

Beste lezer,

Onlangs ben ik afgestudeerd van de master Computer Science met als specialisatie *Artificial Intelligence*. Ik begon met het programma in collegejaar 2021-2022, waarin ik tien vakken gevolgd heb. Het jaar daarop heb ik een semester gependend aan de Royal Institute of Technology in Stockholm, waarna ik – met enige vertraging – mijn Master's Research Project heb doorlopen. Hoewel ik vanwege een bachelor in Sterrenkunde niet de gebruikelijke voorkennis had, ging de opleiding mij goed af, mede dankzij enkele extra vakken in de informatica die ik tijdens mijn bachelor heb gevolgd.

Gedurende het programma heb ik veel gereflecteerd op het onderwijs, specifiek de vakken die ik heb gevolgd – het Master's Research Project is natuurlijk voor iedereen anders. Uiteraard heb ik bij elk vak het evaluatieformulier (uitvoerig) ingevuld, en daarnaast ook een evaluatiebijeenkomst van de opleidingscommissie bijgewoond, waar ruimte was om feedback te geven op de vakken van het afgelopen semester. Wat ik hierbij echter heb gemist is de mogelijkheid om feedback te geven op de opleiding als geheel, om opmerkingen te kunnen maken over vakoverstijgende punten. Daarom schrijf ik u deze brief: opdat ik mijn gedachten over het pakket mastervakken kan delen met iemand die wellicht iets met deze informatie kan.

Natuurlijk begrijp ik dat het perspectief van de studenten bij lange na niet het enige is wat in acht genomen wordt bij het ontwerpen van zo'n programma, en dat de kritiekpunten die ik aanstip niet zomaar op te lossen zijn. Toch wil ik ze onder de aandacht brengen. Ik stel me voor dat veel studenten ze niet snel opbrengen, aangezien ze mogelijke verbeteringen toch niet zelf zullen ervaren en verbeteringen juist ook nadelig kunnen uitpakken voor sommige studenten. Echter denk ik dat zowel de opleiding, het personeel alsook de aanstaande studenten er baat bij zullen hebben als ze opgepakt zullen worden.

Mijn feedback betreft de volgende drie punten, welke ik hieronder verder toelicht

- Disfunctionele groepen bij projecten
- De gebruikelijke beoordelingsmethoden
- De onderwijsvormen

Disfunctionele groepen bij projecten

Als student die niet de bachelor Informatica als vooropleiding had, was het niet vanzelfsprekend dat ik zou worden toegelaten tot de master. Met de extra vakken die ik had gevolgd, kwam ik echter net aan de 90 EC informatica-gerelateerde vakken vereist voor toelating, en kon ik – na goed overleg met de studieadviseur – starten zonder een Pre-Master te hoeven doen. Een groot aantal studenten komt ofwel van een andere opleiding, ofwel van een informatica-opleiding van een andere universiteit (met name internationale studenten).

Hoewel een toegangseis van 90 EC informatica-gerelateerde vakken geen specifieke verwachtingen stelt van wat een student wel en niet moet kunnen voor de opleiding, neem ik aan dat de organisatie achter de opleiding hiermee wel uit gaat van een bepaald profiel van de student. Zelf zou ik verwachten dat studenten voldoende programmeerervaring hebben, zowel zelfstandig als in een team moeten kunnen werken, en verantwoordelijkheid moeten kunnen nemen voor hun werk. Gedurende de opleiding heb ik echter veel voorbeelden gezien van studenten waarbij dit geenszins het geval is, en dit is iets waar andere studenten veel last van ondervinden.

Een groot deel van de beoordeling van de vakken berust namelijk op groepsopdrachten. In vrijwel alle groepsopdrachten die ik heb moeten doen, kwam ik met studenten in de groep die niet of nauwelijks hebben bijgedragen aan de opdracht. Dit probleem speelt minder voor mensen die weten met wie goed samen te werken valt, maar in grotere groepen zit je vaak met studenten die je niet kent. Zulke groepsopdrachten zijn in mijn ogen het meest stressvolle onderdeel van de opleiding, om drie redenen.

Als eerste komt er natuurlijk veel meer werklast te liggen bij de mensen die wel hun best willen doen voor de opdracht. Een opdracht die ontworpen is voor vijf studenten, maar slechts door drie wordt uitgevoerd, zorgt ervoor dat er veel meer gedaan moet worden door die drie mensen. Ten tweede kom je er niet meteen achter dat mensen niet van plan zijn een serieuze bijdrage te leveren. Soms reageren ze in het begin enigszins enthousiast, maar blijkt later dat ze geen tijd vrijmaken voor de opdracht, en soms zeggen ze bezig te zijn met een onderdeel, maar blijkt na een tijdje helemaal niks gedaan te zijn. Het extra werk dat de rest van de groep dan moet doen komt dan pas in een later stadium aan het licht, waardoor die extra last ook niet evenredig over de duur van het project verdeeld wordt. Een derde bron van stress is het ontbreken van een opleidingsbreed gedragen procedure of meldpunt voor deze problematiek. Elk vak gaat op een andere manier om met slecht functionerende groepen: een lager cijfer voor de leden die weinig hebben bijgedragen, een aanvullende opdracht voor die groep, of überhaupt geen consequenties. Naast het feit dat dit gebrek aan procedure alleen maar leidt tot veel mailverkeer tussen docenten en de getroffen studenten in een zoektocht naar een geschikte oplossing, levert geen van deze oplossingen daadwerkelijk een verlichting van de werklast op voor de studenten van wie nu meer werk gevraagd wordt. Daarnaast heerst het gevoel dat de niet-werkende studenten 'er makkelijk mee weg komen', aangezien zelfs voor niets toevoegen aan een project de consequenties doorgaans mild zijn.

Hoewel het fijn is dat de meeste docenten erg meedenkend zijn in dit soort situaties, haal ik uit mijn eigen ervaringen bij verscheidene vakken én de ervaringen van medestudenten dat disfunctionele groepen bij groepsopdrachten een structureel probleem zijn binnen de opleiding. Het probleem ligt bij de studenten die niet meewerken, vaak ofwel omdat ze er toch wel mee weg kunnen komen (de rest van de groep doet het werk toch wel, en het blijkt voor docenten lastig in zo'n situatie de student de studiepunten te ontfangen), ofwel omdat ze niet de kennis hebben om bij te dragen aan het project (bij een opdracht voor *Modern Game AI Algorithms* in het tweede semester gaf een student aan niet te kunnen programmeren in Python, en bij *Robotics* was er iemand die geen ervaring had met Git). Ik kan me goed voorstellen dat sommige studenten in de eerste groep ook behoren tot de tweede groep: ze hebben niet de capaciteiten om bij te dragen, maar zonder dat aan te geven laten ze gewoon niet van zich horen en hopen ze er met een voldoende uit te komen.

Ondanks het beeld dat ik hier schets heb ik ook enkele goede samenwerkingen gehad: wanneer je een goede student vindt, blijf je snel opdrachten met die persoon samen doen. Ik denk dat het mogelijk is om het percentage goedlopende samenwerkingen te verhogen, maar dan moet er enerzijds wat gedaan worden aan de omgang met slecht functionerende groepen, en anderzijds met de toegangseisen van de opleiding of de vakken. Je kan het vak *Modern Game AI Algorithms* niet halen als je niet kan programmeren in Python, dus zo'n student zou dat vak niet mogen doen. Overigens denk ik niet dat de master als geheel te voltooien is zonder kennis van Python, dus zou ik er zelf voorstander van zijn als de eisen voor de opleiding strakker worden geformuleerd én beter worden gehandhaafd. Ik denk dat dit al een zeer positief effect kan hebben op het functioneren van de groepen. Wanneer er daarnaast duidelijke repercussies zijn voor studenten die zich niet inzetten in groepsopdrachten, kan dit probleem verder geminimaliseerd worden.

Dergelijk beleid is op meerdere vlakken in het voordeel voor de opleiding. Het stuklopen van groepsopdrachten leidt er namelijk toe dat studenten vakken niet halen, ofwel omdat ze vanwege gebrek aan bijdragen uit een groep worden gezet, ofwel omdat ze wel hard werken maar door de extra werklust in de knel komen met de deadline of projecten bij andere vakken. Beter beleid zorgt voor minder misgelopen studiepunten, met als gevolg meer studenten die (nominaal) de opleiding weten af te ronden. Dit is goed voor het imago en de statistieken van de opleiding, voor het aantal studenten dat een vak opnieuw moet volgen (wat alleen maar extra werklust is voor personeel), en natuurlijk voor de financiën van de opleiding, omdat er voor hetzelfde onderwijs meer studiepunten en dus meer inkomsten opgehaald worden. Daarnaast zal beter beleid leiden tot een betere mentale gezondheid van de studenten, omdat die minder vaak zullen worden opgezadeld met twee keer zo veel projectwerk als gevraagd. Hoewel ik de vele initiatieven en investeringen vanuit de universiteit rondom het verbeteren van mentale gezondheid toejuich, denk ik dat dit soort structurele verbeteringen in het onderwijs meer effect zullen hebben.

De gebruikelijke beoordelingsmethoden

Zoals hierboven al duidelijk is geworden, zijn groepsprojecten in de huidige vorm mijns inziens geen goede beoordelingsmethode: juist de gemotiveerde studenten komen onder onnodig veel druk te staan. Toch krijgen alle leden van de groep vaak hetzelfde cijfer, of zijn er slechts marginale verschillen die niet representatief zijn voor de inspanning of inzicht dat geleverd is. Echter vind ik ook andere veelgebruikte beoordelingsmethoden niet toereikend, en daar wil ik dit deel van de brief aan wijden.

De beoordeling van elk vak heeft – volgens de onderwijs- en examenregeling (OER) – als doel het inspecteren en evalueren van de vaardigheden van een student wat dat vak betreft. Hoewel het niet als zodanig in het OER staat, zou ik zeggen dat welke vaardigheden dat zijn nader omschreven wordt bij de leerdoelen van dat vak in de studiegids. In mijn ervaring worden veel van die leerdoelen niet of nauwelijks getoetst.

Dit komt grotendeels omdat de meeste vakken een groot eindproject hebben waarin je voornamelijk erg intensief bezig bent met één specifiek onderwerp binnen het vak. Dit zie je vooral wanneer je een project moet opzetten rondom een bepaalde publicatie, wat voor mij het geval was bij *Automated Machine Learning*, *Urban Computing* en *Computational Imaging and Tomography*. Het geheel van papers is een goede weerspiegeling van wat je aan het eind van het vak dient te begrijpen, maar dit gaat verloren wanneer je slechts één paper onder ogen krijgt. Veel van de lesstof wordt dan niet getoetst. Hoewel *Urban Computing* nog twee tussentijdse opdrachten heeft waar een deel van de lesstof in terugkomt, is de beoordeling bij *Automated Machine Learning* alléén gebaseerd op de presentatie en project rondom één paper, terwijl ongeveer twintig belangrijke papers worden behandeld bij het vak. *Computational Imaging and Tomography* is vergelijkbaar: daar presenteert je één paper en maak je een project over een ander, maar ook dat dekt maar een klein deel van alle stof.

Een vergelijkbaar probleem speelt soms ook met kleinere tussentijdse opdrachten, die slecht aansluiten op de lesstof, waardoor de toetsing ook niet overeenkomt met de leerdoelen van het vak. Dit hoeft echter niet erg te zijn als het eindproject of tentamen wel goed de leerdoelen omvat. De opdrachten van bijvoorbeeld *Introduction to Deep Learning* en *Computational Creativity* sloten niet goed aan op de colleges. Toegegeven, *Computational Creativity* had ‘tutorial’ opdrachten die beter aansloten en die volgens de studiegids ook onderdeel waren van de beoordeling, maar uiteindelijk kreeg iedereen een 10 voor dat onderdeel ongeacht of ze iets ingeleverd hadden.

Een ander punt, dat zeker in mijn specialisatie van *Artificial Intelligence* speelt, is dat opdrachten (vaak onbedoeld) op zo'n manier ingericht zijn dat er veel tijd moet worden besteed aan triviale zaken zoals de *set-up* van een project of het trainen van een model. Een van de opdrachten bij *Introduction to Deep Learning* hield in dat een neurale netwerk getraind moest worden op een dataset van 80 000 afbeeldingen. Dit duurt altijd al lang, maar daar kwam in dit geval ook bij dat de opdracht moest gebeuren in Google Colab, een online runtime voor Python notebooks. Deze zijn niet bedoeld voor zulke intensieve berekeningen, waardoor de runtime vaak wordt afgebroken. Hierdoor wordt het grootste deel van de tijd dat je aan een vak besteed wachtend doorgebracht, in plaats van op zo'n manier dat de stof goed overkomt. Bij *Robotics* werd (uiteraard) met hardware gewerkt, en mijn groepje was veel tijd met het opzetten van het project omdat zowel hardware als software enigszins verouderd was, waardoor je minder tijd bezig bent met vakinhoudelijke aspecten.

In mijn ogen is een tentamen een van de meer geschikte manieren op vaardigheden te toetsen: je loopt niet zoals bij groepsopdrachten het risico dat het cijfer niet representatief is voor de individuele studenten, en het biedt de kans om alle deelonderwerpen van een vak aan bod te laten komen. Ik had bij vier van de tien vakken een tentamen, wat over het algemeen goed was uitgevoerd. Alleen bij *Introduction to Deep Learning* waren ongeveer de helft van de vragen 'gerecycled' uit de oefententamens die de docent zelf beschikbaar had gesteld. Daarbij was het tentamen zo kort dat je al na een half uur klaar kon zijn, terwijl er drie uur voor was ingepland.

Het overkoepelende kritiekpunt hier houdt dus in dat sommige methoden op zo'n manier worden ingezet dat ze niet tot een goede beoordeling zullen leiden wat de leerdoelen van de vakken betreft. Gelukkig zijn er uitzonderingen. Het vak *Information Retrieval* had in mijn ogen een erg goede opzet: kleine wekelijkse opdrachten om ervoor te zorgen dat je bezig blijft met de stof, een review van een paper om te kijken of je publicaties uit het veld kan begrijpen, een programmeeropdracht om je technische vaardigheden te beoordelen, maar met de nadruk op een tentamen (60% van het eindcijfer) waarin alle stof de revue passeert. Ik heb weinig weet van hoe de beoordelingsmethoden bepaald worden, maar bij sommige vakken zou ik sterk voorstander zijn van een herijking. Op deze manier kan voorkomen worden dat mensen slagen voor vakken zonder de leerdoelen te beheersen. Naast het feit dit nu voor de betreffende studenten tot nadelige gevolgen kan leiden in een verdere (academische) loopbaan, is het in mijn ogen ook de plicht van de onderwijsinstelling om onderwijs aan te bieden dat zo is ingericht dat het behalen van de leerdoelen gegarandeerd is bij het succesvol afronden van een vak. Daarnaast kunnen ook hier weer interventies leiden tot een beter imago van de opleiding, omdat studenten meer het idee zullen krijgen de inhoud daadwerkelijk te beheersen wanneer ze er adequaat op getoetst zijn.

De onderwijsvormen

Dat de beoordeling van veel vakken naar mijn inzicht doorgaans niet goed aansluit op de leerdoelen, zou een kleiner probleem zijn als die doelen in het onderwijs zelf wél goed aan bod komen. Echter zie ik ook op dit punt systemische problemen. Van mijn drie punten ligt deze het meest gevoelig: het eerste betreft voornamelijk de studenten en het tweede slechts de inzet van beoordelingsmethoden, maar kritiek op de onderwijsmethoden gaat al snel over docenten en ander personeel dat betrokken is bij de ontwikkeling van het vakaanbod.

Om die reden wil ik nogmaals benadrukken dat deze brief is geschreven met als doel constructieve feedback te geven over hoe een student deze opleiding ervaart. De mening van studenten is noch de enige noch de belangrijkste factor die van belang is als het om onderwijsontwikkeling gaat, en

zodoende erken ik dat niet alle problemen die hier aan bod komen haalbare oplossingen hebben in het huidige ecosysteem waar de opleiding deel van is.

De kern van mijn kritiekpunt wat onderwijsmethoden betreft ligt bij een probleem dat aan de 'achterkant' van onderwijsontwikkeling waarschijnlijk een oude bekende is: het vinden van een vakdocent die aan alle eisen van een goede vakdocent voldoet:

- Expertise in het vakgebied
- Goede didactische vaardigheden
- Voldoende beschikbaarheid

Voornamelijk afgaande op wat docenten zoal over zichzelf vertellen bij het eerste college van hun vak, ga ik ervan uit dat het wat expertise betreft over het algemeen goed zit. De andere twee punten lijken echter vaak in het geding te komen, en gecombineerd worden de problemen alleen maar groter. Enerzijds heb ik legio docenten gezien die – ondanks dat het gerenommeerde academici zijn – het niet lukt om overtuigend de lesstof over te brengen aan de studenten, anderzijds zijn er veel docenten die ogenschijnlijk niet de tijd hebben om genoeg aandacht te geven aan hun vak.

Overtuigende kwalitatieve of kwantitatieve argumenten aandragen om het gebrek aan didactische vaardigheden aan te tonen is lastig: het beperkt zich voornamelijk tot algemeenheden zoals docenten die de Engelse taal niet goed machtig zijn of die weinig meer doen dan (delen van) powerpoint slides voorlezen, waardoor colleges in de praktijk makkelijk te vervangen zijn door een snelle scan van het Wikipedia-artikel over het onderwerp van die week of – de effectievere optie tegenwoordig – door ChatGPT te vragen de concepten uit te leggen.

Wanneer docenten daarnaast (of in plaats daarvan) weinig tijd hebben om de colleges of het vak als geheel goed voor te bereiden, heeft dat ook vergaande gevolgen. Mijn beste voorbeeld is *Evolutionary Algorithms*, waarbij alle theorie die wordt besproken gepubliceerd is vóór 2003. Het was bij dit vak overduidelijk dat het nauwelijks was gereviseerd in de minstens tien jaar dat het al gegeven werd. Docenten hebben het te druk om kritisch naar de stof te kijken, met als gevolg dat zij de slides tijdens het college ook weer voor het eerst zien sinds het jaar daarvoor.

Een manier waarop dit probleem lijkt te worden opgelost is om überhaupt geen docenten meer college te laten geven. Bij zowel *Automated Machine Learning* als *Computational Imaging and Tomography* staan bij de meeste colleges studenten voor de klas om papers te presenteren aan medestudenten. Hoewel dit wat mij betreft prima een klein onderdeel kan vormen van het onderwijs (ook hier scoorde *Information Retrieval* goed met korte presentaties in break-out rooms) moet het niet de hoofdmoot vormen. Ondanks dat de didactische vaardigheden verrassend vaak beter zijn – wellicht omdat ze beter voorbereid zijn en er nog op beoordeeld worden – is hun expertise niet van hetzelfde niveau en is een college over concepten en hun context ook niet een-op-een vervangbaar door een presentatie van een specifiek paper. Wanneer daarnaast de gepresenteerde inhoud niet wordt getoetst (wat bij *Automated Machine Learning* en *Computational Imaging and Tomography* inderdaad zo was) is het mij onduidelijk hoe studenten toch geacht worden de leerdoelen van dat vak eigen te maken.

Ik stel me voor dat docenten deze aanpak (uiteraard onder het mom van iets anders) kiezen opdat het veel voorbereidend werk scheelt, en anders vormen colleges met gerecyclede slides van het jaar daarvoor een goed alternatief. Ik weet weinig over het proces achter onderwijsontwikkeling, maar de heersende onderwijsmethoden wekken een sterke indruk dat beperkte financiële en personele middelen de belangrijkste factoren zijn achter die keuze. Eenzelfde indruk bestaat kijkende naar de beoordelingsmethoden: groepsopdrachten beoordelen gaat een stuk sneller dan individuele

tentamens nakijken. Het contrast is erg groot met vakken waar docenten wel voldoende tijd steken in het onderwijs, zoals *Information Retrieval* en *Computational Creativity*. Naast de gevolgen die het heeft voor het (niet) overbrengen van de lesstof, blijven studenten hierdoor weg van colleges en worden degenen die er wel zijn niet geprikkeld om actief mee te doen.

Al het bovenstaande is het gevolg van een lange periode van kritisch nadenken over het onderwijs dat ik zelf heb ervaren gedurende mijn master. Het is bij lange na niet een complete weergave van mijn gedachten, maar door drie rode draden uit mijn ervaringen te halen heb ik wel het idee hier de belangrijkste aandachtspunten behandeld te hebben. Ik heb veel nuttige gesprekken gehad met vrienden hierover, en zodoende sta ik ook meer dan open voor een gesprek om dieper in te gaan op deze materie, zodat ik aan de hand van meer concrete voorbeelden mijn punten kan toelichten. Ik hoor graag of hier interesse voor is.

Met vriendelijke groet,

Wessel van Dam