



Universiteit
Leiden
The Netherlands

Optimal decision-making under constraints and uncertainty

Latour, A.L.D.

Citation

Latour, A. L. D. (2022, September 13). *Optimal decision-making under constraints and uncertainty*. SIKS Dissertation Series. Retrieved from <https://hdl.handle.net/1887/3455662>

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/1887/3455662>

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

Beknopte samenvatting

Wetenschappers, beleidsmakers en individuen maken dagelijkse beslissingen op basis van onzekerheden en kansen. Vaak zijn er ook beperkingen op het aantal en het type beslissingen dat we kunnen nemen. Zeker als de belangen groot zijn, willen we graag dat de uiteindelijke beslissing die we nemen, optimaal is. Een optimale beslissing is een beslissing die naar verwachting de beste uitkomst heeft, volgens een bepaalde maatstaf van succes.

Kansen in een beslissingsprobleem kunnen bijvoorbeeld komen door willekeur in algoritmes of onzekerheid over menselijk gedrag. Als je iemand volgt op Twitter, zie je hun tweets dan in je timeline? En als je een tweet ziet, stop je dan om te lezen, of scroll je verder? De probabilistische component van een beslissingsprobleem kan zijn oorsprong ook hebben in de chaos en willekeur van natuurlijke processen. Welke hoogspanningskabels blijven heel als ze getroffen worden door een aardbeving? Hoeveel huishoudens komen dan zonder elektriciteit te zitten? In alle toepassingen vinden we ook onzekerheden die veroorzaakt worden door het praktische probleem dat sommige dingen lastig exact te meten of te quantificeren zijn.

Daarnaast moeten we ons vaak houden aan bepaalde beperkingen. In een reclamecampagne is er een limiet aan het aantal mensen dat je kunt bereiken via een advertentie in de krant, omdat niet iedereen die krant leest. We hebben ook geen oneindig budget voor het versterken van onze energie-infrastructuur.

De bovengenoemde voorbeelden bevatten ook relaties tussen entiteiten. Mensen kunnen elkaar volgen op sociale media. Elektriciteitsnetwerken verbinden energiecentrales met huishoudens. De onzekerheid in de relaties in de bovengenoemde voorbeelden maakt dat we die relaties kunnen zien als *probabilistische data*, die we kunnen modelleren met *probabilistische netwerken*.

In dit proefschrift presenteren we nieuwe methodes voor het exact oplossen van optimalisatieproblemen waarin we beslissingen moeten nemen over onzekere, relationele data, en relevante beperkingen moeten respecteren. Deze *stochastic constraint (optimisation) problems (SCPs)* zijn moeilijk op te lossen. Om exact te kunnen te redeneren over het effect van onze beslissingen in een onzekere toekomst, moeten we alle mogelijke, en vaak deels overlappende, scenarios in acht nemen. Het feit dat het aantal combinaties van beslissingen dat we kunnen nemen exponentieel kan groeien met het totale aantal individuele beslissingen, maakt het nog lastiger. In het algemeen zijn SCPs $\mathcal{NP}$ -volledig.

Ons doel is nieuwe algoritmes te ontwikkelen die gemakkelijk zijn toe te passen, en daardoor toegankelijk voor een grote groep potentiële gebruikers, inclusief gebruikers met beperkte programmeerervaring. Daarnaast moeten onze methodes algemeen genoeg zijn om een breed scala aan verschillende SCPs op te kunnen lossen, en dat snel genoeg doen om van praktisch nut te zijn. Om deze drie doelen te kunnen bereiken, bouwen wij verder op bestaande methoden voor het redeneren over kansen, en op technologie die helpt om te zoeken door verschillende combinaties van beslissingen: *constraint programming (CP)* en *mixed integer programming (MIP)*.

We presenteren een nieuwe probabilistisch-logische programmeertaal, SC-ProbLog, die speciaal ontwikkeld is om SCPs mee te modelleren. Deze taal is gebaseerd op ProbLog, een programmeertaal die bijzonder geschikt is om probabilistische relaties tussen entiteiten te modelleren. We laten zien hoe we beperkingen op de kansverdelingen die voortkomen uit probabilistische netwerken kunnen formuleren, en hoe we deze beperkingen kunnen modelleren zodat ze kunnen worden opgelost door CP of MIP software. We maken gebruik van bestaande algoritmes om *beslissingsdiagrammen* te genereren die een representatie vormen van deze kansverdelingen. Deze beslissingsdiagrammen helpen ons om efficiënt te redeneren over kansen en kansverdelingen, tijdens het oplossen van SCPs.

Onze oplosmethoden zijn allemaal modulair van aard. Hierdoor kunnen we voor elke deeltaak de best beschikbare relevante software gebruiken. Dit helpt om onze pijplijnen voor het oplossen van SCPs actueel te houden met de laatste ontwikkelingen, zonder dat we de laatste technieken steeds zelf moeten (her)implementeren, of moeten integreren in een monolithisch software-ontwerp.

We volgen hierbij het principe van *programmeren-door-optimalisatie*, en implementeren veel alternatieve oplossingen voor elke deeltaak in de pijplijnen, die hierdoor zeer configureerbaar zijn. Door deze flexibiliteit is het soms lastig om de juiste component voor elke deeltaak en de juiste instellingen voor elk component te vinden voor het oplossen van een specifiek type probleem. Daarom maken we gebruik van technieken die het configureren van onze algoritmes automatiseren. Dit doen we niet alleen om de kwaliteit van onze methodes op een eerlijke en consequente manier te toetsen, maar ook om te achterhalen welke instellingen goed werken voor welk soort problemen.

Ons werk verlegt de grenzen van de al beschikbare technologie voor het oplossen van SCPs door een nieuwe programmeertaal te introduceren voor het modelleren van SCPs, door te laten zien hoe de beschikbare CP en MIP software gebruikt kan worden om SCPs snel op te lossen, en door nieuwe algoritmes te introduceren voor het oplossen van problemen met beperkingen op kansverdelingen. Hierbij kiezen wij een programmeren-door-optimalisatie-aanpak, en zijn, voor zover wij weten, de eersten die dat doen in het vakgebied van probabilistisch redeneren. We demonstreren de effectiviteit van onze methodes op problemen uit de dataminingliteratuur, in het bijzonder op problemen rondom mond-op-mond reclame en *frequent itemset mining (FIM)*.

