Speltheorie voor beginners: Over luie studenten, groepsdynamica en pokeren

Door Michael Muskulus

Speltheorie is een interdisciplinaire onderneming met als doel het opstellen en analyseren van menselijk gedrag in situaties waar beslissingen moeten worden genomen. Zoals elke goede theorie is ook de speltheorie filosofisch gemotiveerd. Uitgangspunt is de aanname dat mensen de volgen van hun handelingen kunnen inschatten en beoordelen met betrekking tot hun wenselijkheid, en dat ze de voor hun meest voordelige uitkomst kiezen.
Het centrale concept in de speltheorie is de wenselijkheid, in het engels *utility* genoemd. De filosoof Jeremy Bentham beschouwde deze als het plezier dat onze handelingen opleveren [1]. In het *utilitarism* van Bentham en John Stuart Mill wordt de maxime van het maximaliseren van de utility toegepast op moralistisch en politiek vlak. In de economie werd dit idee populair door Léon Walras, die een consument beschouwde als een automaat die probeert een utility functie te maximaliseren onder beperkingen die hem door zijn omgeving opgelegd zijn. De utility in economische toepassingen is meestal gelijkgesteld aan de monetaire winst, dat wil zeggen het bedrag aan geld dat een mens door zijn handelingen wint, omdat geld, tenminste in theorie, een universeel ruilmiddel is. Daardoor ontstaan gemakkelijk paradoxen, bijvoorbeeld wanneer wij altruïstisch gedrag willen verklaren. Juist daarom moeten wij utility in zijn oorspronkelijke, ethische dimensie opvatten, namelijk als de winst voor het *goede leven* van een mens [2], waarbij ook esthetische, culturele en sociale aspecten een rol spelen.

Strategische spellen

Omdat mensen niet alleen op de wereld zijn, kan niet iedereen zijn utility zomaar maximaliseren: vaak hebben mensen tegenstrijdige belangen, een situatie die door het argeloos klinkende woord *spel* wordt omschreven.
Wiskundig wordt een spel gemodelleerd door een verzameling van mogelijke acties voor elke speler, en een payoff functie die de utility van de uitkomst voor elke speler beschrijft. Deze functie is afhankelijk van de keuzes van alle spelers. In het geval van een tweepersoons spel kunnen wij de payoff functie van elke speler als een matrix weergeven, met de mogelijke acties van de eerste speler als rijen, en de acties van de tweede speler als kolommen. Verder wordt aangenomen dat utility numeriek meetbaar is, en de elementen van de matrix reële getallen zijn, zodat hun orde overeenkomt met de orde van voorkeuren die elke speler aan de uitkomsten geeft. De eenheid van de utility wordt ook *util* genoemd.
In Figuur 1 is een tweepersoons spel te zien dat ik *werken aan een gemeenschappelijk project* zal noemen, Martin Osborne volgend [3]. Twee studenten moeten samen een werkstuk voor een college indienen. De afspraak is dat iedereen vandaag in zijn eentje werkt, en dat de volgende dag alleen nog maar de resultaten moeten worden samengevoegd. Elke student wil graag een goed cijfer, maar het liefst zonder hard te werken. Eenvoudig gemodelleerd heeft elke student een keuze tussen twee acties: hard werken (H) of helemaal niets doen (N). De beste uitkomst voor elk student is dat hij niets doet, maar zijn partner wel. De slechtst mogelijke uitkomst is dat hij hard werkt, maar zijn partner niet, omdat hij niet wil worden uitgebuit. Wat is de rationele keuze voor beide studenten?
De wiskundige John Nash, de tragische held van Sylvia Nasars boek “A beautiful mind” redeneerde als volgt: Stel dat een speler weet wat zijn partner gaat doen. Dan kan hij een *beste antwoord* kiezen. Bij voorbeeld, als zijn partner hard werkt, zal hij zelf niets doen. Maar als zijn partner niets doet, dan doet hij ook niets, en dus wordt hij ook niet uitgebuit. Een *Nash equilibrium* is nu een strategische keuze voor elke speler, zodanig dat de strategie van elke speler het beste antwoord is op de acties van alle andere spelers. Zoals Ken Binmore heeft aangemerkt verlost een Nash equilibrium ons van de circel die begint met „ik denk dat hij denkt dat ik denk dat...“ [4].
In het geval van ons voorbeeld is helaas de strategie (N, N) het unieke Nash equilibrium. Zou dit de reden zijn waarom het vaak zo moeilijk is om productief met andere studenten samen te werken?

Gemengde strategieën

In Figuur 2 is een spel te zien dat *matching pennies* heet, waarbij twee spelers allebei óf kop (K) óf munt (M) kiezen. Als zij hetzelfde zeggen, wint de eerste speler een util van de tweede, en anders wint de tweede een util van de eerste. Het is duidelijk dat de enige manier voor een speler om niet een heleboel geld te verliezen is om *onvoorspelbaar* te zijn. De actie die een speler kiest moet dus door een element van toeval bepaald zijn. Wiskundig wordt dit gemodelleerd door een kansverdeling over de acties; wij noemen dit een *gemengde strategie*, in tegenstelling tot de *pure strategieën* (K) en (M). Ook voor gemengde strategieën bestaat het concept van een Nash equilibrium, waarbij de deterministische uitkomst door een verwachting moet worden vervangen.
Stel dat de eerste speler de gemengde strategiee (p,1-p) speelt, dat wil zeggen, dat hij elk keer met een waarschijnlijkheid van p>0 kop kiest, en munt met een kans van 1-p. De tweede speler speelt een gemengde strategie (q,1-q). Stel verder dat (p,q) een Nash equilibrium vormen, zodat er voor geen van beide spelers een reden is om van zijn strategie af te wijken. Meer nog, in een

Nash equilibrium is de verwachting van de pure strategieën voor elke speler gelijk: anders zou het namelijk beter zijn om voor een van de pure strategieen te kiezen in plaats van de gemengde [5]. Dit zogenoemde *neutraliteitsargument* leidt tot

$$\mathbf{E}[K] = pq - p(1-q) = \mathbf{E}[M] = (1-p)(1-q) - (1-p)q,$$

wat q=1/2 als oplossing heeft. Op analoge wijze zien wij dat ook p=1/2 optimaal is, en het unieke Nash equilibrium van *matching pennies* is dus (1/2, 1/2).

Groepsdynamica

Het volgende spel wordt vaak *reporting a crime* of *the good samaritan* genoemd. Er zijn N spelers die allen getuigen zijn van een crimineel incident. Iedereen heeft er baat bij dat de politie wordt gebeld omdat daarmee het slachtoffer wordt geholpen. Aan de andere kant geeft iedereen er de voorkeur aan dat iemand anders belt, omdat het toch wat tijd en moeite kost. Elke speler heeft dus de keuze tussen de twee acties *bellen* (B) of *niets doen* (N). De payoff van een speler is 0 als niemand belt, een beloning van $v>0$ als tenminste één andere speler belt, en een beloning van $v-c>0$ als hij belt – en eventueel ook nog iemand anders. Met een neutraliteitsargument kunnen wij, als zoëven, een gemengd Nash equilibrium vinden.

Laten we de kans waarmee een van de spelers belt p noemen. De kans dat niemand van de N-1 andere spelers belt is $(1-p)^{N-1}$. Dan is

$$\begin{aligned}\mathbf{E}[B] = v-c = \mathbf{E}[N] &= v\cdot P[\text{iemand anders belt}] \\ &= v\cdot(1-P[\text{niemand anders belt}]),\end{aligned}$$

wat uiteindelijk $p = 1 - (c/v)^{1/(N-1)}$ oplevert. We zien dat wanneer het aantal spelers toeneemt, de waarde van p afneemt. De kans dat niemand de politie belt is $(1-p)^N = (c/v)^{1+1/(N-1)}$, en deze neemt asymptotisch af tot een waarde van c/v.

Hoe kunnen we dit resultaat nu interpreteren? Neem bijvoorbeeld v = 10 en c = 1, dan wordt bij een voldoende aantal spelers in één van tien gevallen de politie *niet* gebeld. Wat dit in de praktijk voor gevolgen kan hebben, laat ik aan de verbeelding van de lezer over.

Pokeren

De laatste tijd is pokeren erg populair geworden in Nederland, vooral onder jonge mensen en op internet [6]. Ook hier kan natuurlijk speltheorie toegepast worden, maar het kan snel ingewikkeld worden [7]. Het eenvoudige model dat ik nu zal beschrijven is te vinden in een mooi boek van Edward Packel [8] en levert toch al interessante inzichten op.

Stel dat we aan het pokeren zijn met een groep spelers. Het spel heeft een cruciaal punt bereikt: behalve ons is er slechts één tegenstander over, alle anderen zijn niet meegegaan met inzetten. Het is de laatste ronde van bieden en er zijn twee mogelijkheden. Of wíj hebben de betere kaarten, of onze tegenstander. Door onze ervaring in pokeren weten wij zelf zeker wat er aan de hand is, maar onze tegenstander niet. Zo’n situatie komt vaak voor: dat wij tot de laatste ronde een bepaalde kaart nodig hebben voor een winnende hand, maar zonder deze een waardeloze hand hebben, terwijl we redelijk zeker weten wat de tegenstander heeft. Ter illustratie: wij wachten met slechts één *high card* op een *flush* of een *straight*, en onze tegenstander heeft tenminste een *pair*, maar zeker geen *straight* of beter.

Aan de andere kant, stel dat onze tegenstander de kans p die wij op een winnende hand maken goed kan schatten. Laat k de inzet op tafel zijn waar om gespeeld wordt, de zo genoemde *pot*, en laat r het maximaal mogelijke bod in deze laatste ronde zijn. De laatste kaart wordt uitgedeeld en nu moeten we kiezen om het bedrag r in te zetten, of niet. In het laatste geval zal de tegenstander zeker zelf inzetten, en dan laten wij de pot dus aan hem over. In het eerste geval kan de tegenstander kiezen of hij meegaat, waardoor de pot uiteindelijk naar de betere hand gaat (dus met kans p naar ons), of hij kan *folden*, waardoor de pot zeker naar ons gaat.

Omdat wij zeker weten hoe goed onze hand is, lijkt het verstandig in te zetten als wij de betere hand hebben, en zo niet de pot aan de tegenstander over te laten. Maar dat zou näief zijn. Want als onze tegenstander op gegeven moment door heeft dat wij deze strategie spelen, zal hij nooit meer meegaan als wij bieden. Door zo voorspelbaar te zijn maken wij in zo’n situatie een gemiddelde winst van $E_0 = pk$, maar in feite kunnen wij veel meer winnen door soms te *bluffen*, dat wil zeggen, af en toe toch met slechte kaarten verder te spelen. Op die manier verrassen wij onze tegenstander en verhogen *zijn* risico.

Stel dus, dat wij met een kans $x > 0$ bluffen, en onze tegenstander dan met een kans van $y > 0$ mee gaat. In Figuur 3 zijn de verwachte payoffs voor elk mogelijke combinatie van de pure strategieën te zien. Ook hier kunnen wij weer door middel van een neutraliteitsargument een gemengd Nash equilibrium vinden. Met behulp van wat algebra is de oplossing

$$x = p\cdot r \,/\, ((1-p)(k+r)), \qquad y = k \,/\, (k+r),$$

en onze verwachte payoff is dan $E_1 = p\cdot k\cdot(k+2r)/(k+r)$, wat groter (!) is dan E_0.Voor elke andere waarde van x kan onze tegenstander onze payoff verkleinen. Andersom, als hij een andere waarde dan y kiest, kunnen wij onze payoff vergroten tot een maximaale waarde van k (als wij altijd bluffen omdat hij nooit met ons meegaat).

De take-home message hier is dus: bluffen is essentieel om bij het pokeren te winnen. Meer nog, de optimale waarde van x geeft ons (in principe) de optimale frequentie van bluffen.

[1] J. Bentham. *An Introduction to the Principles of Morals and Legislation.* Dover (2007).

[2] Aristotle. *The Nicomachean Ethics.* Penguin Books (2004).

[3] M. J. Osborne. *An Introduction to Game Theory.* Oxford University Press (2004). Het spel *werken aan een gemeenschappelijk project* is een vorm van het bekende *prisoner's dilemma*.

[4] K. Binmore. *Game Theory. A Very Short Introduction.* Oxford University Press (2007).

[5] Dit is een gevolg van de lineariteit van de verwachting.

[6] Veel spelers in Nederland beseffen niet dat de deelname aan poker toernooien buiten het aanbod van Hollands Casino, bijvoorbeeld op internet, illegaal is. Een officiële reden hiervoor is het gevaar van verslaving. Voor meer informatie zie het volgende rapport:

J. Franssen, E.-J. Koning en C. Kolar. *Het gezicht van Poker. Onderzoek naar Poker in Nederland.* College van toezicht op de kansspelen, Den Haag (2007). http://www.toezichtkansspelen.nl/verslagen/Het_gezicht_van_Poker.pdf.

[7] B. Chen en J. J. Ankenman. *The Mathematics of Poker.* ConJelCo (2006). Zie ook de website van de computer poker onderzoeks groep van de University of Alberta: http://www.cs.ualberta.ca/~games/poker.

[8] E. Packel. *The Mathematics of Games and Gambling.* The Mathematical Association of America (2006).

[9] R. Seward. *True Random Number Generator.* http://robseward.com/itp/adv_tech/random_generator/.

	Hard werken	Niets doen
Hard werken	$(2,2)$	$(0,3)$
Niets doen	$(3,0)$	$(1,1)$

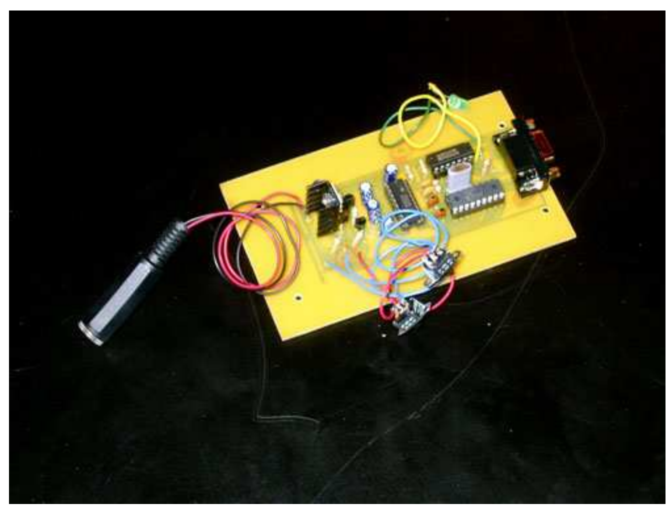
Figuur 1: Payoff matrix voor het spel werken aan een gemeenschappelijk project*. De twee getallen in elk vak zijn de payoffs voor de eerste en de tweede speler,achtereenvolgend.*

	Kop	Munt
Kop	$(1,-1)$	$(-1,1)$
Munt	$(-1,1)$	$(1,-1)$

Figuur 2: Payoff matrix voor het spel matching pennies.

	Altijd Meegaan	Nooit Meegaan
Nooit Bluffen	$p(k+r)+0$	$pk+0$
Altijd Bluffen	$p(k+r)+(1-p)(-r)$	$pk+(1-p)k$

Figuur 3: Verwachte payoffs voor de eerste speler in het poker model, voor elk combinatie van pure strategieën.



Figuur 4: Een toevalsgenerator. Essentieel om onvoorspelbaar te zijn. (Afbeelding van Rob Seward onder Creative Commons [9])