

Verwijderen van irrelevante kennis*

Jaap Kamps Bert Bredeweg Cis Schut

Sociaal Wetenschappelijke Informatica
Universiteit van Amsterdam
Roetersstraat 15 1018 WB Amsterdam
tel: 020-525 6788 fax: 020-525 6896
e-mail: {kamps,bert,schut}@swi.psy.uva.nl

Samenvatting

In dit artikel presenteren wij een techniek voor het automatisch opsporen en verwijderen van irrelevante kennis in kwalitatieve modellen. Het verwijderen van irrelevante kennis betekent dat een algemeen model wordt afgestemd op de specifieke situatie van het systeem (het gebruikte scenario). Gebruikmakend van een bibliotheek met algemene domeinkennis, een scenario en een kwalitatieve gedragsvoorspelling wordt vastgesteld welke kennis essentieel is en welke kennis niet. De techniek gebruikt feedback van een knowledge engineer over het al dan niet handhaven van niet-essentiële kennis. Hierna wordt irrelevante kennis verwijderd uit de domeinkennis, zodat er een efficiënter en duidelijker model ontstaat. De techniek is geïmplementeerd en is beschikbaar binnen de domein onafhankelijke, kwalitatieve redeneeromgeving GARP.

*Dit onderzoek werd mede ondersteund door de Stichting Informatica Onderzoek Nederland (SION) met een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) project nummer: 612-322-307. Een vorige versie van dit artikel is gepresenteerd op de *International Workshop on Qualitative Reasoning about Physical Systems 1993*.

1 Inleiding

Het konstrueren van modellen voor simulaties wordt over het algemeen gezien als het vereenvoudigen van een complex basismodel tot een model dat is afgestemd op de specifieke eigenschappen van het systeem [Ziegler, 1976]. Bij het konstrueren van een kwalitatief model dat het gedrag van een systeem voorspelt [Weld & de Kleer, 1990], omvat dit afstemmen zowel een modelleer als een simulatie activiteit die afwisselend uitgevoerd worden. Het modelleerproces begint met het verzamelen van de algemene domeinkennis die op het systeem van toepassing is. Deze domeinkennis kan worden gerepresenteerd in een bibliotheek met kanonieke model fragmenten [de Kleer & Brown, 1984, Forbus, 1984]. Het algemene model wordt daarna verfijnd door een aantal keer het model te simuleren en de model fragmenten aan te passen. De aanpassingen van het model worden gebaseerd op verschillen tussen het resultaat van de simulatie en het vereiste of verwachte resultaat. Ons onderzoek is erop gericht om het modelleerproces te ondersteunen met hulpmiddelen die fundamentele stappen van het verfijningsproces kunnen uitvoeren.

Een belangrijke reden waarom algemene model fragmenten niet tot een geschikte gedragsvoorspelling hoeven te leiden, is dat de model fragmenten meer domeinkennis beschrijven dan strikt noodzakelijk is voor de voorspelling van het gedrag van het systeem. In dit artikel presenteren wij een techniek die automatisch deze irrelevante kennis opspoort en verwijderd uit het kwalitatieve model. In het bijzonder formuleren wij principes die een eenvoudig model genereren door een algemeen model aan te passen. De techniek gebruikt hiervoor: (i) een beschrijving van de specifieke beginsituatie van het systeem (het scenario); (ii) een verzameling algemene model fragmenten die van toepassing zijn op het systeem en (iii) een kwalitatieve voorspelling van het gedrag van het systeem. Met deze gegevens is de techniek in staat om de overbodige, de onnodig gedetailleerde en de essentiële kennis in het model te identificeren. Aan de hand van deze typen kennis kunnen model vereenvoudigingen worden afgeleid, die worden uitgevoerd zolang het model blijft voldoen aan drie eisen: (i) de gewenste gedragsvoorspelling moet haalbaar zijn; (ii) moet natuurkundig korrekt zijn en (iii) moet cognitief inzichtelijk zijn.

In paragraaf 2 uitgelegd hoe een kwalitatief model wordt gekonstrueerd. In paragraaf 3 wordt de representatie achtergrond van ons onderzoek besproken. Paragraaf 4 bespreekt de kondities waaronder kennis als (ir)relevante kan worden beschouwd. Vervolgens wordt in paragraaf 5 uitgelegd hoe irrelevante kennis is op te sporen en in paragraaf 6 hoe deze is te verwijderen. Tenslotte worden in paragraaf 7 de resultaten van ons onderzoek samengevat en besproken.

2 Automatische model konstruktie

De techniek die in dit artikel wordt beschreven, maakt deel uit van een verzameling hulpmiddelen die worden ontwikkeld om een knowledge engineer te ondersteunen bij het konstrueren van kwalitatieve modellen. Het modelleerproces bestaat uit een aantal taken (zie figuur 1). Een belangrijke stap bij het konstrueren van modellen is de representatie van algemene domeinkennis in model fragmenten zoals beschrijvingen van vloeistoffen, vloeistof stromen, warmte stromen, etc. Deze model fragmenten kunnen worden bewaard in bibliotheken zodat deze herbruikbaar zijn voor toekomstige modellen.

Een tweede belangrijke stap is de konstruktie van een scenario dat belangrijke aspecten

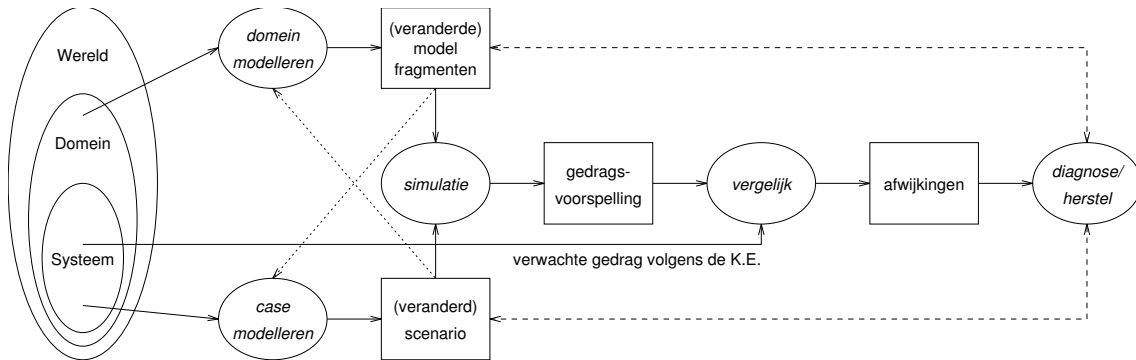


Figure 1: *Taken bij het konstrueren van kwalitatieve modellen.*

van het gemodelleerde systeem specificeert. Het resulterende scenario zal informatie over de fysieke structuur van het systeem bevatten, evenals beginwaarden van variabelen die van belang zijn voor de simulatie.

Het modelleren van de domeinkennis en het scenario gebeurt niet onafhankelijk. Het identificeren van bepaalde fysieke objecten kan de constructie vereisen van model fragmenten die hun gedrag beschrijven. En de aanwezigheid van een bepaald model fragment kan leiden tot de identifikatie van een bepaald object. Het is belangrijk dat de resultaten van de twee modelleersteps op elkaar afgestemd zijn. Het gedrag dat in de model fragmenten gemodelleerd is moet af te beelden zijn op de structuur die in het scenario gespecificeerd is.

Nadat een significante hoeveelheid kennis is gemodelleerd in de model fragmenten en het scenario, kan een kwalitatieve simulator gebruikt worden om het gedrag van het systeem te voorspellen. Vervolgens vergelijkt de knowledge engineer het voorspelde gedrag met het werkelijke gedrag van het systeem. Er zijn vaak verschillen tussen het voorspelde en het werkelijke gedrag, vooral in het begin van het modelleerproces. Daarom is er een diagnose/herstel stap nodig die (i) bepaalt welke fout het afwijkende gedrag veroorzaakt en (ii) de model fragmenten en/of het scenario aanpast zodat het afwijkende gedrag niet meer voor zal komen.

De taken worden binnen het modelleerproces meermalig uitgevoerd tot het voorspelde gedrag overeenkomt met het werkelijke gedrag. Verschillende aspecten van dit modelleerproces kunnen worden ondersteund. De techniek die gepresenteerd wordt in dit artikel ondersteunt de diagnose/herstel stap.

Het uiteindelijke doel is om het diagnose/herstel proces in zijn geheel te automatiseren. Onze strategie daarbij is om afzonderlijke technieken te ontwikkelen voor de verschillende typen fouten die in het model kunnen voorkomen zoals ambiguïteit of ongewenst detail. Bij een inkorrekte voorspelling kiest de knowledge engineer een techniek om de model fragmenten en het scenario te verbeteren. De knowledge engineer heeft de mogelijkheid om van de door een techniek voorgestelde aanpassingen af te wijken.

In [Bredeweg & Schut, 1993] is reeds een techniek beschreven voor het verwijderen van ongewenste ambiguïteit, veroorzaakt door missende kennis over globale gedrageigenschappen van een systeem. In dit artikel beschrijven wij een techniek die een model vereenvoudigt door niet-noodzakelijke kennis te verwijderen. De techniek veronderstelt dat de juiste verzameling toepasbare model fragmenten is geselecteerd en dat de simu-

latie de gewenste gedragstoestanden voorspelt. Alhoewel deze simulatie “korrekt” is, kan het model nog irrelevante kennis bevatten: parameters die nooit een toestandsovergang veroorzaken, onnodig detail in de fysieke structuur van het systeem, etc. De techniek identificeert en verwijdert deze irrelevante kennis in het kwalitatieve model.

3 Representatie omgeving

Deze paragraaf geeft een korte beschrijving van het gebruikte raamwerk voor kwalitatieve voorspelling van gedrag. Dit raamwerk is geïmplementeerd in de domein onafhankelijke, kwalitatieve redeneeromgeving GARP [Bredeweg, 1992]. Deze redeneeromgeving kan door knowledge engineers gebruikt worden om kwalitatieve modellen te ontwikkelen. Evenals de komponent- [de Kleer & Brown, 1984] en de procesgerichte [Forbus, 1984] aanpak, gebruikt GARP model fragmenten om het gedrag te bepalen van systemen in de wereld. Alle model fragmenten bestaan uit een verzameling kondities waaronder het model fragment van toepassing is en een verzameling gevolgen die van toepassing zijn als de kondities vervuld zijn. De kondities en gevolgen worden uitgedrukt in termen van model bouwstenen: *systeem elementen*: abstrakties van entiteiten in de wereld zoals vaten, vloeistoffen, componenten, etc.; *parameters*: variabelen die eigenschappen van de systeem elementen beschrijven zoals druk, temperatuur, voltage, etc.; *parameter waarden*: de waarden (puntwaarden, intervalwaarden en afgeleiden) van variabelen zoals $-$, 0 , $+$, etc.; *parameter relaties*: relaties (of constraints) tussen variabelen zoals ongelijkheden, evenredigheden, inducties etc.; *model fragmenten*: andere model fragmenten (statische, processen, etc.) zoals een warmte stroom proces als voorwaarde voor een kook proces, etc. Het gedrag van het systeem gedurende een bepaalde tijd wordt beschreven door de verzameling toepasbare model fragmenten. De begintoestand van het systeem wordt bepaald door het scenario. Het gedrag over verschillende tijdsperiodes wordt bepaald door de toepassing van overgangsregels. De overgangsregels zijn domein onafhankelijk gedefinieerd en bestaan uit meerdere soorten: er zijn regels die mogelijke veranderingen vaststellen (zoals de verandering van een parameter met negatieve afgeleide) en er zijn regels die de mogelijke veranderingen ordenen (zoals het gelijktijdig veranderen van evenredige parameters). Door het toepassen van de overgangsregels worden steeds de volgende toestanden afgeleid. De simulatie stopt op het moment dat er geen nieuwe toestanden meer kunnen worden afgeleid.

4 Irrelevante kennis

Het verwijderen van irrelevante kennis betekent dat een algemeen model wordt afgestemd op het specifieke scenario. Dit brengt een aantal voordelen met zich mee. Ten eerste wordt de efficiëntie van de gedragsvoorspelling verhoogd (irrelevante parameterwaarden hoeven niet meer te worden berekend). Ten tweede worden de relevante aspecten van het model en de voorspelling benadrukt, waardoor deze eenvoudiger te begrijpen zijn (irrelevante onderdelen hoeven niet langer geïnterpreteerd te worden). Ten derde is dit een stap in de richting van heuristische kennis zoals die door menselijke experts wordt gebruikt. Door het verwijderen van irrelevante kennis ontstaat zg. uitgekompileerde kennis, waarmee het probleem op de snelst mogelijke manier is op te lossen.

Het voorgaande kan geïllustreerd worden door het model van een balans met op beide armen een leeglopend vat (zie figuur 2). De vloeistof stroomt geleidelijk uit de vaten door

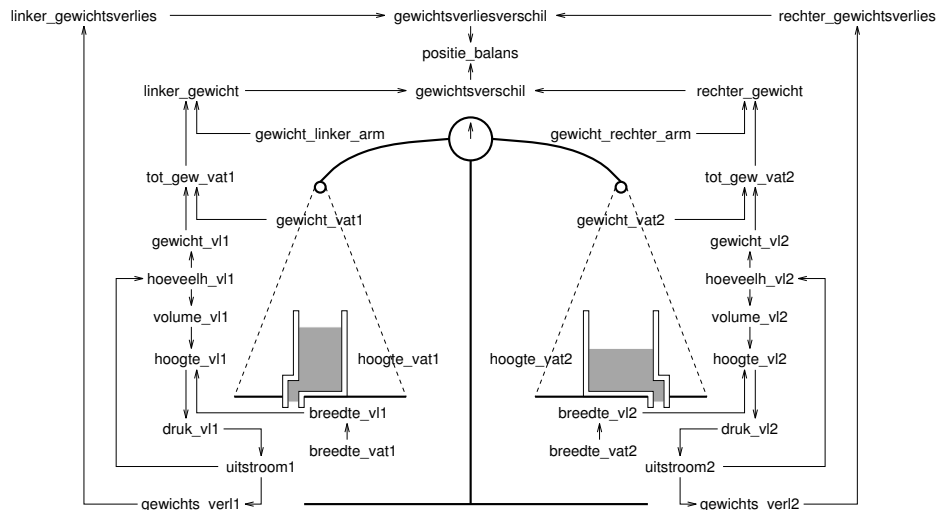


Figure 2: *Een algemeen model voor balans problemen.*

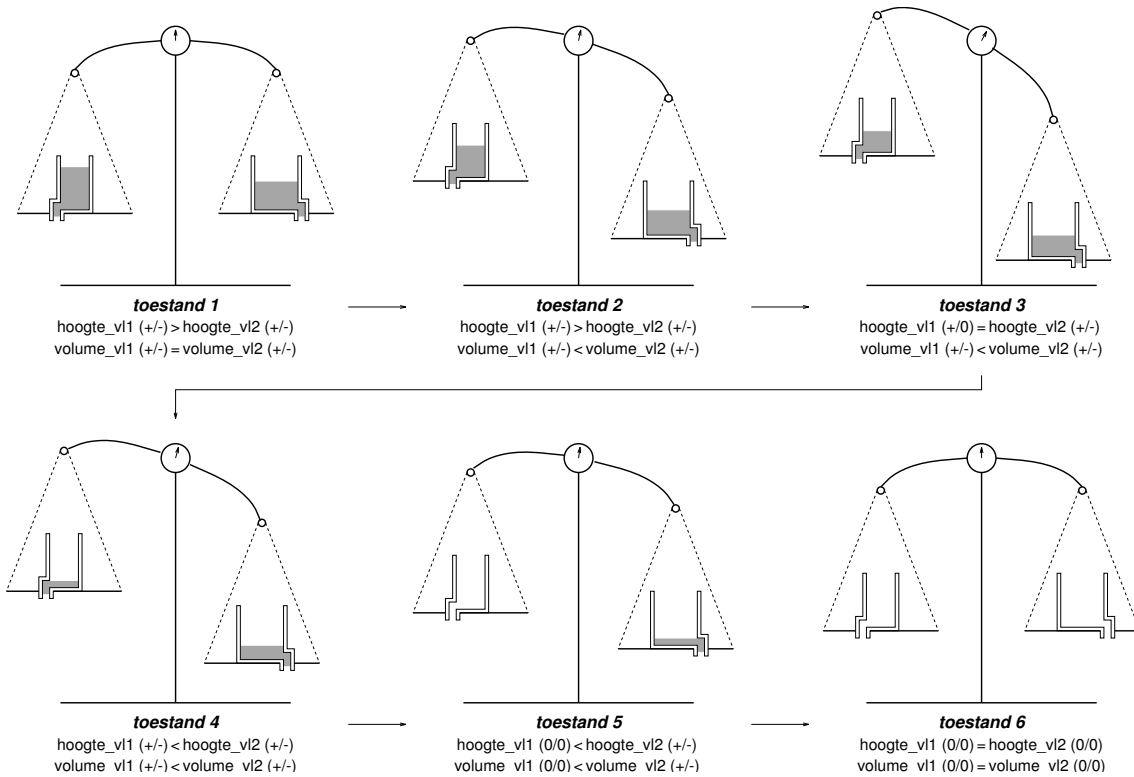


Figure 3: *De gedragsvoorspelling van het balans probleem.*

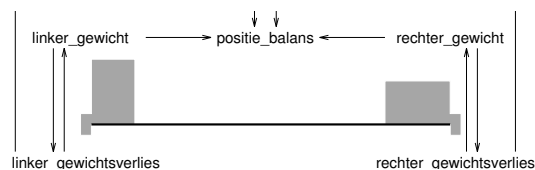


Figure 4: *Een eenvoudig model voor balans problemen.*

een opening in de bodem. Beide vaten hebben hetzelfde gewicht en bevatten (in de beginsituatie) dezelfde hoeveelheid vloeistof. De vorm van beide vaten is echter verschillend.

Het gedrag van dit systeem kan voorspeld worden door gebruik te maken van algemene model fragmenten zoals *vloeistof*, *vloeistof_in_vat*, *vloeistof_stroom* en *balans* (zoals o.a. gebruikt in [Forbus, 1984]). Dit leidt tot het model in figuur 2. Dit model is in staat om een korrekte gedragsvoorspelling te genereren, deze bestaat uit de zes gedragstoestanden in figuur 3. Het model bevat echter veel kennis die niet strikt noodzakelijk is om het juiste gedrag te voorspellen. Een voorspelling die uit dezelfde gedragstoestanden en toestandsovergangen bestaat, kan worden gegenereerd met meer eenvoudige model fragmenten die horen bij het model in figuur 4.

Ons doel is om, uitgaande van een algemeen model, automatisch een eenvoudig model te genereren. Hierbij moet een eenvoudig model voldoen aan de volgende eisen:

Kwalitatieve haalbaarheid Een eenvoudig model moet in staat zijn om het gewenste gedrag te voorspellen. Hiervoor is een minimale verzameling kennis nodig. De minimale verzameling kennis weerspiegelt de eisen van het kwalitatieve redeneermechanisme: de verzameling kennis moet het mechanisme in staat stellen alle gedrags-toestanden af te leiden.

Natuurkundige korrektheid Ten eerste moet de kennis die in het model gebruikt wordt een natuurkundig verantwoorde naam hebben (dus *hoeveelheid(vloeistof₁)* in plaats van $x(y_1)$). Ten tweede kan het gewenst zijn om kennis te behouden of toe te voegen omdat een model daar, vanuit natuurkundig oogpunt, “beter” door wordt. Een voorbeeld is de eis dat alle parameters van fundamentele natuurwetten (Newton, Boyle, etc.) aanwezig moeten zijn, ook indien deze niet nodig zijn voor de kwalitatieve haalbaarheid.

Kognitieve inzichtelijkheid Behalve de natuurkundige korrektheid zijn er eisen aan de “kognitieve inzichtelijkheid” van een model. Indien een model tot doel heeft om uitleg over of “begrip” van een bepaald aspect te geven, kan dit het behoud of de introductie van kennis vereisen om bepaalde aspecten van het model te benadrukken. Een voorbeeld is de aanwezigheid van een *hoogte*, *druk* en *uitstroom* parameter in het balans model om de uitleg te vereenvoudigen. Een ander voorbeeld is het opzettelijk gebruiken van een complexer model om de overeenkomsten met ingewikkelde problemen op hetzelfde domein te vergroten.

Kennis die niet bijdraagt aan de vervulling van deze drie eisen is irrelevant.

5 Opsporen van irrelevante kennis

De eenvoudige model fragmenten die door de techniek gegenereerd worden, zijn toepasbaar op (i) het originele scenario en (ii) variaties van dezelfde parameters die in het originele scenario gemanipuleerd werden. Bij een andere fysieke structuur of bij de manipulatie van andere parameters dan in het originele scenario kan een andere verzameling model fragmenten van toepassing zijn, waardoor andere kennis relevant of irrelevant wordt.

Het opsporen van irrelevante kennis gebeurt met behulp van de algemene model fragmenten, het scenario en een volledige gedragsvoorspelling. Figuur 5 geeft een overzicht van het verwijderen van irrelevante kennis. De eerste taak is het onderscheiden van drie typen

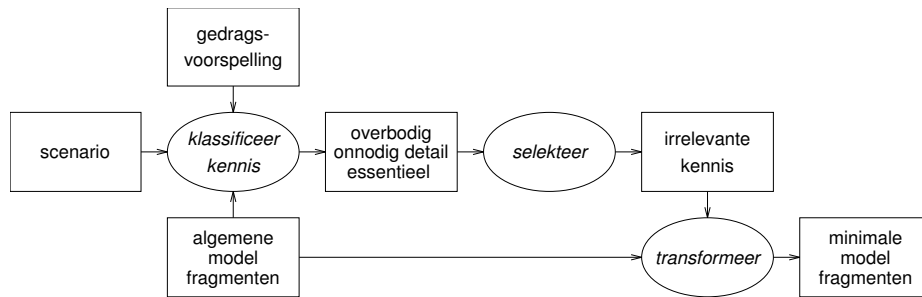


Figure 5: *Taken bij het verwijderen van irrelevante kennis.*

kennis: overbodige kennis, onnodig gedetailleerde kennis en essentiële kennis (beschreven in deze paragraaf). De andere taken uit figuur 5 bestaan uit het beslissen over de handhaving van niet-essentiële kennis en het uiteindelijke verwijderen van irrelevante kennis (beide beschreven in paragraaf 6).

Er zijn drie typen kennis te onderscheiden:

Overbodige kennis Kennis is overbodig indien het geen enkele bijdrage levert aan de voorspelling. Deze kennis kan worden verwijderd, mits de kennis niet van belang is voor de korrektheid of inzichtelijkheid van het model.

Onnodig gedetailleerde kennis Kennis is onnodig gedetailleerd indien meerdere model bouwstenen dezelfde bijdrage aan de voorspelling leveren. Onnodig gedetailleerde kennis kan worden samengevoegd met of vervangen door de kennis die op zichzelf deze bijdrage levert, mits de kennis niet van belang is voor de korrektheid of inzichtelijkheid van het model.

Essentiële kennis Kennis is essentieel indien het een unieke bijdrage levert aan de voorspelling. Zonder deze kennis kan het verwachte gedrag niet voorspeld worden. Deze kennis moet worden behouden omdat verwijdering niet in overeenstemming is met de kwalitatieve haalbaarheidseis.

In alle drie gevallen kan de kennis bestaan uit de model bouwstenen: parameters, waarden, relaties, systeem elementen en model fragmenten (zie paragraaf 3). Omdat irrelevantie zich niet-monotoon gedraagt [Subramanian & Genesereth, 1987], is de volgorde waarin deze bouwstenen worden opgespoord belangrijk. De parameters vormen de belangrijkste groep omdat deze de toestandsovergangen veroorzaken door veranderingen in hun waarde. De andere bouwstenen zijn afhankelijk van de parameters.

5.1 Parameters

5.1.1 Overbodige parameters

Parameters die niet van waarde veranderen, zoals parameters zonder relaties, zijn overbodig. Indien een model uit meerdere ongerelateerde delen bestaat, moeten de delen verwijderd worden of het model opgesplitst. Een voorbeeld van een overbodige parameter is de *hoogte* van de vaten in het balans model.

5.1.2 Onnodig gedetailleerde parameters

Onnodig gedetailleerde parameters zijn verzamelingen parameters die dezelfde bijdrage aan het model leveren. Parameters uit deze verzameling mogen worden verwijderd door ze te vervangen door andere parameters uit deze verzameling. Parameters uit deze verzameling mogen ook worden samengevoegd door een nieuwe parameter te creëren die meerdere “oude” parameters uit deze verzameling omvat. Niet alle parameters mogen worden verwijderd omdat dan de kwalitatieve gedragsvoorspelling niet meer haalbaar is. Parameters zijn onnodig gedetailleerd indien deze aan drie voorwaarden voldoen.

Korresponderende parameters Een voorwaarde voor onnodige gedetailleerde parameters is dat de parameters korresponderen. Twee parameters korresponderen indien ze dezelfde toestandsovergangen veroorzaken. Dit houdt in dat de parameters hetzelfde kwalitatieve gedrag vertonen: in elke toestand hebben beide parameters een korresponderende waarde en dezelfde afgeleide. Een voorbeeld van korresponderende parameters zijn de *hoogte*, *druk* en *uitstroom* van een vloeistof in het balans model. Met behulp van de gedragsvoorspelling kan worden vastgesteld dat deze parameters hetzelfde kwalitatieve gedrag vertonen: in de toestanden waar er nog vloeistof is hebben ze allemaal waarde + en afgeleide – en in de toestanden waar er geen vloeistof meer is waarde 0 en afgeleide 0.

Konflikterende relaties Een tweede voorwaarde is dat de parameters geen konfliktierende relaties hebben. In de beginsituatie van het balans model hebben zowel de *hoogte* als het *volume* van *vloeistof*₁ een korresponderende waarde evenals de beide parameters van *vloeistof*₂ (zie figuur 3). In de begintoestand is echter $volume_{vl1} = volume_{vl2}$ en $hoogte_{vl1} > hoogte_{vl2}$ door de verschillend gevormde vaten. Dit betekent dat de hoogte en volume parameters niet dezelfde kennis vertegenwoordigen en dus niet kunnen worden samengevoegd. Deze relaties kunnen voorkomen omdat waarden uit punten of uit intervallen kunnen bestaan. Indien twee parameters een korresponderende puntwaarde hebben, zijn ze ook gelijk. Indien twee parameters een korresponderende intervalwaarde hebben, zegt dat niets over de gelijkheid: de ene parameter kan immers een hogere *kwantitatieve* waarde hebben dan de andere parameter.

Veranderingen binnen objecten Een derde voorwaarde is dat de parameters tot hetzelfde systeem element behoren. In het balans model voldoen de *hoeveelheid* en het *volume* van de vloeistoffen aan deze voorwaarde en het *gewicht* van *vat*₁ en *vat*₂ voldoen niet aan deze voorwaarde¹ (zie figuur 6). Het komt echter vaak voor dat korresponderende

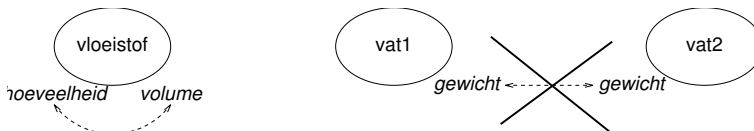


Figure 6: Alleen veranderingen indien de parameters tot hetzelfde systeem element behoren.

parameters niet tot hetzelfde systeem element behoren maar tot verschillende onderdelen van een samengesteld object. In het balans model geldt dit voor de *breedte* van het

¹Hetgeen in dit geval geen probleem is omdat deze parameters niet veranderen en dus overbodig zijn.

vat en de vloeistof (zie figuur 7). Deze parameters mogen niet worden samengevoegd of verwijderd, tenzij de delen van het samengestelde object ook worden samengevoegd. Indien de delen van het samengestelde object worden samengevoegd, ontstaat er een nieuw

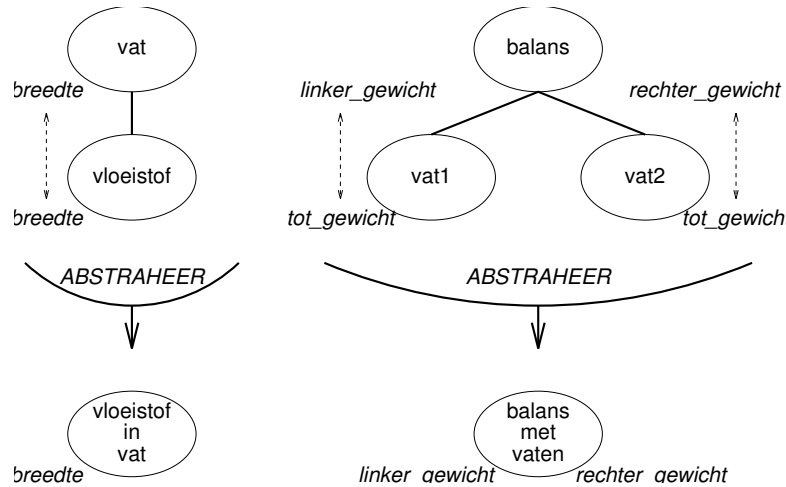


Figure 7: *Samenvoegen van parameters van verschillende systeem elementen.*

systeem element die de “oude” omvat. De parameters behoren nu wel tot hetzelfde systeem element en dus is samenvoeging of verwijdering toegestaan. Een ander voorbeeld is het *totaal_gewicht* van de vaten en het *linker_gewicht* en *rechter_gewicht* van de balans (zie ook figuur 7). Samenvattend: de vraag of parameters van verschillende systeem elementen in een samengesteld object onnodig gedetailleerd zijn, hangt samen met de vraag of hun systeem elementen onnodig gedetailleerd zijn. De reden voor deze voorzichtigheid is dat één van de delen in een of meerdere gedragstoestanden zelfstandig kan voorkomen. Een verandering zal ook dit zelfstandig voorkomend deel beïnvloeden, hetgeen betekent dat er een verandering over de grenzen van een object heen optreedt.

5.1.3 Essentiële parameters

Indien een toestandsovergang uitsluitend wordt veroorzaakt door de waardeverandering van één parameter of door de verandering van de gelijkheidsrelatie tussen twee parameters dan zijn die parameters essentieel. Alle parameters die niet overbodig of onnodig gedetailleerd zijn, worden voor de zekerheid als essentieel beschouwd. Bijvoorbeeld een parameter die in een aantal toestanden korrespondeert met een parameter en in de andere toestanden met een andere parameter. In het balans model is de *positie* parameter van de balans essentieel.

5.2 Overige bouwstenen

Waarden en relaties Niet toepasbare waarden en relaties zijn overbodig. Identieke waarden en relaties zijn onnodig gedetailleerd. Deze identieke voorkomens kunnen ontstaan indien parameters worden samengevoegd. Verder zijn waarden en relaties overbodig, onnodig gedetailleerd of essentieel indien ze horen bij een parameter die overbodig, onnodig gedetailleerd of essentieel is.

Systeem elementen Systeem elementen die geen parameters hebben en die niet voorkomen in de kondities van een model fragment zijn overbodig. Deze systeem elementen kunnen ontstaan doordat hun parameters irrelevant blijken te zijn. Zoals reeds eerder is opgemerkt, hangt de vraag of systeem elementen onnodig gedetailleerd zijn samen met de vraag of hun parameters onnodig gedetailleerd zijn. Systeem elementen die in alle toestanden samen voorkomen en die korresponderende parameters hebben, zijn onnodig gedetailleerd en kunnen worden samengevoegd. De overige systeem elementen zijn essentieel. In het balans model is het niet nodig om expliciet te redeneren over de *vaten* die de *vloeistof* bevatten (zie figuur 7). Er kan een nieuw systeem element *vloeistof_in_vat* worden gekreëerd die de “oude” systeem elementen omvat.

Model fragmenten Niet toepasbare model fragmenten zijn overbodig. Model fragmenten met een lege verzameling gevolgen (indien dit model fragment toepasbaar is, wordt er geen kennis toegevoegd) en die niet in de kondities van een ander model fragment voorkomen, zijn overbodig. Lege model fragmenten kunnen ontstaan doordat hun inhoud irrelevant blijkt te zijn. Model fragmenten die onnodig gedetailleerde systeem elementen beschrijven, zijn ook onnodig gedetailleerd. Onnodig gedetailleerde model fragmenten kunnen worden samengevoegd. De overige model fragmenten zijn essentieel.

6 Verwijderen van irrelevante kennis

Essentiële kennis mag niet worden verwijderd omdat dan de gedragsvoorspelling niet meer haalbaar is. Onnodig gedetailleerde kennis mag worden verwijderd of samengevoegd en overbodige kennis mag worden verwijderd, mits deze niet van belang zijn voor de natuurkundige korrektheid of de cognitieve inzichtelijkheid van het model.

De knowledge engineer zal moeten beslissen over de natuurkundige korrektheid en de cognitieve inzichtelijkheid van het model. Met andere woorden: mogelijke veranderingen (verwijderen van onnodig gedetailleerde en overbodige kennis) moeten eerst worden goedgekeurd door de knowledge engineer, alvorens deze worden uitgevoerd.

Er zijn een aantal heuristieken die knowledge engineers gebruiken om te beslissen over de handhaving van niet-essentiële kennis variërend van het verwijderen van alle niet-essentiële kennis tot meer conservatieve regimes. Een heuristiek is het altijd verwijderen van overbodige kennis. Deze kennis levert geen enkele bijdrage aan de gedragsvoorspelling en heeft geen relatie tot andere delen van het model. Het is een redelijke aanname dat deze kennis de natuurkundige korrektheid of de cognitieve inzichtelijkheid niet vergroot. Een andere heuristiek is het altijd behouden van kennis uit het scenario. Het scenario wordt gespecificeerd door de knowledge engineer en het is een redelijke aanname dat een knowledge engineer geen ongewenste model bouwstenen specificeert. Dit zou ook een reden kunnen zijn om binnen een verzameling onnodig gedetailleerde parameters andere parameters te verwijderen ten gunste van een parameter uit het scenario. Deze heuristieken kunnen ook worden geautomatiseerd. Het is dan mogelijk om nadat de voorkeur van een knowledge engineer vastgesteld is, deze automatisch uit te voeren.

Nadat de irrelevante kennis is vastgesteld wordt deze verwijderd uit de domein kennis (model fragmenten) en het scenario. Voor details over de techniek wordt verwezen naar [Kamps, 1993]. Deze techniek is niet specifiek voor het verwijderen van irrelevante kennis maar toepasbaar voor model veranderingen in het algemeen.

7 Resultaten en discussie

Het in dit artikel beschreven onderzoek heeft geleid tot een domein onafhankelijke methode voor het verwijderen van irrelevante kennis. In de methode worden drie typen kennis en drie eisen gerelateerd (zie figuur 8). Ten eerste onderscheiden we essentiële

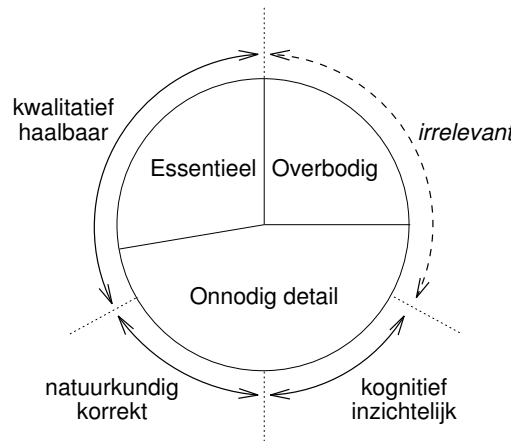


Figure 8: *Typen kennis.*

kennis, deze kennis levert een unieke bijdrage is aan de voorspelling: zonder deze kennis is de gewenste kwalitatieve gedragsvoorspelling niet haalbaar. Ten tweede onderscheiden we onnodig gedetailleerde kennis, dit is een verzameling bouwstenen die op zichzelf dezelfde bijdrage aan de voorspelling leveren: een van deze bouwstenen is voldoende voor de haalbaarheid van de kwalitatieve gedragsvoorspelling. Andere bouwstenen uit deze verzameling kunnen de natuurkundige korrektheid of de cognitieve inzichtelijkheid vergroten. De rest van deze verzameling is irrelevant. Ten derde onderscheiden we overbodige kennis, deze kennis levert geen bijdrage aan de voorspelling. Deze kennis is irrelevant, omdat het een redelijke aanname is dat deze kennis de natuurkundige korrektheid of de cognitieve inzichtelijkheid van het model niet vergroot. De methode is geïmplementeerd in SWI-Prolog [Wielemaker, 1992] en is te gebruiken binnen de domein onafhankelijke, kwalitatieve redeneeromgeving GARP.

In [Subramanian & Genesereth, 1987] wordt een formele beschouwing gegeven van irrelevantie. Zij behandelen onder andere Quines principe [Quine, 1963]: “indien objecten in een bepaalde context niet van elkaar te onderscheiden zijn, dan moeten deze objecten in die context als identiek worden behandeld”. Dit komt overeen met onze notie van onnodig gedetailleerde kennis. Alhoewel Subramanian en Genesereth in dit artikel een basis geven voor het redeneren over irrelevantie, stellen zij ook vast dat het opsporen van irrelevante kennis in specifieke domeinen een open probleem blijft. In dit artikel hebben wij vastgesteld onder welke voorwaarden model bouwstenen irrelevant zijn voor het kwalitatief voorspellen van het gedrag.

Ook anderen hebben onderzoek verricht naar het redeneren over (ir)relevantie bij het konstrueren van kwalitatieve modellen. Andere benaderingen zoals [Nayak et al., 1992, Falkenhainer & Forbus, 1991] gebruiken relevantie redeneren [Levy et al., 1992] om vast te stellen welke model fragmenten van belang zijn voor het gebruikte scenario. In vergelijking met deze benaderingen gaan wij een stap verder in het vaststellen van het geschikte model:

wij richten ons niet alleen op de (ir)relevantie van hele model fragmenten, maar ook op de (ir)relevantie van de bouwstenen in de model fragmenten.

We hebben aangetoond dat er binnen model fragmenten nog irrelevante kennis kan voorkomen en we hebben beschreven hoe deze opgespoord en verwijderd kan worden. Door deze benadering is het niet nodig om grote bibliotheken aan te leggen met model fragmenten die slechts minimaal van elkaar verschillen. Onze benadering levert hierdoor een belangrijke bijdrage aan het oplossen van het probleem om geschikte kwalitatieve modellen te konstrueren.

References

- [Bredeweg, 1992] Bredeweg, B. (1992). *Expertise in Qualitative Prediction of Behaviour*. PhD thesis, University of Amsterdam.
- [Bredeweg & Schut, 1993] Bredeweg, B. & Schut, C. (1993). Reducing ambiguity by learning assembly behaviour. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence '93*, pages 980–985. Chambéry, France, August 1993.
- [de Kleer & Brown, 1984] de Kleer, J. & Brown, J. S. (1984). A qualitative physics based on confluences. *Artificial Intelligence*, 24:7–83.
- [Falkenhainer & Forbus, 1991] Falkenhainer, B. & Forbus, K. D. (1991). Compositional modeling: finding the right model for the job. *Artificial Intelligence*, 51:95–143.
- [Forbus, 1984] Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, 24:85–168.
- [Kamps, 1993] Kamps, J. (1993). Alleviation of irrelevant knowledge: Support knowledge for constructing qualitative models. Master's thesis, SWI UvA / VU Amsterdam.
- [Levy et al., 1992] Levy, A. Y., Iwasaki, Y., & Motoda, H. (1992). Relevance reasoning to guide compositional modeling. In *Proceedings of the 6th International Workshop on Qualitative Physics*, pages 7–21, Intelligent Automation Laboratory, Heriot-Watt University, Edinburgh, Scotland.
- [Nayak et al., 1992] Nayak, P. P., Joskowicz, L., & Addanki, S. (1992). Automated model selection using context-dependent behaviors. In *Proceedings of the tenth National Conference on Artificial Intelligence*, pages 710–716. Morgan Kaufmann Publishers Los Altos CA.
- [Quine, 1963] Quine, W. V. O. (1963). *From a logical point of view*. Harper Torchbooks Inc. NY, second revised edition.
- [Schut & Bredeweg, 1993] Schut, C. & Bredeweg, B. (1993). Automatic enhancement of model parsimony. In *Proceedings of the 7th International Workshop on Qualitative Reasoning about Physical Systems*, pages 194–203. Orcas Island, Washington, U.S.A., May 1993.
- [Subramanian & Genesereth, 1987] Subramanian, D. & Genesereth, M. R. (1987). The relevance of irrelevance. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 416–422.
- [Weld & de Kleer, 1990] Weld, D. S. & de Kleer, J., editors (1990). *Readings in Qualitative Reasoning about Physical Systems*. Morgan-Kaufman CA.
- [Wielemaker, 1992] Wielemaker, J. (1992). SWI-prolog 1.6 reference manual. Technical report, Dept. of Social Science Informatics, University of Amsterdam.
- [Ziegler, 1976] Ziegler, B. P. (1976). *Theory of modelling and simulation*. Wiley-Interscience, New York, London.