

Abstract (EN):

Reinforcement learning comprises a set of techniques that allow a computer system to gradually learn how to solve a particular task at hand. Intrinsically, these techniques relate to very simple but efficient ways people learn on a daily basis. For instance, when parents teach their child how to ride a bike for the first time, they do not provide it a large manual with all the possible situations (and the optimal actions associated with it) that can occur. On the contrary, they simply let the child try and learn from its failures and its successes. Over the years, reinforcement learning has gained more and more attention amongst researchers for solving single-objective problems. However, many real-life problems inherently involve multiple objectives that often give rise to a conflict of interest. An example can be found in the manufacturing industry where the goal of a company is to come up with a policy that, at the same time, maximises the profit and the user satisfaction while minimising the production costs and the environmental impact. For these types of problems, the traditional reinforcement learning framework is insufficient and lacks expressive power.

In this dissertation, we argue the need for particular reinforcement learning algorithms that are specifically tailored for multi-objective problems. These algorithms are multi-objective reinforcement learning (MORL) algorithms that provide one or more optimal trade-off solutions of the problem's original objectives. What this balance intrinsically conveys depends on the emphasis and the preferences of the decision maker. In the case the decision maker's preferences are clear and known a priori, this can be used to guide the search to the single best policy that matches this information. However, in case the preferences are unclear (or even unknown), the problem becomes much more complex and it might be appropriate to present the decision maker with a diverse set of optimal trade-off solutions that each comprise a different balance. In this dissertation, we propose an approach that is specifically tailored to offer a palette of solutions that can then help the decision maker to evaluate and compare the different possibilities.

Abstract (NL):

Reinforcement learning omvat een verzameling van technieken die een computer toelaten om gradueel te leren hoe een taak het best opgelost wordt. Intrinsiek zijn deze technieken nauw verwant met eenvoudige maar efficiënte methodes die mensen op een dagelijkse basis gebruiken om te leren. Bijvoorbeeld, wanneer ouders hun kind willen aanleren om met een fiets te rijden, gaan ze dit kind niet eerst een groot boek laten lezen over alle mogelijke situaties (en bijhorende optimale acties) die kunnen voorvallen tijdens het fietsen. Integendeel, ze laten het kind simpelweg meerdere keren proberen waarna het zal leren uit de fouten en de successen. Reinforcement learning krijgt vanuit de onderzoekswereld de laatste jaren meer en meer aandacht om problemen met een enkel objectief op te lossen. Echter omvatten veel problemen uit de echte wereld meerdere doelstellingen die vaak aanleiding geven tot strijdige belangen. Een voorbeeld hiervan is te vinden in de industrie waar het doel van een onderneming is om te komen tot een beleid dat op hetzelfde moment de winst en de tevredenheid van de gebruiker maximaliseert, terwijl het eveneens de productiekosten en de milieu-impact tracht te minimaliseren. Voor dit soort problemen is het traditionele reinforcement learning raamwerk ontoereikend en ontbreekt het expressiviteit.

In dit proefschrift pleiten we voor de behoefte aan specifieke reinforcement learning algoritmen die speciaal zijn aangepast voor problemen met meerdere objectieven. Deze algoritmen heten vervolgens multi-objectieve reinforcement learning (MORL) algoritmen die een of meerdere optimale compromissen van de doelstellingen verstrekken. Wat dit evenwicht intrinsiek uitstraalt is afhankelijk van de klemtoon en de voorkeuren van de beleidsmaker. In het geval dat de voorkeuren van de beleidsmaker duidelijk en bij voorbaat bekend zijn, kunnen deze gebruikt worden om het zoeken naar het compromis te begeleiden die het best overeenkomt met deze informatie. In het geval dat de voorkeuren van de beleidsmaker onduidelijk (of zelfs onbekend) zijn voor het optimalisatieproces plaatsvindt, wordt het probleem veel complexer en zou het aangewezen zijn om de beleidsmaker een diverse verzameling van optimale oplossingen voor te stellen. In dit proefschrift stellen we een aanpak voor die specifiek is afgestemd om een palet van oplossingen aan te bieden die vervolgens de beleidsmaker kunnen helpen om de verschillende mogelijkheden te evalueren en met elkaar te vergelijken.