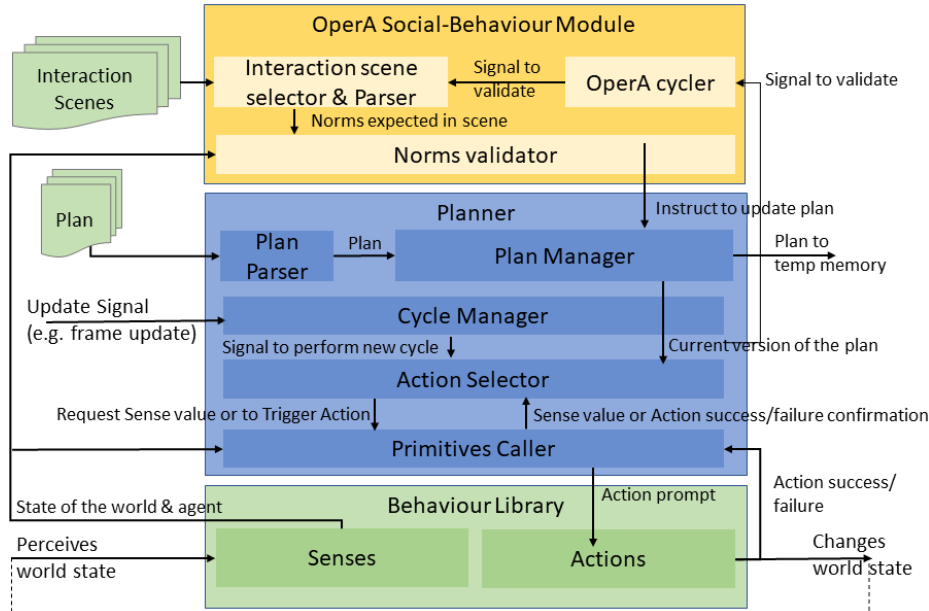


Methnani, L., Antoniadou, A., & Theodorou, A. (2022). Embracing AWKWARD! Real-time Adjustment of Reactive Planning Using Social Norms. *arXiv preprint arXiv:2204.10740*.



UMEÅ UNIVERSITY
LEILA.METHNANI@CS.UMU.SE

AWKWARD



Methnani, L., Antoniadou, A., & Theodorou, A. (2022).
Embracing AWKWARD! Real-time Adjustment of Reactive
Planning Using Social Norms. *arXiv preprint*
arXiv:2204.10740.



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

INTERACTION SCENE

- Landmarks indicating scene start and end.
- Roles involved in interaction.
- Rules for scene duration.
 - Deontic operators describing permissions and obligations for each role.

```
IF ... THEN
    OBLIGED(role A, behaviour)
ELSE
    NOT PERMITTED(role B, behaviour)
ENDIF
```

Dignum, V. (2004). *A model for organizational interaction: based on agents, founded in logic.*
PhD thesis.



UMEÅ UNIVERSITY

NORM ENFORCEMENT

- Checks state using reactive planner's sense element.
- Agent behaviour compared against norms encapsulated by scene rules.
 - Upon norm violation, OperA manipulates agent's plan.
 - E.g., removing a drive that is not permitted within the specified context.



IMPLEMENTATION



UMEÅ UNIVERSITY

DOTA 2



- Multiplayer real-time strategy game.
- Complex and dynamic environment.
- Varied hero abilities; social agent interaction is key.



FARMING BEHAVIOUR



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

INTERACTION SCENES

Table 3: Interaction Scene for Priority Lane Farming

scene	priority lane farming
roles	carry and support
landmarks	partner and creeps nearby
results	partner not nearby
rules	IF highest priority around THEN OBLIGED to farm ELSE NOT PERMITTED to farm



Position 1



Position 5



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

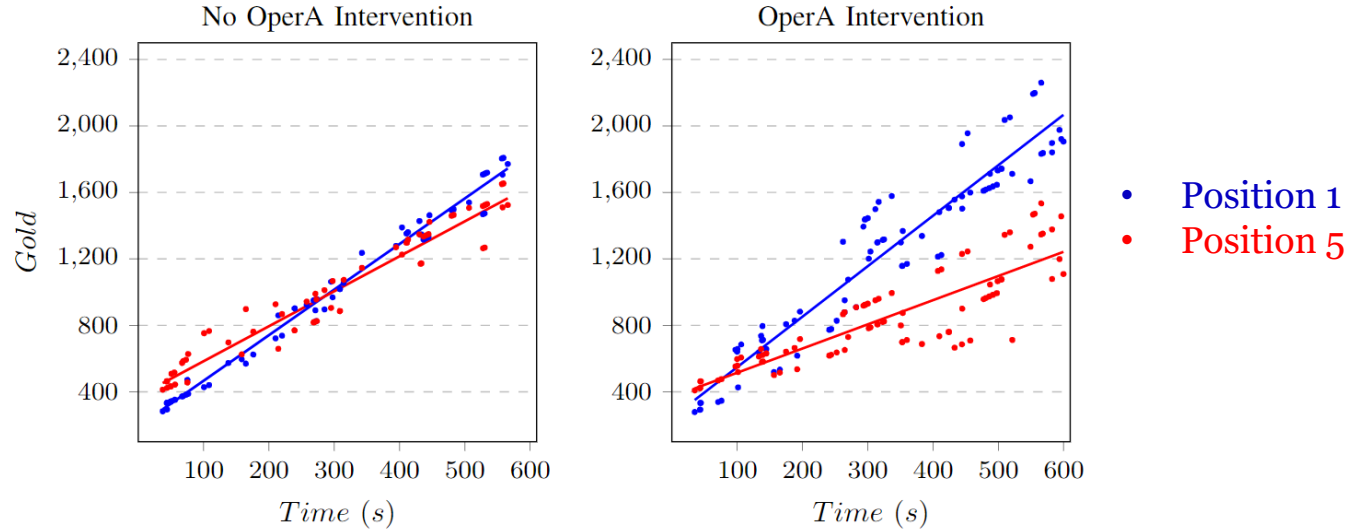
SOCIAL INTERACTION



- OperA 'Priority Farm' Scene is running.
- Enforcing the 'No Farm' Norm for hero of lower priority, i.e. Witch Doctor: Position 5.
- Harassing enemy heroes instead.



IN-GAME RESULTS



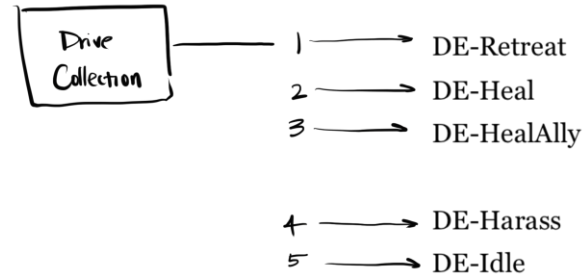
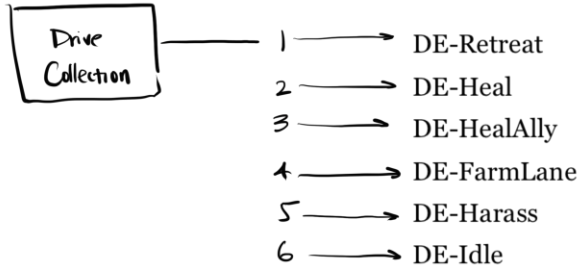
Methnani, L., Antoniadou, A., & Theodorou, A. (2022).
Embracing AWKWARD! Real-time Adjustment of Reactive
Planning Using Social Norms. *arXiv preprint*
arXiv:2204.10740.



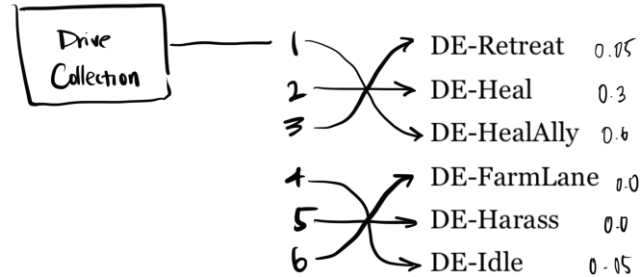
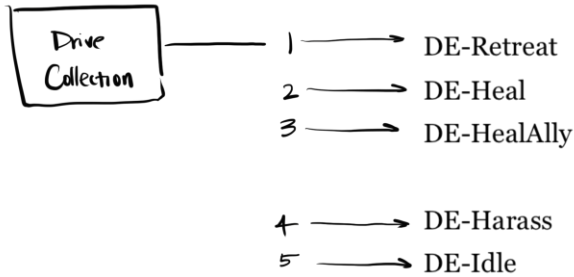
UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

OUT-OF-SCENE VS IN-SCENE PLAN STRUCTURES



IS THERE A BETTER STRUCTURE?



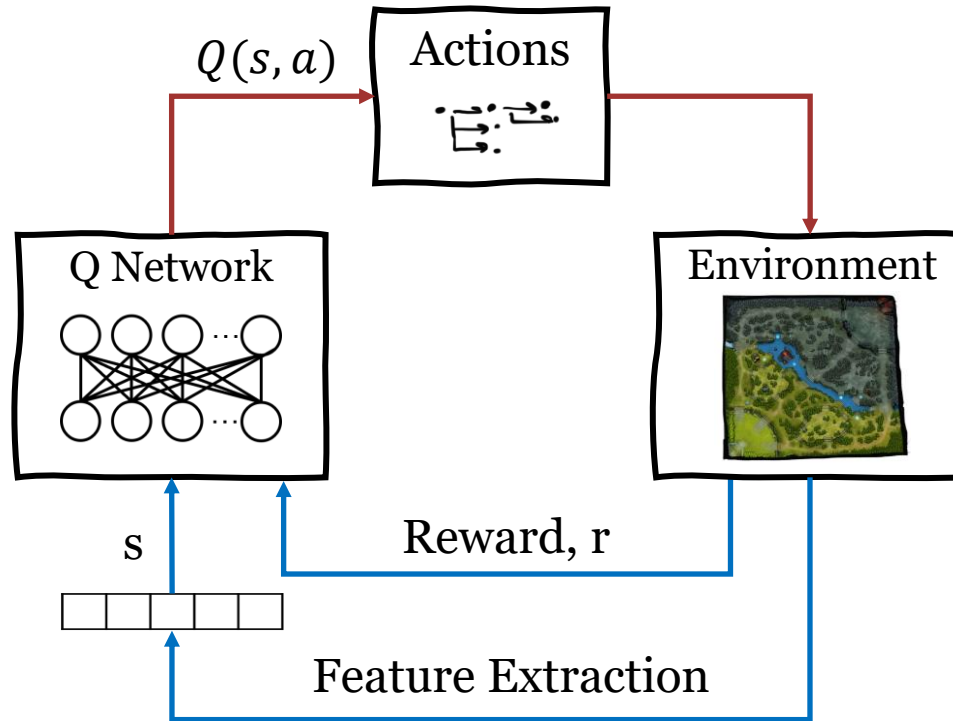
UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

ONGOING EXTENSIONS



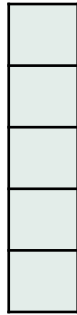
UMEÅ UNIVERSITY



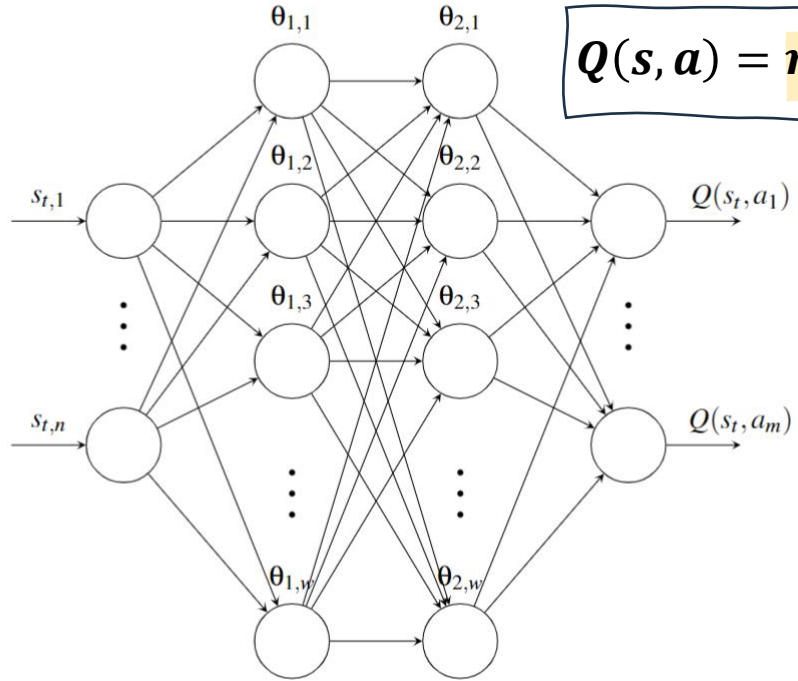
UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

Distance
 Agent Gold
 Partner Gold
 Agent Health
 Partner Health



Input layer Hidden layer Hidden layer Output layer



$$Q(s, a) = r(s, a) + \gamma \max_a Q(s', a)$$

- DE-Retreat
- DE-Heal
- DE-HealAlly
- DE-FarmLane
- DE-Harass
- DE-Idle

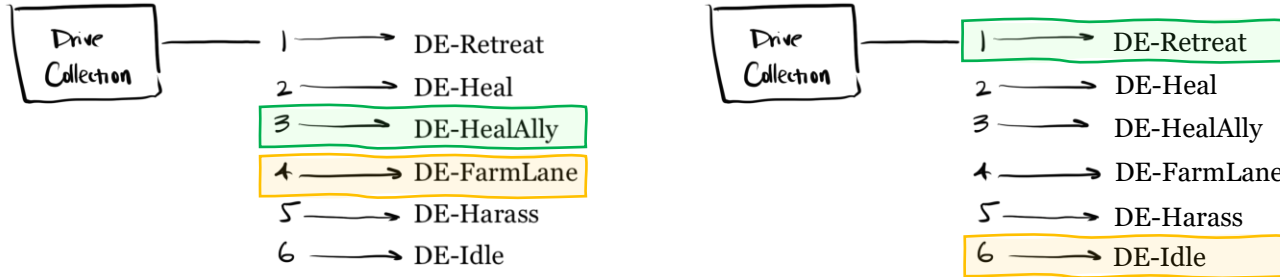


PROBLEM FORMULATION

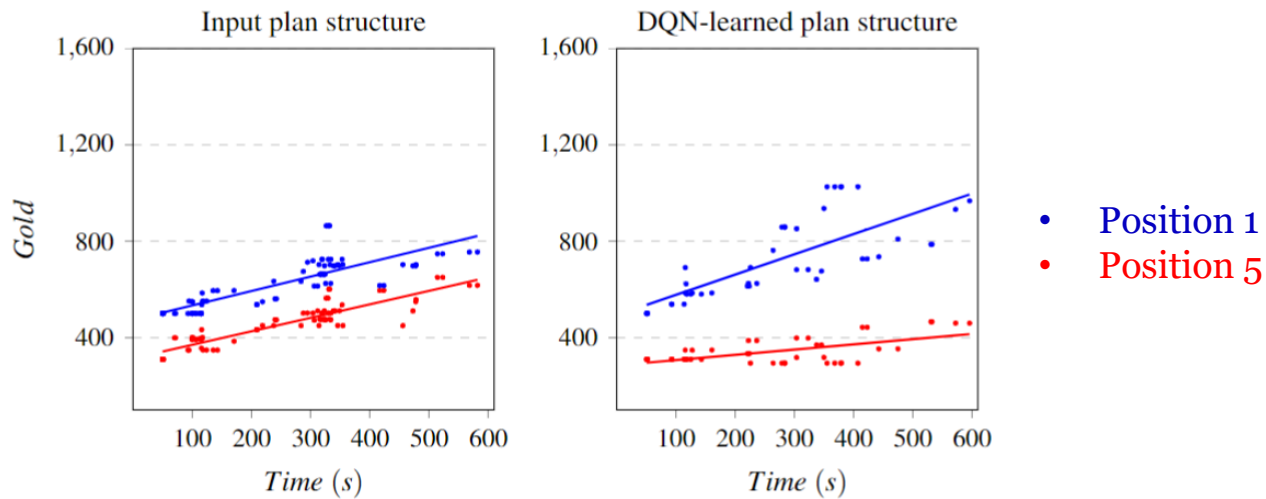
- **Episode:**
 - Duration of an interaction scene.
- **Goal (social outcome):**
 - Maximise Position 1 gold acquisition.
- **Reward system:**
 - Increased position 1 gold = +50
 - Increased self gold = +10
 - Otherwise = +0



PRELIMINARY RESULTS



RESULTS



UMEÅ UNIVERSITY

LEILA.METHNANI@CS.UMU.SE

CONSIDERATIONS

- Well, if we can induce the same (or an even better) outcome without RL ... why use it?
- Larger action-state space → automating plan structures become more useful.
- State representation and shaping the reward function matters; more study is needed here.



MODULE DESIGN

- Feature extraction
 - What state variables are relevant?
- Reward shaping
 - What supplemental rewards can motivate the big payoff?



CHALLENGES

- State-action space complexity.
 - Sample efficiency.
 - Training time.
- Designing a suitable reward.
- Reproducibility.



FUTURE WORK

- Expand the scope:
 - Interaction Scenes.
 - Behaviour Drives.
- User studies for transparency.
- Optimise the DQN.
- Include a human in the loop (both as module and as interactive RL).
- Packaging as an educational/research platform.



HMIEAI?

- New, “weird” → AWKWARD (hybrid).
- Interpretable, practical, useful → motivator for design choices.
- AI is not just machine learning ... but it is also machine learning → somewhat “plug and play”.



THANK YOU

Questions?



UMEÅ UNIVERSITY