



Foto: Tom van Limpt

MET EEN BEETJE MATHEMATISCHE BESLISKUNDE KOM JE EEN HEEL EIND

JOAQUIM GROMICHO

We zitten nu – februari 2021 – in de tweede lockdown tijdens de voortslepende COVID-19-pandemie. Voor de tweede keer in minder dan een jaar – dat twee opeenvolgende schooljaren beslaat – zitten alle leerlingen de hele dag thuis, met uitzondering van kwetsbare kinderen en kinderen van ouders met vitale beroepen. Een grote beproeving voor de kinderen, de leerkrachten en uiteraard ook voor de ouders. Dat was het ook tijdens de eerste lockdown in het voorjaar van 2020. Toen het kabinet op

21 april 2020 besloot om de basisscholen vanaf 11 mei weer te openen, was iedereen blij, maar voor scholen was het een flinke (wiskundige) puzzel om alles te organiseren. Scholieren zouden ongeveer de helft van hun lestijd weer les op school gaan krijgen, maar wel in kleinere groepen. De andere helft van de tijd zouden ze thuis de lessen volgen.

Zelf heb ik drie dochters op dezelfde basisschool. In het schooljaar 2019–2020 zaten er 553 kinderen op die school,

Hoe een wiskundige de school van zijn kinderen tijdens de lockdown helpt bij maken van een indeling van de klassen waarbij rekening gehouden moet worden met de wensen van de ouders en de leerkrachten.

afkomstig uit 407 gezinnen en verdeeld over 22 groepen. Elke groep moest dus in twee delen worden gesplitst.

Voor elk kind moest het besluit worden genomen: deel A of deel B? Deel A ging naar school op maandag en donderdag en op de oneven weken op woensdag, deel B op dinsdag en vrijdag en op de woensdagen van de even weken.

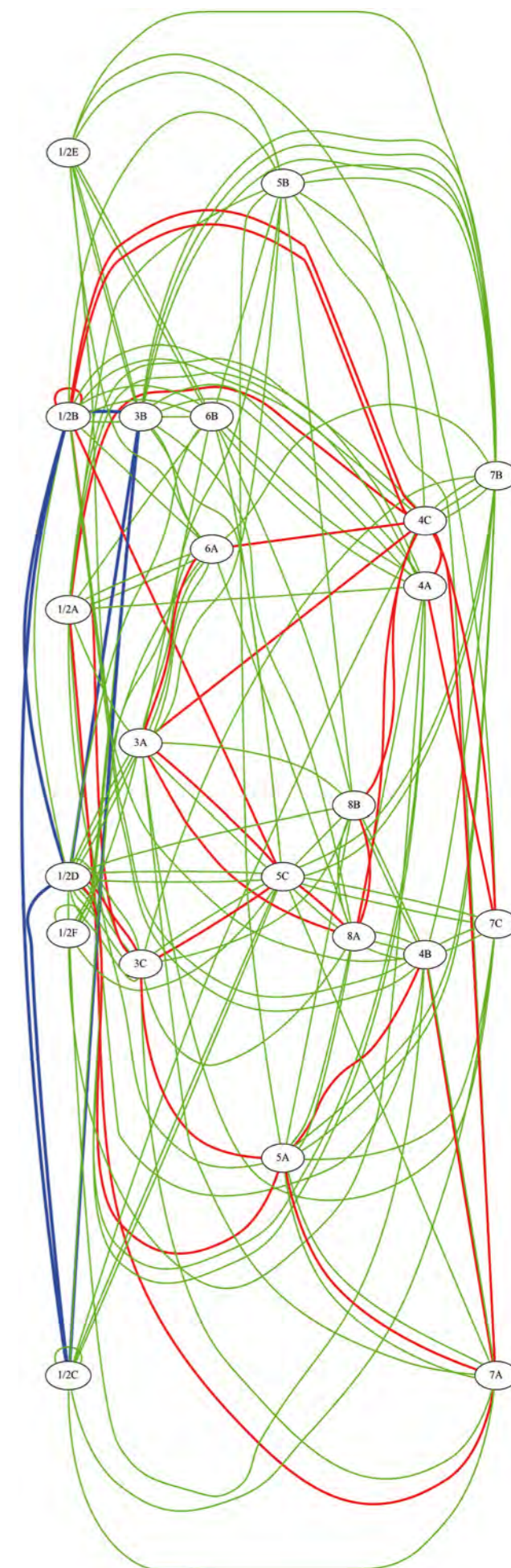
Het was een nadrukkelijke wens van de ouders dat alle kinderen uit hetzelfde gezin op dezelfde dagen naar school zouden gaan, oftewel in hetzelfde deel zou komen. Ook zouden de ouders met vitale beroepen mogen kiezen in welk deel hun kinderen zouden komen. Daarnaast was er de wens van de leerkrachten (en de kinderen zelf) dat de splitsing rekening zou houden met de sociale cohesie onder de kinderen. Ze hadden het al zwaar te verduren, dus dan maar het liefste samen met de hartsvriendinnen en de boezemvrienden. Verder was de wens dat voor groepen met meerdere deeltijdleerkrachten het aantal verschillende leerkrachten per groep per week zo laag mogelijk zou zijn. Uiteraard moesten ook de deelgroepen van gelijke grootte zijn (plus minus een voor oneven aantallen). En laatst maar niet het minst, gesteld dat een leerkracht van risicogroepen vervangen moest worden, dat er ook dan een oplossing met evenveel lesdagen op school voor alle kinderen zou worden gevonden, idealiter zou ieder kind een vol schema A of B krijgen.

Alleen al de wens ‘alle kinderen van hetzelfde gezin op A of op B’ was voor menig schooldirecteur onbegonnen werk om met pen en papier op te lossen. Neem bijvoorbeeld de graaf die voor de groepen van de school van mijn kinderen aangeeft met een groene lijn dat twee kinderen van hetzelfde gezin in de verbonden groepen zitten of met een rode lijn dat er alle drie kinderen van hetzelfde gezin zich over die groepen verdelen en in blauw voor de vier kinderen uit hetzelfde gezin. Dan begrijpen we dat alleen al deze doelstelling de pet van de gemiddelde overbelaste schooldirecteur ver te boven gaat.

Gesneden koek

Maar zo’n puzzel is gesneden koek voor wie een beetje mathematische besliskunde heeft geleerd. Het meest bekende deel ervan, gemengde geheeltallige lineaire optimalisatie, kan deze puzzel namelijk prima oplossen.

Gewapend met een Excelbestand van gezinssamenstellingen, met Python en met mijn kennis van besliskunde kon ik aan alle wensen voldoen en zonder moeite de puzzel meerdere malen opnieuw oplossen, telkens na het ontstaan van nieuwe uit voortschrijdend inzicht geboren wensen.*



Mijn model implementeerde alle wensen als zachte restricties, dit om een toegelaten oplossing te kunnen garanderen. De gewichten waarmee schendingen in de doelfunctie werden beboet waren gekozen overeenkomstig de prioriteit van de wensen. De allerbelangrijkste wens was dat alle kinderen van hetzelfde gezin op dezelfde dagen naar school zouden gaan.

Als belangrijkste beslissingsvariabelen heb ik x_{ij} voor elke leerling i en elk deel $j \in \{A, B\}$ binair genomen: voor elke kind i moet idealiter een van x_{iA} of x_{iB} gelijk aan 1 zijn, en de andere is 0. Om een zachte restrictie te implementeren werd $x_{iA} + x_{iB} \leq 1$ genomen, terwijl de som over alle kinderen werd gemaximaliseerd. Voor elke gezin G worden voor alle kinderen $i \in G$ alle x_{iA} gelijk aan elkaar gesteld en ook x_{iB} . Om de delen zo gelijk aan elkaar mogelijk te krijgen werden twee variabelen per groep gedefinieerd voor de aantal kinderen in de kleinste en de grootste deel respectievelijk en het totale verschil werd geminimaliseerd. Alle andere randvoorwaarden werden op een gelijke manier gemodelleerd.

Dit is de gesimplificeerde code (zie ook de hierbij afgebeelde code) waarbij elk gezin in een deel als harde restrictie wordt gemoduleerd en enkel naar delen van zo gelijk mogelijke grootte wordt geoptimaliseerd. Het hierboven geschetste probleem is binnen 1,5 seconden op te lossen met de gratis CBC-solver (<https://github.com/coin-or/Cbc>).

De volledige puzzel – tegemoetkomend aan **alle wensen** wordt met dezelfde gratis solver in minder dan een minuut opgelost. Om de veranderingen niet te groot te

laten zijn werd, bij elke toevoeging van een nieuwe wens, getracht om de nieuwe oplossing zo gelijk mogelijk aan de vorige versie te krijgen, dat wederom met zachte restricties. Als extra service heb ik, omdat ik toch in de Pythoncode zat, automatisch door middel van PyLaTeX de indelingsbrieven in pdf laten genereren. Geen handwerk, geen knip- en plakwerk, geen fout.

Toen Thomas Davenport in 2012 data scientist het meest sexy beroep op aarde noemde, presenteerden zich in korte tijd op LinkedIn enorme aantallen data scientists. Helaas beschikt slechts een minderheid daarvan over serieuze kennis van wiskunde, statistiek of (nog minder!) mathematische besliskunde. Dat is jammer, want alleen dan kan men dit soort problemen uit het dagelijks leven op een snelle en gemakkelijke manier oplossen.

In deze tijden van lockdown raad ik van harte het geweldige boek *Operationele analyse* (2002) van onze trouwe columnist Henk Tijms aan. Samen met *Pyomo; optimization modeling in Python* van William Hart en Carl D. Laird (2017) komt u een heel eind!

* De medezeggenschapsraad – waarvan ik deel uitmaak – gaf toestemming om de gegevens over gezinssamenstelling onder strikte voorwaarden en enkel voor dit doel te gebruiken. De gegevens zijn anoniem en onherkenbaar verwerkt.

JOAQUIM GROMICHO is sinds januari 2021 hoogleraar Business Analytics aan de Universiteit van Amsterdam. Ook is hij Science & Education Officer bij ORTEC en hoofdredacteur van *STATOR*. E-mail: Joaquim.Gromicho@ortec.com

```
# example data
groups = { 1 : ['a','c'], 2 : ['b','d'] }
families = [ ['a','b'], ['c'], ['d'] ]

import pyomo.environ as pyo

m = pyo.ConcreteModel( 'split' )
m.part = [ 'A', 'B' ]
m.child = {c for f in families for c in f}
m.group = groups.keys()

def familyPairs(m):
    return [ (i,j) for f in families for i,j in zip(f[:-1],f[1:]) if len(f) > 1 ]
m.together = pyo.Set( dimen = 2, initialize = familyPairs )

m.x = pyo.Var( m.child, m.part, within=pyo.Binary )
m.minPart = pyo.Var( m.group, within=pyo.NonNegativeReals )
m.maxPart = pyo.Var( m.group, within=pyo.NonNegativeReals )

m.objective = pyo.Objective( expr = sum( m.maxPart[g]-m.minPart[g] for g in m.group ),
                             sense= pyo.minimize )

m.choose = pyo.Constraint( m.child,
                           rule = lambda m,c : sum( m.x[c,p] for p in m.part ) == 1 )
m.join = pyo.Constraint( m.together, m.part,
                        rule = lambda m,i,j,p : m.x[i,p] == m.x[j,p] )
m.below = pyo.Constraint( m.group, m.part,
                          rule = lambda m,g,p : sum( m.x[c,p] for c in groups[g] ) >= m.minPart[g] )
m.above = pyo.Constraint( m.group, m.part,
                          rule = lambda m,g,p : sum( m.x[c,p] for c in groups[g] ) <= m.maxPart[g] )
```

Code (gesimplificeerd) voor het indelen van schoolklassen met een aantal restricties



VVSOR shows a growing collaboration between disciplines focused on the visibility of statistics and operations research. Despite the troubling times we all faced during 2020, there was a huge increase in the presence of data analysis in society. Never before were phrases like “curve flattening”, “false positive diagnostic tests” and “vaccine efficacy” so widespread.

Several of our members became the face of the scientific response to the pandemic, of the cross-discipline collaboration, data literacy and healthy skepticism. One noteworthy example is prof. dr. Casper Albers. With his columns for the newspaper *de Volkskrant*, Casper Albers encourages his readers to arm themselves with statistics to fight the threat of fake news. In the University of Groningen university newspaper, he empowers both students and employees to be the future of the university. And many news outlets seek his advice, like Nu.nl and *Het Dagblad van het Noorden*, or television shows like Human’s *Medialogica*. Casper is an important voice in the scientific landscape that has argued for critical changes to decrease sloppy research and for a bottom-up approach to scientific progress.

It is with great pride that we announce the candidacy of Casper Albers to be the next president of the VVSOR. At the Annual Meeting of March 18th, if the meeting votes in favour, Casper will succeed Fred van Eeuwijk who served as the VVSOR’s president for four years. Next, Casper gives us a quick introduction.

THE VVSOR DAILY BOARD



Foto: Bob Bronshoff

Letter from the president elect

During the Annual Meeting, the Society can elect me as successor of Fred van Eeuwijk. As the next President, I hope to contribute to the visibility of our society, towards our own members as well as towards society at large. The current daily board made great improvements to the administrative structure of our society. The new board can therefore focus on the future: how do we achieve a steady stream of new, young members and how do we stay relevant for the next ten years or so? I hope to contribute to those challenges.

CASPER ALBERS

Casper Albers received his PhD in Groningen in 2003, under the supervision of Willem Schaafsma. After two postdocs he returned to Groningen in 2009, to join the Faculty of Social and Behavioural Sciences. Here, he holds the chair on Applied Statistics and Data Visualisation. His research focuses on longitudinal models for psychological data and statistics communication. Within VVSOR, Casper has served as the president of the social science section for several years. Outside the VVSOR he fulfilled roles as Research Director of the Heymans Institute for Psychological Research, board member of the International Association for Statistical Computing and as a board member of the Groningen University Fund.