Een theorie van granulariteit toegepast op telwoorden

Auteur: Viviënne Buisman (5637570)

Begeleider: dr. R.W.F. Nouwen
Tweede beoordelaar: dr. J. Korbmacher

Scriptie 7.5 ECTS, Bachelor Kunstmatige Intelligentie, UU

28 juni 2019

Samenvatting

In natuurlijke taal is vaak sprake van vaagheid en imprecisie. Imprecisie is niet per se een negatief fenomeen. Het verschilt per context welk niveau van precisie, oftewel granulariteit, we relevant vinden. Mensen kunnen over het algemeen goed inschatten welk niveau van granulariteit gewenst is en hoe ze uitingen van anderen moeten interpreteren. Voor computers is dit niet zo simpel. Als we echter een intelligentie machine willen hebben, is het essentieel dat deze een theorie van granulariteit in zich heeft. De vraag is hoe we granulariteit goed kunnen modelleren. Hobbs (1985) heeft een theorie van granulariteit voorgesteld die met name gaat over kennisrepresentatie van de wereld. In dit onderzoek is gekeken of we deze theorie ook kunnen gebruiken voor een meer talig fenomeen: telwoorden. Hiervoor is gekeken naar de semantiek en granulariteit van telwoorden aan de hand van bestaande literatuur en vervolgens is er geprobeerd een koppeling te maken met Hobbs' theorie. Met wat aanpassingen bleek Hobbs' theorie inderdaad toepasbaar te zijn.

Trefwoorden: granulariteit, imprecisie, vaagheid, simplificatie, kunstmatige intelligentie, schaalgranulariteit, telwoorden, rondheid, approximators.

Inhoudsopgave

1 Inleiding

In natuurlijke taal is vaak sprake van vaagheid en imprecisie. Er wordt namelijk vaak vrij losjes gesproken. Niemand kijkt raar op als Piet om 17:10 uur uit zijn werk komt, terwijl hij had gezegd dat hij om vijf uur klaar zou zijn met werken. Tevens, als een verkeersbord aangeeft dat een dorp 20 km verderop ligt, dan is dat vaak niet precies 20000,00 meter, maar kan dat ook een paar honderd meter meer of minder zijn. We zijn op dat moment simpelweg niet geïnteresseerd in die precisie, omdat we in het verkeer grofweg willen weten hoe ver iets is. We zouden alleen maar heel vermoeid raken als we van alles de precieze details te horen zouden krijgen.
Imprecisie is dus niet per se een negatief fenomeen. Het verschilt per context welk niveau van detail we relevant vinden. Zo is in het verkeer kort en bondige informatie gewenst. Als we namelijk heel veel details zouden zien, zouden we de relevante informatie er niet uit kunnen halen. Dit zou dan veel gevaarlijke situaties kunnen opleveren. Bij een sprintwedstrijd daarentegen is gedetailleerde informatie juist wel relevant, want dan kunnen milliseconden bepalend zijn voor wie er sneller is. We kijken dus telkens op een bepaalde, zogenoemde, *granulariteit* naar dingen in de wereld. Met granulariteit bedoelen we dus het niveau van detail. Hoe fijnkorreliger de granulariteit, hoe preciezer het detailniveau.

Mensen kunnen doorgaans makkelijk inschatten welk niveau detail op welk moment gewenst is en ze weten vaak goed hoe ze uitingen van anderen moeten interpreteren. Voor computers is dit nog een ander verhaal. Als je een computer zou vragen wanneer je voor het laatst je ketel hebt laten checken, zou hij kunnen antwoorden met "402 dagen geleden" (of zelfs nog gedetailleerder). We zouden op zo'n moment eigenlijk liever iets willen horen als "Meer dan een jaar geleden" of beter "Te lang geleden". Tevens andersom, als een mens zou aangeven dat de ketel een jaar geleden is gecontroleerd, dan zou een computer dat interpreteren als 365 dagen geleden, terwijl dit waarschijnlijk niet echt het precieze aantal dagen is. We willen dat een computer dingen iets ruimer kan interpreteren, dus bijvoorbeeld in dit geval als tussen de 320 en 410 dagen. Het zou dus handig zijn als een computer een idee heeft van granulariteit in taal. De vraag is hoe we zulke verschillende granulariteitsniveaus kunnen modelleren.

In het gebied van kunstmatige intelligentie is er een soortgelijk probleem voor granulariteit te vinden, alleen dan over kennisrepresentatie van de wereld. Als we een intelligente machine willen hebben, is het volgens Hobbs, die een paper over granulariteit heeft geschreven (1985), essentieel dat deze machine een theorie van granulariteit in zich heeft. Volgens Hobbs is het vermogen om de wereld op verschillende granulariteiten te conceptualiseren en hiertussen te schakelen namelijk fundamenteel voor de intelligentie van de mens. Men kan hierdoor de wereld om zich heen eenvoudig in kaart brengen. Zonder de mogelijkheid om bijvoorbeeld naar een grovere granulariteit te vereenvoudigen, zou kennisrepresentatie inefficiënt zijn. Het zou dan namelijk alle, ook irrelevante, details registeren. Hobbs heeft een kader gegeven voor zo'n theorie van granulariteit. Zijn theorie gaat over het construeren van eenvoudige theorieën uit complexere theorieeï, dus over het versimpelen van theorieën die *agents* over hun omgeving hebben.

Hobbs is in zijn paper niet nader ingegaan op granulariteit van taal. In die tak van granulariteit ben ik juist geïnteresseerd. Ik ben nieuwsgierig naar hoe we ervoor kunnen zorgen dat een machine de complexiteit van menselijke taaluitingen begrijpt en taal op de juiste manier kan interpreteren in verschillende contexten. Hoe we er dus voor kunnen zorgen dat een machine weet wanneer uitingen precies moet worden geïnterpreteerd en wanneer juist niet. Dus dat een machine niet zou vastlopen als je zou zeggen "Geef de 100 mensen op het feest een glas champagne", terwijl er maar 95 mensen zijn. We willen dan dat de computer begrijpt dat '100' met een marge bedoeld zou kunnen zijn en dat zijn taak is volbracht wanneer hij alle 95 mensen een glas champagne heeft gegeven.

Ik wil in dit onderzoek kijken of ik Hobbs' theorie van granulariteit toepasbaar kan maken voor talige fenomenen die met granulariteit te maken hebben. Ik hoop dat ik hierdoor meer kan leren over hoe granulariteit in elkaar steekt. In deze studie kan ik helaas niet het gehele assortiment aan talige fenomenen analyseren, dus ik ga me in dit onderzoek toespitsen op telwoorden. Ik zal proberen een antwoord te geven op de volgende onderzoeksvraag: "(Hoe) kunnen we Hobbs' theorie van granulariteit toepassen op telwoorden?"

Om deze vraag te beantwoorden ga ik eerst de voor dit onderzoek relevante stukken uit Hobbs' theorie bespreken, dan ga ik het hebben over de semantiek en granulariteit van telwoorden en ten slotte probeer ik Hobbs' theorie te koppelen aan de semantiek van telwoorden.

2 Hobbs' theorie van versimpeling

Om te kunnen kijken of we Hobbs' theorie van granulariteit kunnen toepassen op taal, is het noodzakelijk om te kijken wat zijn theorie precies inhoudt. Ik zal in dit hoofdstuk de voor dit onderzoek relevante delen van zijn theorie bespreken.

Zoals in de inleiding is besproken, kun je granulariteit vinden in taal, maar ook bijvoorbeeld in het opslaan en representeren van kennis over een wereld. Wanneer je naar de wereld kijkt, kun je dit in verschillende granulariteitsgroottes doen en je beschouwt meestal alleen datgene waar je op dat moment geïnteresseerd in bent. Wanneer je op straat loopt, kun je de hele straat in het oog nemen of focussen op een rij auto's die er staat. Je kan ook nog dieper inzoomen, bijvoorbeeld op een deuk op een van de auto's. Objecten en gebeurtenissen kunnen dus in verschillende niveaus van detail worden bekeken en beschreven.

Laten we kijken naar hoe Hobbs' theorie in elkaar steekt aan de hand van een voorbeeld. Stel je voor dat er een hele berg vol verschillende snoepjes voor je ligt. Deze snoepjes vallen allemaal onder de klasse 'snoepjes' en niet onder een klasse als 'koekjes'. Dit klinkt heel voor de hand liggend, maar als je dieper gaat inzoomen, zie je dat categoriseren steeds complexer wordt. Dan zie je namelijk dat je de snoepjes ook kan onderverdelen in vorm of kleur. Je kan dan groepen maken van rode snoepjes, blauwe snoepjes etcetera. Nog verder inzoomend, kom je veel verschillende tinten rood en blauw tegen. Je kan de wereld dus zo complex of simpel maken als je zelf wilt.

Als je iets wilt zeggen over deze wereld, moet je dus weten op welke granulariteitsgrootte je dat wil doen. Het kan echter moeilijk zijn om over een hele complexe wereld, zoals die met alle verschillende tinten snoepjes, te praten. Dan is het handig om die grote, complexe wereld iets te versimpelen, zonder er al te ver van af te stappen. Het zou al makkelijker zijn om dingen over deze wereld te kunnen zeggen als we de snoepjes categoriseren in 'rode snoepjes', 'roze snoepjes', etc. Om dit te doen moet je wel kijken welk snoepje in welke categorie, oftewel klasse, kan zitten en welke je dus kan beschouwen als niet te onderscheiden. Met niet te onderscheiden bedoel ik hier dat je kan zeggen dat bepaalde tinten kleuren eigenlijk zo dicht bij elkaar zitten, dat het kleine verschil dat er is, er niet toe doet. Of iets er niet toe doet, is wel relatief aan wat je op dat moment relevant vindt. We zijn op dit moment niet geïnteresseerd in kleine details, dus we willen verscheidene tinten rood als soortgelijken zien; die willen we allemaal zien als gewoon 'rood'. Hobbs heeft een theorie van granulariteit voorgesteld waarmee je dit zou kunnen doen. Hij heeft een theorie bedacht om uit complexere theorieën, eenvoudige theorieën te construeren. Ik ga nu op een formeler niveau laten zien hoe zijn theorie in elkaar steekt.

Een beschrijving van een wereld is wat Hobbs een theorie noemt. Stel je hebt als wereld een bibliotheek vol boeken. Die boeken, zeg boek 1, 2, 3, etc., zijn onze entiteiten en die zijn ons domein van interpretatie, $\mathbf{S}_0$. De boeken hebben allemaal een specifieke kleur in een kleurcode (#ff1713 en #ff1714 zijn bijvoorbeeld twee tinten rood), die kleuren zitten in de verzameling predicaten $\mathbf{P}_0$. $\mathbf{P}_0$ heeft dus alle kleurcodes in zich, maar ook predicaten als '*rood*' en '*blauw*'. Deze wereld met $\mathbf{S}_0$ en $\mathbf{P}_0$ noemen we $\mathbf{T}_0$. In $\mathbf{T}_0$ zit ook nog een verzameling proposities die geldt voor deze wereld. Dus bijvoorbeeld *rood*(1), *blauw*(2) en *#ff1714*(1) zijn waar, en alle andere (niet-genoemde) proposities zijn onwaar. Dus dan is het waar dat boek 1 rood is, maar niet waar dat boek 1 blauw is.

Nu zou het fijn zijn als we ook alleen naar een grovere, meer bruikbare, indeling van kleur in deze wereld zouden kunnen kijken. Dus dat we boeken categoriseren op de grove kleurklassen van onder andere 'rood' en 'blauw', in plaats van dat elk boek los gecategoriseerd staat op zijn kleurcode. Dan hebben we een indeling in circa tien klassen in plaats van honderden, waardoor we een meer overzichtelijke wereld hebben met simpelweg 'oranje boeken', 'rode boeken', etcetera. We willen dus van de complexere wereld $\mathbf{T}_0$, naar de eenvoudigere, grofkorrelige wereld $\mathbf{T}_1$. In $\mathbf{P}_0$ hadden we kleurcodes als predicaten, maar ook grovere kleuren zoals '*rood*' en '*blauw*'. Predicaten zoals deze laatstgenoemden zijn voor ons nu alleen relevant. Deze vallen dus in de verzameling $\mathbf{R}$, de verzameling van relevante predicaten.

Om tot de grove kleurindeling, die van de relevante predicaten, te komen, moeten we kleine kleurverschillen tussen boeken in $\mathbf{T}_0$ 'onzichtbaar maken' aangezien die niet relevant zijn voor $\mathbf{T}_1$. Er moet nu worden gekeken welke entiteiten niet te onderscheiden zijn qua kleur. Dit drukken we uit met de ononderscheidbaarheidsrelatie '$\sim$':

$(\forall x,y) x \sim \mathrm{y} \equiv (\forall p \in \mathbf{R})(p(x) \equiv p(y))$

Twee entiteiten x en y zijn niet te onderscheiden als voor alle predicaten die relevant zijn, oftewel voor de predicaten die in $\mathbf{R}$ zitten, geldt dat p waar is voor x dan en slechts dan p waar is voor y. Aan de hand van de relevante predicaten zien we dus welke entiteiten niet te onderscheiden zijn. Uit deze ononderscheidbaarheidrelaties worden equivalentieklassen ingedeeld.

κ is de toewijzing die elk element van $\mathbf{S}_0$ in zijn equivalentieklasse in $\mathbf{S}_1$ opneemt. $\mathbf{S}_1$ is dus de reeks equivalentieklassen van $\mathbf{S}_0$ ten opzichte van de ononderscheidbaarheidsrelatie $\sim$. De definitie van κ is als volgt:

$(\forall x,y) x \sim \mathrm{y} \equiv \kappa(x) = \kappa(y)$

Twee entiteiten zitten in dezelfde equivalentieklasse als ze niet te onderscheiden zijn. $\mathbf{P}_1$ is de reeks predicaten $\{\kappa(p) | \mathrm{p} \in \mathbf{R}\}$.

Laten we teruggaan naar ons voorbeeld en kijken hoe dit er dan uit komt te zien. Stel we kijken naar boek 1 en 2. In $\mathbf{T}_0$ gelden dat de proposities *#ff1713*(1),

#ff1714(2), *rood*(1) en *rood*(2) waar zijn. In onze versimpelde wereld $\mathbf{T}_1$ is hiervan alleen *'rood'* nog een relevant predicaat; kleurcodes zijn niet meer relevant. We krijgen dan met de bovenstaande definitie van $\mathbf{P}_1$: $\mathbf{P}_1 = \{\kappa(rood)\}$.
Als we nu willen kijken of boek 1 en 2 niet te onderscheiden zijn, kijken we naar de definitie van de ononderscheidbaarheidsrelatie. We moeten kijken of voor alle relevante predicaten hetzelfde geldt voor 1 als voor 2 en omgekeerd. Alleen '*rood*' is nog een relevant predicaat en daarvoor geldt *rood*(1) en *rood*(2), oftewel: $(\forall p \in \mathbf{R})(p(1) \equiv p(2))$. Boek 1 en 2 zijn dus niet te onderscheiden en ze komen dan ook in dezelfde equivalentieklasse in $\mathbf{S}_1$. Dus $\mathbf{S}_1 = \{[1,2]\}$.
Als we nu in $\mathbf{T}_1$ zijn, kunnen we niet zomaar checken of *rood*(1) geldt, aangezien deze vorm bij $\mathbf{T}_0$ hoort. In $\mathbf{T}_1$ kennen we alleen equivalentieklassen, en predicaten als κ(p). In $\mathbf{T}_1$ geldt dat $\kappa(p)(\kappa(x))$ waar is dan en slechts dan als $p(x)$ waar is. Dus $\kappa(rood)(\kappa(1))$ is waar dan en slechts dan als *rood*(1) waar is en we weten van $\mathbf{T}_0$ dat dat zo is.

Laten we nu nog even kijken naar het grotere plaatje. Dus stel we hebben $\mathbf{S}_0 = \{1,2,3,4\}$ en $\mathbf{P}_0 = \{$*#ff1713*, *#ff1714*, *#e5be01*, *#ffa500*, *rood*, *geel*, *oranje*$\}$ en er geldt dat *#ff1713*(1), *#ff1714*(2), *rood*(1), *rood*(2), *#e5be01*(3), *#ffa50- 0*(4), *geel*(3) en *oranje*(4) waar zijn (alle andere proposities in $\mathbf{T}_0$ zijn onwaar) en $\mathbf{R} = \{$*rood*, *geel*, *oranje*$\}$. Dan krijgen we dus $\mathbf{P}_1 = \{\kappa(rood), \kappa(geel), \kappa(oranje)\}$ en $\mathbf{S}_1 = \{[1,2], [3], [4]\}$. Entiteiten 3 en 4 hebben namelijk niet dezelfde relevante predicaten en zijn dus wel van elkaar te onderscheiden. Ze komen dus niet samen in een equivalentieklasse terecht. Entiteit 1 en 2 zitten wel in dezelfde equivalentieklasse omdat we net al hadden aangetoond ze niet te onderscheiden zijn. We hebben nu dus alle entiteiten in grovere klassen ingedeeld en onze versimpelde wereld $\mathbf{T}_1$ bereikt.

We hebben nu dus door middel van Hobbs' theorie uit een complexere wereld, een simpelere wereld geconstrueerd. Later in dit onderzoek wordt er gekeken of we dit kunnen toepassen op het concept van telwoorden. We gaan hiervoor eerst kijken naar wat telwoorden precies zijn en hoe de granulariteit en semantiek van telwoorden eruitziet.

3 Telwoorden

3.1 Wat zijn telwoorden?

Telwoorden zijn woorden die het aantal of (rang)nummer van iets aangeven. Woorden die een aantal of nummer weergeven, zijn hoofdtelwoorden en woorden die een rangvolgorde (eerste, tweede, etc.) weergeven, zijn rangtelwoorden. Dan kunnen deze telwoorden ook nog bepaald of onbepaald zijn. In dit onderzoek wordt er alleen gekeken naar bepaalde hoofdtelwoorden; woorden die specifieke hoeveelheden aangeven. Ik zal 'bepaalde hoofdtelwoorden' echter voor het gemak gewoon aanduiden als 'telwoorden' in dit onderzoek. Tevens heb ik het hier expliciet over 'specifieke' hoeveelheden omdat een specifiek aantal wegens imprecisie in de werkelijkheid niet altijd exact dat aantal is, ook al wordt een specifieke hoeveelheid gegeven.
Er wordt dus gekeken naar zinnen als: "Jan kocht 25 bonbons" of "De brug is 20 meter lang". Zinnen als "Er waren enkele bonbons verdwenen" worden buiten beschouwing gelaten.

Daarnaast zullen ook *approximators* kort aan bod komen. Dit zijn uitdrukkingen die beweringen meer of minder nauwkeurig maken (Sauerland & Stateva, 2007), zoals 'exact' en 'grofweg'. Dit zijn dan zinnen als "Er zaten ongeveer 50 mensen op de boot".

3.2 Imprecisie en granulariteit van telwoorden

Beschouw de volgende twee zinnen:

1. *Er zijn 200 T-shirts verkocht.*
2. *Er zijn 194 T-shirts verkocht.*

De tweede zin lijkt preciezer te zijn dan de eerste zin. In de literatuur wordt er een koppeling gemaakt tussen precisie en granulariteit. In de taalkundige literatuur wordt granulariteit vaak gezien als de dichtheid van punten op een meetschaal. (1) kan betrekking hebben op een schaal met veelvouden van 50 (..., 150, 200, 250, ...), of bijvoorbeeld van 25 of 10 (..., 190, 200, 210, ...), terwijl (2) een fijnere schaal heeft met intervallen van 1. Wat zijn dan de voor- en nadelen van het gebruik van (1) in plaats van (2) en andersom? Het gebruik van een fijnere granulariteitschaal zoals bij (2) levert meer precisie op. Dit is bijvoorbeeld belangrijk voor administratie van een kledingwinkel. Echter, precisie is niet altijd gewenst. Wanneer er bijvoorbeeld een grote verkoop is en een medewerker wil weten of er al meer shirts uit het magazijn moeten worden gehaald, dan wil hij alleen grofweg weten hoeveel shirts er al verkocht zijn. In deze context is (1) dan een meer gewenst niveau van granulariteit.

Hoe grover de granulariteitschaal, hoe meer imprecisie. Dit is zoals we net al zagen niet per se negatief. Vaak zien we dat getallen op een grovere schaal ronder zijn. Ronde getallen komen ook vaak frequenter voor dan niet-ronde getallen (Pollman & Jansen, 2001). Een verklaring hiervoor kan onder andere zijn dat in natuurlijke taal in het algemeen vaak geldt dat korte en simpele expressies worden geprefereerd. Doorgaans zijn aantallen die verwijzen naar de machten van tien, of veelvouden daarvan (zoals 30, 400 of 7000), korter dan aantallen die verwijzen naar aangrenzende natuurlijke cijfers (Krifka, 2002). Ronde getallen krijgen dus vaak de voorkeur, ook al kunnen deze imprecisie bevorderen. In veel situaties is beknoptheid namelijk een waardevoller doel om op te richten dan precisie, aldus Krifka. Zo wil je bijvoorbeeld dat informatie in het verkeer niet gedetailleerd is, maar beknopt, zodat er geen ongelukken ontstaan.

Welke getallen rond zijn is echter niet eenduidig. Pollman en Jansen (2001) vonden dat in culturen met een decimaal getalsysteem de mate van ronding nauw verbonden is met zogenoemde 10-heid, 2-heid, 5-heid en $2\frac{1}{2}$-heid, en met verdubbeling en halvering. Dit is bijvoorbeeld te zien bij jubilea (5, 10, 25, 50 etc) en geldstukken (1 euro, 2, 5, 10, 20 etc.). Kijk je echter naar eierdoosjes, dan zijn 6 en 12 ronde getallen. Bij graden en klokken zijn dat 15, 30, 45, 60, 90, etc. Welke getallen rond zijn, is dus afhankelijk van de context. Het is als het ware afhankelijk van hoe je meetlint eruitziet - op wat voor granulariteitschaal je je bevindt. Een eierdoos geeft een heel andere schaal dan je hand.

Precisie, granulariteit en rondheid zijn dus veelal met elkaar verbonden. Hoe grover de granulariteit, hoe minder precisie en hoe ronder de getallen. Andersom ook: Hoe meer precisie, hoe fijner de granulariteit en hoe minder rond de getallen.

3.3 Semantiek van telwoorden

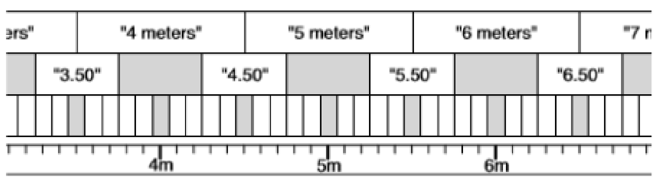
Ik wil nu laten zien hoe ik de semantiek van zinnen met telwoorden ga representeren. De compositionele semantiek doet er voor dit stuk niet toe, dus ik veronderstel dat ik voor zinnen aan een simpele betekenisrepresentatie kan komen die getalsinformatie uitdrukt. Voor meer over informatie over compositionaliteit kun je *Semantics for Counting & Measuring* (Rothstein, 2017) raadplegen.

De zinnen die ik beschouw zijn opgebouwd uit een telwoord (en eventueel een *approximator*) en hetgeen waar het telwoord over gaat. Ik ga het telwoord aanduiden als 'Getalspredicaat' en waar het telwoord over gaat, hetgeen waaraan je 'Getalspredicaat' kan relateren, met 'Aantal'. Dus 'Getalspredicaat' drukt de cardinaliteit uit van de waarde die 'Aantal' uitdrukt. Dit geef ik als volgt weer: Getalspredicaat(Aantal). Dus stel je hebt de zin "Er staan 50 schapen in de wei", dan wordt het: 50(|schapen in de wei|). Met de sluistekens om de zin geef ik aan dat het gaat om de waarde van de zin. Als we een zin met een *approximator* hebben, zoals "Ingrid kocht ongeveer 20 rozen", dan krijgen we Ongeveer(20(|rozen die Ingrid kocht|)), oftewel Approximator(Getalspredicaat(Aantal)).

Aantal is dus de daadwerkelijke waarde van wat er wordt uitgedrukt. Dus stel dat er maar 48 schapen in de wei staan terwijl er wordt gezegd dat er 50 staan, ziet dit er als volgt uit: 50(48). We zullen straks in hoofdstuk vier meer hierover zien.

3.4 Schaalgranulariteit

In de literatuur probeert men ook imprecisie te modelleren. Zo hebben Krifka (2007) en Solt (2014) beiden een alternatieve analyse van imprecisie voorgesteld op basis van schaalgranulariteit. Getallen kunnen in plaats van één punt een interval aanduiden (Krifka 2007). Zo kan worden gezegd dat een walvis 15 meter lang is, terwijl de daadwerkelijke lengte misschien ergens in het interval van 14.5 en 15.5 meter ligt. Hierbij is de schaal verdeeld in segmenten van 1 meter, zoals te zien is in de rij van '5 meters' in figuur 3.1 (Sauerland & Stateva, 2007, p. 232)



Figuur 3.1: Opstelling van lengtetermen met schaalintervallen. *Herdrukt van "Scalar vs. epistemic vagueness: evidence from approximators", door Sauerland, U., & Stateva, P., 2007, Semantics and Linguistic Theory, Vol. 17, pp. 232*

Sauerland en Stateva (2007), die de theorie van Krifka verder uitwerken, veronderstellen dat een granulariteitsfunctie elk punt toewijst aan een interval dat dit punt bevat. In de volgende voorbeelden zie je hoe zij bijvoorbeeld op deze manier '5 meter' interpreteren.

(a) gran_{fijn}(5m) = [4.95m, ..., 5.00m, ..., 5.05m]
(b) gran_{mid}(5m) = [4.75m, ..., ..., 5.00m, ..., ..., 5.25m]
(c) gran_{grof}(5m) = [4.50m, ..., ..., ..., 5.00, ..., ..., ..., 5.50m]

Bij de fijne granulariteitschaal gebruiken ze intervallen van 0.05 m, bij gemiddeld 0.25 m, en bij de grove schaal 0.50 m. Wanneer voor een punt meerdere granulariteiten worden overwogen, wordt deze toegewezen aan de meest grove granulariteit mogelijk, zodanig dat de expressie de kortste expressie mogelijk is. Dit ziet er dan als volgt uit:

(d) $[[5 \text{ meter}]]^{gran} = \text{gran}_{grof}(5\text{m}) = [4.50\text{m}, 5.50\text{m}]$
(e) $[[4 \text{ meter } 50]]^{gran} = \text{gran}_{mid}(4.5\text{m}) = [4.25\text{m}, 4.75\text{m}]$
(f) $[[4 \text{ meter } 90]]^{gran} = \text{gran}_{fijn}(4.9\text{m}) = [4.85\text{m}, 4.95\text{m}]$

Dan hebben we ook nog de *approximators*. Bij nauwkeurige uitdrukkingen als ‘exact’ wordt de granulariteitsparameter ingesteld op zo fijn mogelijk, terwijl minder nauwkeurige uitdrukkingen als ‘ongeveer’ het grofste niveau krijgt toegewezen door Sauerland en Stateva. Zo kunnen bijvoorbeeld uitdrukkingen waarvan we net zeiden dat ze een grovere interpretatie zouden krijgen, een fijnere granulariteit toegewezen krijgen. Stel, we hebben ‘exact 5 meter’. Zonder de approximator krijgt het een grovere granulariteit toegewezen zoals we zien bij (d), maar de granulariteit wordt nu in het nieuwe geval bepaalt door de approximator. ‘Exact’ zorgt er nu dus voor dat 5 meter een fijne interpretatie krijgt, namelijk: $[[\text{exact } 5 \text{ meter}]]^{gran} = \text{gran}_{fijn}(5\text{m}) = [4.95\text{m}, 5.05\text{m}]$.

We hebben nu gezien hoe Krifka (2007) en Sauerland & Stateva (2007) imprecisie en granulariteit van telwoorden modelleren. Ik wil nu echter kijken of ik Hobbs’ theorie van granulariteit kan gebruiken. Ik zal dan ook gebruik maken van intervallen zoals hier is gedaan, maar ik zal de granulariteit directer proberen in te pluggen. In plaats van “$[[50 \text{ meter}]]^{gran} = \text{gran}_{grof}(50\text{m}) = [40\text{m}, 60\text{m}]$”, zeg ik namelijk simpelweg 50_{10}. Ik zal in het volgende hoofdstuk meer uitleggen over deze notatie.

4 Telwoorden en Hobbs' versimpeling

4.1 De versimpeling

We gaan nu kijken of we Hobbs' theorie kunnen toepassen op telwoorden en of we ongeacht de uitkomst meer inzicht kunnen krijgen in granulariteit. Laten we weer het voorbeeld van "Er staan 50 schapen in de wei" gebruiken en kijken hoe het telwoord in deze zin zou worden gerepresenteerd volgens Hobbs' theorie van granulariteit.

In de complexe wereld, $\mathbf{T}_0$, zijn alle granulariteiten voor deze zogenoemde '50' nog mogelijk. Dat wil zeggen dat iemand het daadwerkelijk over 50 schapen kan hebben, maar dat het er ook 48 of zelfs 40 kunnen zijn. De '50' in ons voorbeeld kan in verschillende niveaus van granulariteit worden geïnterpreteerd. Over welk niveau we praten, geven we weer door middel van een lage markering, dus 50_5 betekent de 50 van een schaalgrootte met intervallen van 5, oftewel [48 - 52]. 55_5 betekent dan dat het aantal in interval [53 - 57] valt en 50_1 betekent de 50 van de schaal met intervallen van 1 schaap. We hebben vergeleken Krifka (2007) en Sauerland & Stateva (2007) de granulariteit nu dus directer ingeplugd.
De 50_1, 50_5, 55_5, etc. worden onze predicaten voor $\mathbf{T}_0$.

Zoals we al bij de semantiek in paragraaf 3.3 zagen, kunnen we telwoorden weergeven met Getalspredicaat(Aantal). De predicaten in $\mathbf{T}_0$ worden de Getalspredicaten, en de entiteiten $\mathbf{S}_0$ worden de Aantallen (dus wat het daadwerkelijke aantal is). Om Aantallen makkelijker te onderscheiden van Getalspredicaten, zal ik ze in de tekst onderscheiden door Aantallen dik te drukken en Getalspredicaten niet. We krijgen dan het volgende voor $\mathbf{T}_0$: $49_1(\mathbf{49})$, $50_1(\mathbf{50})$, $51_1(\mathbf{51})$, $50_5(\mathbf{49})$, $50_5(\mathbf{50})$, $50_5(\mathbf{51})$, ..., $50_{25}(\mathbf{61})$, etc. Onze verzameling predicaten is dus als volgt: $\mathbf{P}_0 = \{..., 49_1, 50_1, 51_1, ..., 50_5, ..., 50_{10}, ..., 50_{25}, ...\}$ en onze entiteiten, dus $\mathbf{S}_0, = \{..., \mathbf{49}, \mathbf{50}, \mathbf{51}, \ldots\}$

Als we dan $\mathbf{T}_0$ willen versimpelen naar een meer grofkorrelige wereld $\mathbf{T}_1$, dan willen we de fijnste schaalgranulariteit, in dit geval van interval 1, zogezegd onzichtbaar maken. We willen dat **49**, **50**, **51**, etc. niet van elkaar te onderscheiden zijn, omdat predicaten als 49_1, 50_1, etc. irrelevant geworden zijn. Dit doen we door middel van de ononderscheidbaarheidsrelatie $\sim$, zoals eerder gedefinieerd in hoofdstuk twee. Laten we hiervoor eerst kijken naar welke predicaten nu in **R** zitten, de predicaten die relevant voor ons zijn. Alleen de predicaten vanaf de granulariteitschaal van vijftallen zijn nog relevant. Ik neem nu alleen even de vijftallen als relevante predicaten (ik zal later laten zien waarom), dus onze **R** is als volgt: $\mathbf{R} = \{...,45_5, 50_5, 55_5, ..\}$. Twee entiteiten zijn ononderscheidbaar als ze voor deze relevante predicaten niet van elkaar verschillen. Stel we willen nu kijken of **49** en **51** niet te onderscheiden zijn, dan moet voor alle predicaten die in $\mathbf{R}$ zitten, gelden dat p waar is voor **49** dan en slechts dan als p waar is voor **51**. Predicaat 49_1 en 51_1, de enige predicaten waarin **49** en **51** verschilden, zitten niet in $\mathbf{R}$ en doen er dus niet toe. Wel gelden $50_5(\mathbf{49})$ en $50_5(\mathbf{51})$ nog. **49** en **51** vallen allebei in hetzelfde interval, dat van 50_5, en verschillen dus niet. Dus voor alle relevante predicaten geldt dezelfde waarheidswaarde voor **49** en **51**, dus $\mathbf{49} \sim \mathbf{51}$.

Entiteiten die ononderscheidbaar zijn krijgen in $\mathbf{T}_1$ dezelfde equivalentieklasse in $\mathbf{S}_1$ toegewezen. Zoals ik bij het stuk van Hobbs noemde, is κ de toewijzing die elk element van $\mathbf{S}_0$ in zijn equivalentieklasse in $\mathbf{S}_1$ opneemt en $\mathbf{P}_1$ is de reeks predicaten $\{\kappa(p)|\mathrm{p} \in \mathbf{R}\}$. Dus $\mathbf{P}_1 = \{..., \kappa(45_5), \kappa(50_5), \kappa(55_5), ...\}$ en $\mathbf{S}_1 = \{..., [\mathbf{43}, \mathbf{44}, \mathbf{45}, \mathbf{46}, \mathbf{47}], [\mathbf{48}, \mathbf{49}, \mathbf{50}, \mathbf{51}, \mathbf{52}], [\mathbf{53}, \mathbf{54}, \mathbf{55}, \mathbf{56}, \mathbf{57}], ...\}$.
Stel, we willen weten of $\kappa(50_5)(\kappa(\mathbf{49}))$ waar is, dan moeten we kijken of $50_5(\mathbf{49})$ waar is. $\kappa(50_5)(\kappa(\mathbf{49})) = \kappa(50_5)([\mathbf{48}, \mathbf{49}, \mathbf{50}, \mathbf{51}, \mathbf{52}])$. En we zien dat **49** inderdaad in een equivalentieklasse onder 50_5 valt, dus $\kappa(50_5)(\kappa(\mathbf{49}))$ is waar.

$\mathbf{S}_1$ bestaat nu dus uit equivalentieklassen van vijftallen. Dus we hebben nu de entiteiten in grovere klassen ingedeeld en onze versimpelde wereld $\mathbf{T}_1$ bereikt. Maar ik heb nu alleen vijftallen als relevante predicaten genomen, terwijl alle relevante predicaten niet enkel vijftallen of enkel tientallen zouden moeten zijn nu, maar gewoonweg alle mogelijke 'n-tallen' die groter zijn dan de intervallen van 1. Er is echter een reden waarom ik dat niet heb gedaan en dat zal ik nu in de volgende paragraaf laten zien.

4.2 Problemen oorspronkelijke toepassing Hobbs' theorie

Stel, we hebben $\mathbf{T}_0$: $\mathbf{P}_0 = \{..., 49_1, 50_1, 51_1,..., 50_5, ..., 50_{10}, ..., 50_{25} ,...\}$ en $\mathbf{S}_0 = \{..., \mathbf{49}, \mathbf{50}, \mathbf{51}, \dots\}$. We willen deze $\mathbf{T}_0$ weer gaan versimpelen en de fijnste schaalgranulariteit onzichtbaar maken. Predicaten met interval 1 zijn niet meer relevant. Ik had net alleen de vijftallen genomen als relevante predicaten voor $\mathbf{T}_1$, maar ik zal nu laten zien wat er gebeurt als de andere grovere predicaten ook nog relevant zijn, dus als $\mathbf{R} = \{..., 45_5, 50_5, 55_5, ..., 50_{10}, ..., 50_{25}, ...\}$.

Met deze $\mathbf{R}$ worden onze predicaten in $\mathbf{T}_1$: $\mathbf{P}_1 = \{..., \kappa(45_5), \kappa(50_5), \kappa(55_5), \kappa(50_{10}), \kappa(50_{25}), ...\}$. We moeten nu kijken wat onze $\mathbf{S}_1$ met equivalentieklassen wordt. Wat we intuïtief willen dat er nu zou gebeuren, is dat wat geldt voor de vijftallen, ook nog klopt als o.a. de tientallen bij de relevante predicaten zitten. We zagen net bijvoorbeeld dat als we alleen vijftallen hebben als relevante predicaten, dat voor 45_5 **43**, **44**, **45**, **46** en **47** dan niet te onderscheiden zijn en in dezelfde equivalentieklasse komen. Intuïtief zouden deze elementen ook nog ononderscheidbaar moeten zijn als we een stap grover gaan naar de tientallen. Echter, als we gaan kijken welke entiteiten nog te onderscheiden zijn als we ook tientallen als relevante predicaten hebben, dan zien we dat dit niet gebeurt.

Stel dat we namelijk willen checken of **44** en **45** niet te onderscheiden zijn, dan moet weer voor alle predicaten die in $\mathbf{R}$ zitten, gelden dat p waar is voor **44** dan en slechts dan als p waar is voor **45**. In figuur 4.1 zien we een illustratie van een aantal predicaten zodat we wat beter kunnen visualiseren waar elk predicaat naar verwijst. Laten we kijken welke predicaten waar zijn voor **44** en **45**. We hebben o.a. $45_5(\mathbf{44})$ en $45_5(\mathbf{45})$. Het predicaat 45_5 is waar voor zowel **44** als **45** dus ze verschillen niet. Maar dan geldt ook nog dat $40_{10}(\mathbf{44})$, maar dit predicaat is niet waar voor **45**. En er geldt dat $50_{10}(\mathbf{45})$, maar niet $50_{10}(\mathbf{44})$. Niet voor alle relevante predicaten geldt voor **44** en **45** hetzelfde, dus $\mathbf{44} \not\sim \mathbf{45}$. **44** en **45** komen dus niet in dezelfde equivalentieklasse terecht. We zouden nu voor $\mathbf{S}_1$ wel o.a. [**43**, **44**] en [**45**, **46**, **47**] als equivalentieklassen krijgen, maar dit is niet het intuïtief correcte resultaat wat we wilden bereiken. We willen namelijk dat **43**, **44**, **45**, **46** en **47** allemaal in dezelfde equivalentieklasse terecht komen. Dus we zien nu dat als 50_5 relevant is, 50_{10} (etcetera) niet ook relevant zijn.

$\vdots$

$45_5 =$ [**43**, **44**, **45**, **46**, **47**]

$50_5 =$ [**48**, **49**, **50**, **51**, **52**]

$55_5 =$ [**53**, **54**, **55**, **56**, **57**]

$40_{10} =$ [**35**, **36**, **37**, **38**, **39**. **40**, **41**, **42**, **43**, **44**].

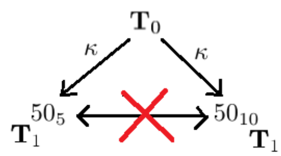
$50_{10} =$ [**45**, **46**, **47**, **48**, **49**, **50**, **51**, **52**, **53**, **54**]

$\vdots$

Figuur 4.1: Illustratie van intervallen van predicaten

4.3 De oplossing

Het doel van Hobbs' theorie is om vanuit een complexe wereld te versimpelen naar een eenvoudigere, grofkorreligere wereld. Je wilt hiervoor dus de fijnste details in de complexe wereld onzichtbaar maken. Ik heb geprobeerd om dit te doen, maar we zagen net dat het problemen oplevert als we enkel dit doen. De versimpelde wereld zoals deze intuïtief zou moeten zijn, is er namelijk niet meer als er meerdere granulariteitsniveaus naast elkaar bestaan in één wereld. Daarom heb ik in paragraaf 4.1 toen we naar $\mathbf{T}_1$ gingen versimpelen niet alleen het fijnste detailniveau eruit gehaald, maar ook de andere granulariteitniveaus, op één na. Op die manier is $\mathbf{T}_1$ wel correct. We kunnen nu dan alleen niet vanuit deze $\mathbf{T}_1$ verder versimpelen, terwijl we wel nog de mogelijkheid willen hebben om naar een nog grovere granulariteit te simplificeren. We kunnen dit oplossen door meerdere versimpelde werelden $\mathbf{T}_1$ te maken, met telkens een ander granulariteitsniveau. Dus in plaats van te versimpelen van $\mathbf{T}_0$, naar $\mathbf{T}_1$, naar $\mathbf{T}_2$, etc., willen vanuit $\mathbf{T}_0$ naar meerdere $\mathbf{T}_1$'s versimpelen. We willen vanuit $\mathbf{T}_0$ verschillende toewijzingen κ hebben naar verschillende $\mathbf{T}_1$. Dus aparte toewijzingen voor vijftallen, aparte toewijzingen voor tientallen etc., waarbij andere n-tallen niet gelijktijdig in dezelfde werelden zijn. Zie figuur 4.2 voor een kleine illustratie hiervan.



Figuur 4.2: Illustratie van toewijzingen naar verschillende $\mathbf{T}_1$

4.4 Interpretatie

We hebben nu dus gezien hoe we werelden met telwoorden kunnen versimpelen door wat aanpassingen te doen. We weten nu alleen nog niet precies hoe we een zin als "Er staan 50 schapen in de wei", moeten interpreteren als we die tegenkomen. We weten namelijk niet welke predicaten nu echt relevant zijn in de werkelijkheid. Daar gaan we nu ook geen specifiek antwoord op vinden in dit onderzoek. Ik wil nu wel globaal laten zien hoe je zinnen kunt interpreteren.

Je kan in het algemeen vaak kijken naar de context van een zin om het granulariteitsniveau te bepalen. Wanneer je in een precieze context zit (dus bijvoorbeeld als je voor een belastingformulier moet opgeven hoeveel schapen je precies hebt), wil je de fijnst mogelijke granulariteit kiezen. Dit doen we in $\mathbf{T}_0$ door het predicaat te kiezen met de kleinste lage markering, dus in dit geval dan 50_1. In een grovere context willen we een iets grotere predicaat kiezen, zoals 50_5. Dus in een precieze context kunnen we $[[\text{Anna kreeg 50 rozen}]]^{\mathbf{T}_0}$ interpreteren als bijvoorbeeld: $[[50_1(\mathbf{50})]]^{\mathbf{T}_0}$ en in een grovere context als bijvoorbeeld $[[50_5(\mathbf{50})]]^{\mathbf{T}_0}$. Let wel dat ik nu **50** heb gekozen als Aantal, maar dit kan ook bijvoorbeeld **48** zijn. Het daadwerkelijke aantal is nu namelijk nog niet bekend. Om alsnog wel de methode te kunnen laten zien, zal ik voor het gemak nu telkens **50** gebruiken als Aantal.

Stel dat we echter een zin met een *approximator* hebben zoals "Anna kreeg ongeveer 50 rozen", dan willen we dat die *approximator* de richting van de interpretatie aangeeft. Het gebruik van een *approximator* is dan ook een manier om een gegeven context te veranderen. Ook al bevind je je in een precieze context, een *approximator* zoals 'ongeveer' kan alsnog een grovere granulariteit geven. Een *approximator* als 'exact' zou dan juist een fijnere granulariteit geven.

Dus stel, we hebben "Anna kreeg ongeveer 50 rozen" in een precieze context, dus $[[\text{Anna kreeg ongeveer 50 rozen}]]^{\mathbf{T}_0} = [[\text{Ongeveer}(50_1(\mathbf{50}))]]^{\mathbf{T}_0}$.
Voordat ik de toepassing van Hobbs' met telwoorden had gemaakt en alleen bekend was met hoe Hobbs' theorie op zichzelf werkte, had ik verwacht dat we als volgt zouden simplificeren: van $\mathbf{T}_0$ naar $\mathbf{T}_1$, naar $\mathbf{T}_2$, etc. De interpretatie zou er dan als volgt hebben uitgezien: $[[\text{Ongeveer}(50_1(\mathbf{50}))]]^{\mathbf{T}_0} = [[50_{10}(\mathbf{50})]]^{\mathbf{T}_2}$ of $[[\text{Ongeveer}(50_1(\mathbf{50}))]]^{\mathbf{T}_0} = [[50_{25}(\mathbf{50})]]^{\mathbf{T}_3}$. Dus als ik in een precieze context "Anna kreeg ongeveer 50 rozen" zou zeggen in $\mathbf{T}_0$, dan zou dit in bijvoorbeeld $\mathbf{T}_2$ of $\mathbf{T}_3$ kunnen worden geïnterpreteerd. Dit zou in het eerste geval betekeken dat het daadwerkelijke aantal rozen voor 'ongeveer 50' dan in het interval [45 - 54] zit en in het tweede geval in [38 - 62].

We hebben net echter gezien hoe we Hobbs' theorie kunnen toepassen op telwoorden en daarbij gaan we niet meer van $\mathbf{T}_0$ naar $\mathbf{T}_1$, naar $\mathbf{T}_2$, maar van $\mathbf{T}_0$ naar meerdere $\mathbf{T}_1$'s. De bovenstaande interpretatie is dus niet meer in overeenstemming met onze toepassing van Hobbs' theorie. Hoe kunnen we de ons voorbeeld dan wel interpreteren? We willen namelijk nog steeds tussen werelden kunnen schakelen van granulariteit.

We hadden: $[[\text{Anna kreeg ongeveer 50 rozen}]]^{\mathbf{T}_0} = [[(\text{Ongeveer}(50_1)(\mathbf{50}))]]^{\mathbf{T}_0}$.
Ik wil nu echter laten zien hoe we in het algemeen, dus vanuit willekeurige waardes voor de predicaten, tot een interpretatie kunnen komen. Dus:
$[[\text{Anna kreeg ongeveer 50 rozen}]]^{\mathbf{T}_0} = [[\text{Ongeveer}(50_i(\mathbf{50}))]]^{\mathbf{T}_0}$.
In welke granulariteit we ook zitten, als we 'ongeveer' willen toepassen en dus naar grovere granulariteiten willen gaan, dan willen we dat we naar een interpretatie kunnen gaan die in ieder geval grover is dan de i. We willen dan dat $[[\text{Ongeveer}(50_i(\mathbf{50}))]]^{\mathbf{T}_0} = [[50_{i+x}(\mathbf{50})]]^{\mathbf{T}_1}$, waarbij $x > 0$.

Zoals ik eerder vermeldde, weten we dus niet welke predicaten nu daadwerkelijk relevant zijn, maar ik heb nu wel geprobeerd een beeld te schetsen van hoe je zinnen met telwoorden zou kunnen interpreteren.

In dit hoofdstuk hebben we dus gekeken naar hoe Hobbs' theorie van versimpeling werkt bij telwoorden. In het laatste hoofdstuk, de conclusie en discussie, zal ik reflecteren wat ik heb gedaan.

5 Conclusie en discussie

In dit onderzoek is gekeken of een algemene theorie van granulariteit, die van Hobbs (1985), ook toepasbaar is voor een talig fenomeen als telwoorden. Hiervoor heb ik in de inleiding de volgende onderzoeksvraag geformuleerd: "(Hoe) kunnen we Hobbs' theorie van granulariteit toepassen op telwoorden?". Om deze vraag te beantwoorden heb ik eerst gekeken naar hoe Hobbs' theorie in elkaar steekt en daar de relevante delen uit geselecteerd. Vervolgens heb ik gekeken naar de semantiek en granulariteit van telwoorden aan de hand van al bestaande literatuur. Daarna heb ik geprobeerd om zelf een andere theorie van granulariteit, die van Hobbs, toe te passen op telwoorden.

Voor de onderzoeksvraag was het eerst de vraag of het überhaupt mogelijk is om Hobbs' theorie van granulariteit toe te passen op telwoorden. Dit bleek inderdaad mogelijk te zijn, maar dan wel met een paar aanpassingen. Dat brengt ons op het 'hoe-gedeelte' van de onderzoeksvraag. Ik heb ontdekt dat je in $\mathbf{T}_1$, en verdere werelden, niet meerdere granulariteitsniveaus naast elkaar kan hebben, omdat ze elkaar dan in de weg zitten. Een oplossing die ik hiervoor heb gevonden, is om specifiek vanuit $\mathbf{T}_0$ te kijken naar welke granulariteitsgrootte je toe wilt versimpelen en de andere granulariteiten dan even buiten beschouwing laten in die wereld.

Ik wil graag nog vermelden dat er eigenlijk veel meer predicaten zijn dan degenen die ik heb gebruikt in mijn simplificatie. Je zou namelijk ook 48_3 en 48_7 kunnen hebben. Ik heb mijn $\mathbf{P}_0$ echter laten bepalen door rondheid. We zagen al in paragraaf 3.2 dat rondheid afhankelijk is van de context. In mijn context had ik te maken met intervallen, en schaalintervallen gaan meestal in intervallen van 1, 5, 10, etc. zoals ook te zien is in figuur 3.1 (Sauerland & Stateva, 2007, p. 232). We tellen dan ook niet in zeventallen, maar wel vaak in vijftallen, tientallen etc. Dat is de reden dat ik ervoor gekozen heb om deze intervallen aan te houden voor mijn predicaten. In een context met tellen met eieren zou je echter misschien predicaten als 12_2 hebben. De predicaten die we kiezen zijn dus afhankelijk van rondheid en rondheid is afhankelijk van de context, van op wat voor granulariteitschaal je je bevindt. Hier zien we dus weer de relatie tussen granulariteit en rondheid terugkomen. Rondheid is ook de reden dat we meerdere predicaten met 50 hebben, want 50 komt op meerdere schaalgroottes voor (die van 1, 5, 10, 25, 50, etcetera).

In hoofdstuk drie zagen we dat in de literatuur ook is geprobeerd om imprecisie te modelleren voor telwoorden. Ik heb geprobeerd om dit op een efficiëntere manier te doen door middel van Hobbs' theorie. Ik heb ten eerste namelijk granulariteit directer in kunnen pluggen in uitdrukkingen. In plaats van een hele reeks te hebben zoals $[[50 \text{ meter}]]^{gran} = \text{gran}_{grof}(50\text{m}) = [40\text{m}, 60\text{m}]$, zeg ik namelijk simpelweg 50_{10}. Ten tweede is het toepassen van Hobbs' theorie mogelijk een algemene manier om granulariteit in taal te modelleren, niet alleen voor telwoorden. We zagen namelijk in de inleiding dat we graag granulariteit voor alle talige fenomenen willen kunnen modelleren en nu kunnen we dus mogelijk één theorie gebruiken voor alle fenomenen, in plaats van voor elk fenomeen een aparte theorie. Hier is nog wel vervolgonderzoek voor nodig; er moet worden gecheckt of Hobbs' theorie echt toepasbaar is voor andere talige fenomenen dan alleen telwoorden. Zo heeft Bosman (2019) bijvoorbeeld ook succesvol Hobbs' theorie van granulariteit toegepast op een talig fenomeen, namelijk op de locatieve indexical 'hier'. Wel heb ik in dit onderzoek enkel gekeken naar bepaalde hoofdtelwoorden en heb ik andere soorten telwoorden buiten beschouwing gelaten, maar deze andere telwoorden zouden ook nog kunnen worden bekeken in volgend onderzoek.

Verder zou er in vervolgonderzoek kunnen worden gekeken naar het achterhalen van welke predicaten nu daadwerkelijk relevant zijn. Ik heb in dit onderzoek namelijk telkens gewoon aannames gemaakt over wat ik relevant vind en dus als **R** neem. Echter is het in de werkelijkheid niet gemakkelijk te bepalen wat nu relevant is, het is iets wat we proberen in te schatten. Voor Hobbs' theorie echt geïmplementeerd kan worden, moet dus nog verder worden onderzocht hoe je de daadwerkelijk relevante predicaten kan achterhalen. Tevens zou het uiteindelijk mooi zijn om te kijken of er in een computer echt een theorie van granulariteit kan worden toegepast, bijvoorbeeld voor natuurlijke taalherkenning.

De achterliggende gedachte van dit onderzoek was dat als we een intelligente machine willen hebben, het essentieel is dat deze een theorie van granulariteit in zich heeft, aangezien dat een van de krachten is van de menselijke intelligentie. Hierbij hoort ook granulariteit van taal. In het vakgebied van de kunstmatige intelligentie is men volop bezig met het ontwikkelen van natuurlijke taalherkenning voor computers. Er is al veel bereikt, maar er is nog een hoop te doen voor het helemaal goed werkt - zoals het goed modelleren van granulariteit. In dit onderzoek heb ik geprobeerd granulariteit van telwoorden te modelleren. Dit is echter maar een heel klein stukje van de gehele wereld van granulariteit en dit bleek al niet heel eenvoudig te zijn om te modelleren. Ik denk echter dat het uiteindelijk zeker mogelijk kan zijn om in de kunstmatige intelligentie een goed model te hebben voor granulariteit. Er is gewoon nog een lange, maar wel zeer interessante weg te gaan.

6 Bibliografie

1. Bosman, L. (2019). Granulariteit in taal: Een bestaande granulariteitstheorie toegepast op de locatieve indexical 'hier'. Geraadpleegd in scriptiearchief Universiteit Utrecht, http://studenttheses.library.uu.nl

2. Hobbs, J.R. (1985). Granularity: Proceedings 9th International Joint Conference on Artificial Intelligence, Los Angeles, 432-435.

3. Jansen, C.J., & Pollmann, M.M. (2001). On round numbers: Pragmatic aspects of numerical expressions. Journal of Quantitative Linguistics, 8(3), 187-201.

4. Krifka, M. (2002). Be brief and be vague: and how bidirectional optimality theory allows for verbosity and precision. In D. Restle & D. Zaefferer (eds.), Sounds and Systems. Studies in Structure and Change: A Festschrift for Theo Vennemann, 439–458. Berlin: Mouton de Gruyter.

5. Krifka, M. (2007), Approximate Interpretation of Number Words: A Case of Strategic Communication, in Bouma, G., et al. (eds.), Cognitive Foundations of Interpretation, Proceedings of the KNAW colloquium, Amsterdam, pp. 111–126.

6. Rothstein, S. (2017). Semantics for Counting and Measuring (Key Topics in Semantics and Pragmatics). Cambridge: Cambridge University Press. doi:10.1017/9780511734830

7. Sauerland, U., & Stateva, P. (2007). Scalar vs. epistemic vagueness: Evidence from approximators. In Semantics and Linguistic Theory, Vol. 17, pp. 228-245.

8. Solt, S. (2014). An alternative theory of imprecision. In: Semantics and Linguistic Theory,vol. 24, 514–533