



Universiteit
Leiden
The Netherlands

Learning in automated negotiation

Renting, B.M.

Citation

Renting, B. M. (2025, December 11). *Learning in automated negotiation*.
Retrieved from <https://hdl.handle.net/1887/4284788>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4284788>

Note: To cite this publication please use the final published version (if applicable).

SAMENVATTING

Dit proefschrift onderzoekt het ontwerp en de evaluatie van geautomatiseerde onderhandelingsagenten binnen het domein van multi-agent systemen, met de nadruk op de ontwikkeling van lerende onderhandelingsagenten om de prestaties en generaliseerbaarheid van zulke agenten te verbeteren. Traditionele benaderingen voor het ontwikkelen van onderhandelingsagenten zijn gebaseerd op handmatig ontworpen heuristieken en strategieën. Hoewel dergelijke methoden effectief zijn in specifieke scenario's, hebben ze te kampen met beperkingen, waaronder menselijke ontwikkelingskosten, slechte generaliseerbaarheid voor verschillende onderhandelingsproblemen en soorten tegenstanders, en menselijke voorkeuren in het ontwerp van strategieën. Het eerste deel van dit proefschrift gaat in op deze beperkingen door methoden voor het leren van onderhandelingsstrategieën te onderzoeken. We behandelen verschillende methoden waarmee autonome agenten effectief onderhandelingsgedrag kunnen leren.

Hoofdstuk 3 presenteert een geautomatiseerde algoritmeconfiguratie techniek, specifiek Sequential Model-based Algorithm Configuration (SMAC) [75], voor de optimalisatie van de parameters van een handmatig gedefinieerde onderhandelingsstrategie ruimte. Deze aanpak toont een verbetering aan ten opzichte van handmatig ontworpen strategieën en configuraties die volgen uit literatuur, terwijl nog steeds wordt vertrouwd op een vooraf gedefinieerde onderhandelingsstrategie ruimte en handmatig ontworpen instantiekenmerken om het configuratieproces te begeleiden.

Omdat we beseffen dat één enkele geoptimaliseerde strategie waarschijnlijk niet optimaal is voor elke situatie, gaan we in Hoofdstuk 4 verder met het onderzoeken van portfolio-gebaseerde benaderingen. Met behulp van methoden zoals Hydra [162] voor het samenstellen van portefeuilles van strategieën en AutoFolio [100] voor het selecteren van de beste strategie per situatie, wordt automatisch een portefeuille van complementaire onderhandelingsstrategieën gegenereerd en wordt een selector getraind. Hierdoor kan de agent dynamisch de meest geschikte strategie uit zijn portefeuille kiezen op basis van de kenmerken van het huidige onderhandelings situatie en de tegenstander, wat leidt tot een verbeterde aanpassingsvermogen en algehele prestaties in vergelijking met het gebruik van één enkele beste strategie.

Om de afhankelijkheid van handmatig ontwerp verder te verminderen en door mensen ingebouwde voorkeuren mogelijk nog substantiëler te beperken, onderzoekt Hoofdstuk 5 een benadering met behulp van reinforcement learning. We maken gebruik van Proximal Policy Optimisation (PPO) [135] en Graph Neural Networks (GNN's) [86] om de variabele dimensionaliteit in de observatie- en actie-ruimtes van diverse onderhandelings situaties te kunnen hanteren. Hierdoor kan de onderhandelingsstrategie rechtstreeks geleerd worden uit historische interacties, zonder expliciet ontwerp van instantiekenmerken of een handmatig gedefinieerde

onderhandelingsstrategieruimte. Dit maakt het mogelijk om door mensen ingebouwde voorkeuren verder te reduceren en te generaliseren naar onderhandelingsproblemen van verschillende omvang en complexiteit. Er zijn nog wel uitdagingen wat betreft het effectief aanpassen van de strategie ten opzichte van verschillende tegenstanders.

In het tweede deel van dit proefschrift gaan we verder dan de ontwikkeling van lerende agenten door de methoden die worden gebruikt voor hun evaluatie kritisch te bekijken. Hoofdstuk 6 presenteert een uitgebreide empirische analyse, waarbij gebruik wordt gemaakt van gegevens van de Automated Negotiation Agents Competition (ANAC) 2022, die we zelf hebben georganiseerd en waarbij we deelnemers specifiek hebben uitgedaagd om lerende agenten te ontwikkelen. De analyse toont de beperkingen van de huidige evaluatiemethoden aan. Ze laat zien dat de rangschikking van agenten afhankelijk is van de keuze van prestatiecriteria, zoals individuele score, gesommeerde groepsscore of speltheoretische evenwichten. Ze toont ook aan dat de veelgebruikte maatstaf, van gemiddelde individuele score, gevoelig is voor de samenstelling van de groep en niet geschikt is voor het beschrijven van niet-transitieve verhoudingen tussen agenten. De conclusie is dat er momenteel geen enkele, universeel robuuste maatstaf bestaat voor de evaluatie van onderhandelingsagenten, met name voor lerende agenten die niet-stationair gedrag vertonen.

Om tegemoet te komen aan de vastgestelde behoefte aan duidelijke evaluatiecriteria en onderzoeksuitdagingen binnen toepassingen, stelt Hoofdstuk 7 multi-agent kalenderplanning voor als een rijk, relevant en complex toepassingsgebied. Deze taak bevat veel uitdagingen voor de automatische onderhandelingen onderzoeksgemeenschap en kan mogelijk toekomstig onderzoek sturen om vooruitgang te boeken buiten de huidige grenzen van deze gemeenschap.

Samengevat levert dit proefschrift methoden voor het ontwikkelen van onderhandelingsagenten die in staat zijn om te leren en hun strategieën aan te passen, waardoor menselijk ingebouwde voorkeuren worden verminderd en de generaliseerbaarheid wordt vergroot. Het biedt ook een kritische analyse van evaluatiemethodologieën in geautomatiseerde onderhandelingen, waarbij de tekortkomingen ervan worden benadrukt en wordt gepleit voor onderzoek naar evaluatiemethoden, mogelijk met behulp van toepassingsgebieden zoals agendabeheer.