

Empirische toetsing van modellerstrategieën



Auteur
Studentnummer
Versie
Afstudeernummer
Begeleider
Referent

Joris Reijbroek
0421804
1.0
36 IK
Dr. S.J.B.A Hoppenbrouwers
Prof. Dr. H.A. Proper



1 Samenvatting

Het onderzoek dat is beschreven in deze scriptie, probeert enige modelleerstijlen aan het licht te brengen, zoals deze in de praktijk voorkomen. Daarnaast is deze empirisch gebaseerde informatie gerelateerd aan theorie zoals deze op dit moment binnen de onderzoeksgroep IRIS wordt ontwikkeld.

Om de benodigde informatie te verzamelen, hebben een negental deelnemers elk twee cases gemodelleerd met behulp van ORM. De deelnemers bestonden uit twee groepen: ervaren en onervaren modelleurs. De deelnemers moesten de cases op twee verschillende manieren oplossen. Tijdens de eerste case werden de deelnemers geheel vrij gelaten in het modelleren, vervolgens werd een stappenplan overhandigd waarmee de tweede case gemodelleerd moest worden. Door de cases op twee verschillende manieren op te lossen, kon het effect van sturing door middel van een stappenplan worden onderzocht.

Het modelleerproces tijdens deze experimenten is op verschillende manier vastgelegd zodat deze achteraf kon worden geanalyseerd. Tijdens deze analyse kwamen een aantal veel voorkomende fenomenen naar voren die in deze scriptie worden behandeld als modelleerstijlen.

In de literatuur zijn weinig vergelijkbare onderzoeken te vinden. Daarom is het belangrijk om een onderzoek als dit goed te evalueren en de methode duidelijk te beschrijven. Eén van de subvragen luidt daarom: Hoe kan een theorie die betrekking heeft op modelleerstijlen empirisch getest worden? De beantwoording van deze vraag gaat niet in op het onderzoek, maar gaat *over* het onderzoek zelf. Ook is het onderzoek erg exploratief (verkenkend) opgezet.

2 Voorwoord

Deze scriptie vormt het sluitstuk van mijn afstudeeronderzoek aan de Radboud Universiteit Nijmegen. Gedurende de afgelopen zes maanden heb ik een onderzoek uitgevoerd naar modelleerprocessen en hoe deze zich in de praktijk manifesteren. Met dit onderzoek sluit ik tevens mijn masterstudie informatiekunde binnen het NIII (Nijmeegs Instituut voor Informatica en Informatiekunde) af.

Een praktisch onderzoek als dit is niet mogelijk zonder mensen daadwerkelijk te laten modelleren. De deelnemers aan mijn onderzoek zou ik daarom bij deze hartelijk willen bedanken voor hun inzet en tijd. In alfabetische volgorde waren dit: Patrick van Bommel, Pieter Buitenhuis, Ferry Bruijsten, Maikel Couwenberg, Stijn Hoppenbrouwers (pilot), Thijs Leusink, Ger Paulussen, Erik Proper en Arnoud Vermij.

Naast de deelnemers aan de experimenten zou ik Guido Bakema willen bedanken voor het beschikbaar stellen van zijn tijd. Mijn gesprek met hem heeft geleid tot het aanscherpen van mijn onderzoek. Tot slot wil ik Stijn Hoppenbrouwers bedanken voor zijn begeleiding en inhoudelijke feedback.

Joris Reijbroek,

Nijmegen 26 oktober 2006

Inhoud

1 Samenvatting.....	2
2 Voorwoord.....	3
3 Inleiding.....	6
3.1 Probleemgebied.....	6
3.2 Probleemdefinitie.....	8
3.3 Begripsbepaling.....	8
3.4 Onderzoeksvragen.....	9
3.4.1 Hoofdvraag.....	9
3.4.2 Subvragen.....	10
3.5 Afbakening.....	10
4 Methode.....	12
4.1 Populatie.....	12
4.2 Cases.....	12
4.3 Procedure.....	13
4.4 Verwerking.....	14
5 Theorie.....	16
5.1 Modelleerproces.....	16
5.2 CSDP.....	18
5.3 Modelleragenda en strategiecatalogus.....	18
5.4 Eerdere experimenten.....	20
6 Resultaten.....	22
6.1 Modelleerstijlen.....	22
6.1.1 Olievlek.....	23
6.1.2 Constraints.....	27
6.1.3 Instantie- en typeniveau.....	32
6.2 Veranderingen bij gestuurde strategie.....	33
6.2.1 Ervaren groep.....	33
6.2.2 Onervaren groep.....	34
6.3 Mate van ervaring.....	35
6.4 Voorkennis.....	35
6.5 Afleidbare informatie.....	36
7 Evaluatie onderzoek.....	37
7.1 Methode.....	37
7.2 Eigen rol.....	38
7.3 Voorleggen van strategieën.....	39
7.4 Think aloud protocol.....	40
7.5 Populatie.....	40
7.6 Vervolgonderzoek.....	40
8 Conclusie.....	42
8.1 Subvragen.....	42
8.1.1 Subvraag 1.....	42
8.1.2 Subvraag 2.....	42
8.1.3 Subvraag 3.....	43
8.2 Hoofdvraag.....	44
9 Referenties.....	46

10 Woordenlijst.....	47
11 Bijlagen.....	48
11.1 Bijlage 1: Formule 1 case.....	49
11.2 Bijlage 2: Voetbal case.....	51
11.3 Bijlage 3: Voorgelegde strategie.....	52
11.4 Bijlage 4: Aantekeningen.....	53
11.5 Bijlage 5: Foto's van een modellerproces.....	55

3 Inleiding

De ontwikkeling op het gebied van informatiemodelleren zou op sommige punten vergeleken kunnen worden met een andere belangrijke ontwikkeling op het gebied van informatietechnologie: programmeren. Toen de ICT veertig of vijftig jaar geleden nog in de kinderschoenen stond, was het programmeren van systemen voorbehouden aan een selecte groep mensen. Deze veelal academisch opgeleide programmeurs waren erg vakkundig en waren volledig bekend met de systemen waarmee ze werkten. Omdat de technologie nog niet volwassen was, waren de toepassingen beperkt en vergden zelfs de simpelste applicaties zeer veel vakmanschap om te ontwikkelen. Zaken die deze programmeurs zouden kunnen helpen zoals frameworks, uitgebreide ontwikkelomgevingen en communities, bestonden niet of nauwelijks.

Om tegenwoordig een applicatie te ontwikkelen is niet meer *hetzelfde* vakmanschap vereist als van deze eerste pioniers. Om een applicatie te ontwikkelen is het bijvoorbeeld niet meer noodzakelijk dat de programmeur van tegenwoordig direct de hardware aanstuurt waar deze applicatie gebruik van maakt. Doordat de techniek de afgelopen tientallen jaren enorm is ontwikkeld, kan de programmeur van tegenwoordig zich richten op zaken waar zijn collega's van veertig jaar geleden niet aan toe kwamen. Zaken die vaak dicht bij de gebruiker of opdrachtgever staan en minder raakvlakken hebben met onderliggende techniek.

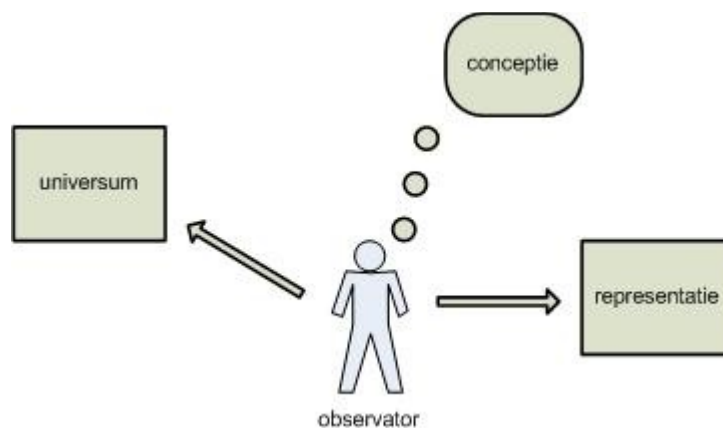
Informatiemodelleren is op dit moment nog niet zo ver ontwikkeld. Zoals ook uit dit onderzoek is gebleken vereist het goed modelleren van zelfs een simpele case nog enorm veel vakmanschap en ervaring. De vraag is vaak zelfs nog wat "goed modelleren" eigenlijk precies inhoudt.

Het onderzoek op dit gebied van de onderzoeksgroep IRIS is erop gericht om modellen, modelleren en modelleurs dezelfde ontwikkeling te laten doormaken als de programma's, het programmeren en de programmeurs de afgelopen veertig jaar hebben gedaan. Dit onderzoek draagt hieraan bij door empirische gegevens aan te leveren over verschillende modelleerprocessen. Daarnaast wordt het onderzoek zelf geëvalueerd zodat hier in de toekomst eventueel op kan worden voortgebouwd.

3.1 Probleemgebied

In de wereld van systeemontwikkeling wordt veelvuldig gebruik gemaakt van modellen. Deze modellen worden bijvoorbeeld gebruikt om te communiceren over een domein of om overeenstemming te krijgen tussen opdrachtgever en opdrachtnemer over de specificatie van een domein. In alle gevallen worden modellen gebruikt om kennis over te dragen. Kennis over het domein die gedeeld moet worden tussen de betrokken partijen.

Een model kan ontstaan als een observator een gedeelte van de werkelijkheid (“universe of discourse”) beschouwd zoals hier onder is aangegeven.



Door deze beschouwing ontstaat er een perceptie in zijn hoofd. Deze perceptie wordt geïnterpreteerd met behulp van bijvoorbeeld voorkennis en vooroordelen over het specifieke gedeelte van de werkelijkheid en hierdoor ontstaat een conceptie van de werkelijkheid. Een model ontstaat vervolgens als de observator de conceptie gaat representeren om deze te delen met anderen. Dit model is dan de representatie.

Om beter grip te krijgen op de kwaliteit van modellen zijn er verschillende klassen van doelen gedefinieerd[1]

1. Algemene doelen
Doelen die betrekking hebben op algemene zaken als bijvoorbeeld compleetheid van een model.
2. Validatie doelen
Doelen die betrekking hebben op het niveau waarop overeenstemming wordt tussen de verschillende betrokkenen.
3. Argumentatie doelen
Doelen die betrekking hebben op de argumentatie achter een bepaalde modelleerkeuzes
4. Meta-model geörienteerde doelen
Doelen die betrekking hebben op de modelleertaal die gebruikt wordt.
5. Abstractie doelen
Doelen die betrekking hebben op het juiste abstractieniveau van een model
6. Interpretatie doelen
Doelen die betrekking hebben op de mate waarin alle betrokkenen de gebruikte concepten kunnen begrijpen en gebruiken.

Een manier om deze doelen te bereiken (of in elk geval zo goed mogelijk na te streven), is door eisen te stellen aan het modelleerproces dat vooraf gaat aan het verkrijgen van het definitieve model. In een laboratoriumtest uit 2003 [4] zijn twee methoden naast elkaar getest. (zie paragraaf 5.3). Uit dit onderzoek bleek dat de rule-based methode de beste resultaten opleverde. Deze rule-based methode hield in dat het modelleerproces werd gestuurd met behulp van een aantal beperkende regels

3.2 Probleemdefinitie

Onderzoeksgroep IRIS is op dit moment bezig met het bestuderen, beschrijven en valideren van verschillende modelleerstrategieën. Deze inspanningen moeten in de toekomst leiden tot betere modelleerstrategieën en meer grip op het modelleerproces. Uiteindelijk zal dat weer resulteren in kwalitatief betere modellen.

Voordat modelleerstrategieën in de praktijk kunnen worden toegepast, is het belangrijk dat ze praktisch toepasbaar zijn en aansluiten bij de werkwijze die modelleurs gewend zijn. Hiervoor is het belangrijk dat er gegevens beschikbaar zijn over het modelleerproces van (verschillende) modelleurs en hoe deze omgaan met sturing tijdens een modelleerproces.

3.3 Begripsbepaling

Voor een goed begrip van het domein van dit onderzoek, zijn duidelijke begrippen noodzakelijk. Vaak zijn de begrippen die gebruikt worden niet geheel duidelijk of ambigu en daarom wil ik in deze paragraaf een aantal begrippen duidelijk definiëren.

Model

De definitie van een model zoals in [7] wordt gebruikt luidt “A purposely abstracted and unambiguous conception of a domain”. Vertaald betekent dit

“Een met opzet geabstraheerde en ondubbelzinnige conceptie van een domein”

Hierbij moet worden aangetekend dat deze definitie niets zegt over de representatie van een model. Bij het gebruik van het begrip “model” in deze scriptie gaat het over de fysieke representatie van een model zoals gebruikt in [7].

Modelleerproces

Aan elk conceptueel model is een modelleerproces vooraf gegaan. Tijdens dit proces proberen de betrokken partijen overeenstemming te bereiken over een gemeenschappelijk model van een deel van de werkelijkheid. Elke vorm van modelleren kent een modelleerproces en hoeft niet noodzakelijkerwijs een plan of sturing te bevatten. Dit kan natuurlijk wel.

Modelleerstijl

Een modelleerstijl is een modelleerproces of een onderdeel van een modelleerproces met specifieke kenmerken. Als er één modelleerproces wordt beschouwd heeft dit proces natuurlijk altijd eigen, specifieke kenmerken en daarom worden modelleerstijlen pas interessant om te bestuderen als deze voorkomen in meerdere modelleerprocessen.

Hoewel het begrip “modelleerstijl” kan worden verward met modelleerstrategie, wil ik hier toch een duidelijk onderscheid in aanbrengen. De modelleerstijlen zoals hier bedoeld zijn alleen achteraf te bepalen en hebben niets te maken met een vooropgesteld plan (zoals modelleerstrategieën dat wel hebben). Dit onderscheid is in het kader van mijn onderzoek belangrijk omdat ik de impliciete modelleerstijlen die aanwezig zijn in de populatie van mijn onderzoek, vergelijk met de expliciete modelleerstrategieën zoals deze bijvoorbeeld worden ontwikkeld binnen de onderzoeksgroep IRIS.

Het is mogelijk dat één modelleerproces verschillende modelleerstijlen bevat. Verschillende stijlen kunnen elkaar overlappen maar ze kunnen elkaar niet uitsluiten.

Voorbeeld van overlap

Als een modelleur de gewoonte heeft om eerst de feiten uit een domein te bepalen en daar de objecten aan te relateren (stijl 1) zegt dit nog niets over het moment waarop hij totale-rol constraints toevoegt (stijl 2).

Voorbeeld van uitsluiting

Tijdens één modelleerproces is het niet mogelijk dat een modelleur alle constraints bepaalt nadat alle feittypen en objecttypen bekend (stijl 1) *en* alle constraints direct plaatst nadat het betreffende feittypen en objecttypen zijn geformuleerd (stijl 2)

Modelleerstrategie

Modelleerstrategieën zijn doelgerichte, vooropgezette, regels en principes die bedoeld zijn om een modelleerproces richting en sturing te geven. Belangrijk hierbij is dat een strategie niet “beknellend” moet werken maar ruimte laat voor de expertise en vakmanschap van de systeem analist. Een modelleerstrategie is volgens deze definitie geen modelleerproces dat met behulp van regels en principes is uitgevoerd. Het gaat bij een strategie om de regels en de principes zelf.

3.4 Onderzoeksvragen

Zoals in de probleemdefinitie is aangegeven, is er behoefte aan informatie die kan helpen bij het opstellen van bruikbare strategieën. Deze informatie wil ik tijdens dit onderzoek formuleren met behulp van een aantal vragen.

3.4.1 Hoofdvraag

Hoe verhoudt de theorie over modelleerstrategieën, zoals gedefinieerd binnen de onderzoeksgroep IRIS, zich met de modelleerstijlen die in de praktijk worden toegepast?

Belangrijk om bij deze vraag op te merken is dat ik geen uitspraken doe over de mate waarin modelleerstijlen voorkomen of de effectiviteit ervan, maar puur wil ingaan op de relatie tussen mijn bevindingen en de theorie. Met behulp van mijn bevindingen kan de bestaande theorie worden gevalideerd.

3.4.2 Subvragen

De hoofdvraag wordt ondersteund door een aantal subvragen. De antwoorden op deze subvragen helpen vervolgens het antwoord op de hoofdvraag te ondersteunen.

1. Welke modelleerstijlen zijn er te onderscheiden tijdens de experimenten?
Tijdens de experimenten ben ik op zoek gegaan naar manieren van modelleren die bij een aantal deelnemers terugkwamen.
2. In welke mate heeft de ervaring van de deelnemer invloed op het modelleerproces?
Het onderzoek is uitgevoerd onder twee verschillende groepen: ervaren en onervaren groep. Bij de beantwoording van deze vraag wordt ingegaan op de verschillende manieren van modelleren van de twee groepen.
3. Wat gebeurt er als er van vrij modelleren wordt overgeschakeld op een strategie?
Beide groepen deelnemers hebben een case vrij en gestuurd op moeten lossen. Het antwoord van deze vraag zal ingaan op de verschillen tussen deze twee manieren van modelleren.
4. Hoe kan een theorie die betrekking heeft op modelleerstijlen empirisch getest worden?
Deze subvraag staat redelijk los van de rest van het onderzoek. Deze vraag gaat *over* het onderzoek in plaats van dat hij deel uit maakt van het onderzoek. Hierdoor heb ik een apart hoofdstuk (hoofdstuk 7) gewijd aan de beantwoording.

3.5 Afbakening

Een modelleerproces kan erg breed worden opgevat. Tijdens het eerste contact tussen domeinexpert en systeemanalist zou men bijvoorbeeld al van een modelleerproces kunnen spreken. Dit is echter niet waarop ik mij heb gericht. Mijn onderzoek richt zich op een fase waarin de systeemanalist en domeinexpert tot een expliciet, gezamenlijk model proberen te komen. Hoewel deze fase in de praktijk kan worden opgevat als een conversatie waar veel interactie is tussen de systeemanalist en de domeinexpert, ligt de nadruk in mijn onderzoek op de taak van de systeemanalist (zie paragraaf 5.1).

Tijdens de beoordeling van de opgeleverde modellen wordt er geen gebruik gemaakt van objectieve kwaliteitseisen. De nadruk van het onderzoek ligt op het proces van het modelleren en minder op de opgeleverde modellen.

4 Methode

Om empirische gegevens te vergaren heb ik een aantal experimenten uitgevoerd. Tijdens deze experimenten heb ik een aantal proefpersonen een tweetal cases laten modelleren in ORM. De ORM modellen zijn, samen met opnamen van de uitwerking (audio en foto materiaal) en een beschrijving van de houding en opinie van de deelnemer, geanalyseerd zodat een uitspraak kan worden gedaan over het modellerproces.

4.1 Populatie

In één van de deelvragen wordt gevraagd of er verschil is tussen de modelleerstijlen van ervaren en onervaren modelleers. Om deze vraag te kunnen beantwoorden is de populatie van het onderzoek verdeeld in twee groepen:

- **Studenten**
Deze groep bestaat uit studenten die allemaal het vak Domeinmodelleren of Modelleren van Organisaties op de Radboud Universiteit hebben gevolgd, zodat ze over voldoende ORM kennis beschikken om de cases te modelleren.
- **Professionals**
De groep bestaat uit hoogleraren en docenten van de Radboud Universiteit. Zij maken in hun dagelijks werk regelmatig gebruik van ORM en andere modelleertechnieken en beheersen ORM goed.

Deze twee groepen zijn interessant omdat ze wat betreft modelleerervaring duidelijk te onderscheiden zijn, maar toch allemaal ORM beheersen. Uit praktische overwegingen zijn de experimenten uitgevoerd bij negen proefpersonen (5 ervaren en 4 onervaren). Idealiter zou voor een dergelijk onderzoek een grotere populatie onderzocht moeten worden, maar dat is gezien de omvang van de afstudeeropdracht niet mogelijk. Bovendien is gebleken dat zelfs deze beperkte populatie al een aantal interessante verschijnselen aan het licht brengt.

Bij één van de deelnemers is een pilot experiment uitgevoerd. Hoewel er na deze pilot een aantal wijzigingen zijn doorgevoerd, zijn de gegevens uit dit experiment wel verwerkt in mijn onderzoek. De wijzigingen waren niet zodanig dat de uitslag van de pilot niet in het onderzoek opgenomen kon worden.

4.2 Cases

Het onderzoek is uitgevoerd door de deelnemers twee verschillende cases te laten modelleren. Deze cases [bijlagen 1 en 2] hebben betrekken op redelijk eenvoudige onderwerpen die door alle deelnemers makkelijk en snel konden worden geïnterpreteerd. Op die manier is ervoor gezorgd dat de voorkennis van alle deelnemers zoveel mogelijk gelijk was op het moment dat ze aan het modelleren begonnen.

De deelnemers moesten tijdens de experimenten de cases op twee verschillende manieren oplossen (hier wordt later nog op ingegaan). Het was niet mogelijk om voor beide manieren dezelfde case te gebruiken. In dat geval zou de domeinkennis die tijdens het eerste modellerproces is opgedaan, worden gebruikt voor het tweede experiment. Om dit probleem te omzeilen zijn er twee verschillende cases gebruikt die qua complexiteit en omvang vrijwel identiek zijn.

Om toch eventuele verschillen in voorkennis op te vangen, is er aan de deelnemers van tevoren gevraagd of ze bekend waren met het domein. Dit moesten ze aangeven met een cijfer van 1 tot en met 10. Het zou namelijk mogelijk kunnen zijn dat door gebrek aan kennis van het domein, een deelnemer minder eenvoudig een ORM diagram kan opstellen. Door de kennis van tevoren te kwalificeren kan hiermee rekening worden gehouden in de analyse achteraf.

De deelnemers hebben tijdens het onderzoek twee verschillende cases gekregen, die ze op twee verschillende manieren op hebben moeten lossen:

- Vrij
Bij deze case was de deelnemer vrij om zijn eigen aanpak te kiezen en was het toegestaan om zelf keuzes te maken bij het opstellen van het ORM diagram. Voorafgaand aan het modelleren van deze case, werd expliciet aangegeven dat de deelnemer zijn eigen manier van werken moest volgen. Het doel van deze fase is vooral om erachter te komen hoe het modellerproces van de deelnemers normaliter verloopt (er vanuit gaande dat de deelnemers geen strategie gebruiken als zij iets modelleren).
- Gestuurd
Bij deze case werd een stappenplan [bijlage 3] uitgereikt die de deelnemer moet volgen. Het stappenplan is gebaseerd op een “strict fact based” strategie die bijvoorbeeld ook bij FCO-IM wordt gebruikt. De strategie is mede tot stand gekomen door de input van Guido Bakema. Naar aanleiding van een gesprek met hem, heb ik de strategie die gebruikt is tijdens de pilot verbeterd.

De deelnemer kreeg van tevoren ruim de tijd om een idee te vormen van de strategie en om vragen te stellen. Anders dan tijdens de vrije case, is het tijdens deze case niet van belang hoe de deelnemers gewend zijn om te modelleren. Tijdens deze fase werd nagegaan hoe de deelnemers om gaan met een verplichte werkwijze.

De deelnemers moesten de eerste case modelleren naar eigen inzicht, bij de tweede case moest de strategie worden gevolgd. Er is voor deze volgorde gekozen om de strategie geen invloed te laten hebben op de vrije case. Als de deelnemers eerst de case met behulp van de strategie hadden moeten oplossen en daarna de vrije case, was het mogelijk geweest dat ze elementen uit de strategie hadden gebruikt in hun eigen manier van modelleren.

4.3 Procedure

Alle experimenten vonden plaats in een afgesloten ruimte waar alleen de deelnemer en ikzelf aanwezig waren. Op deze manier werden de deelnemers niet afgeleid of gestoord. De duur van de experimenten liep erg uiteen. Een aantal ervaren deelnemers had slechts een half uur beschikbaar om mee te doen aan het experiment. Toch bleven ze vaak ruim binnen die tijd. Onervaren deelnemers hadden soms meer dan anderhalf uur nodig om tot een vergelijkbaar model te komen.

Voorafgaand aan het experiment heb ik aan alle deelnemers een aantal zaken expliciet vermeld. Deze zaken moesten worden verteld om de deelnemer duidelijkheid te geven het experiment en zijn taak hierin. De volgende zaken werden voor elk experiment uitdrukkelijk gezegd.

- “Er worden twee cases voorgelegd”
- “Tijdens de eerste case mag er gemodelleerd worden zoals je gewend bent.” Hierbij geef ik expliciet aan dat de deelnemer bijvoorbeeld opnieuw mag beginnen, meerdere modellen mag gebruiken etc.
- “De tweede case wordt opgelost volgens een stappenplan.” (Wat dit stappenplan precies inhoud, vertel ik in deze fase nog niet.)
- “Lees de case eerst rustig door voordat je begint”
- “Denk tijdens het hele experiment hardop”

De modellen werden opgesteld met behulp van een potlood en papier. Om ervoor te zorgen dat de deelnemers genoeg ruimte hadden om de modellen uit te werken, had het papier een A2 formaat. Voorbeeldzinnen, schetsen en het model konden zo op één werkblad worden weergegeven wat het overzicht ten goede kwam.

Tijdens het oplossen van de cases nam ik de rol van domeinexpert op me. De deelnemers konden vragen aan mij stellen als er in hun ogen onduidelijkheden waren. Behalve het beantwoorden van deze vragen, probeerde ik ook wedervragen te stellen om erachter te komen tegen welke problemen de deelnemers aanliepen.

Het modellerenproces zelf is vastgelegd met behulp van aantekeningen, foto's, audio-opnamen en de modellen zelf. Met behulp van deze gegevens was het mogelijk om achteraf de opgeleverde modellen te vergelijken en het modellerenproces te analyseren. Aan alle deelnemers is gevraagd om “hard op te denken” (think aloud protocol [5]). Hierdoor was het mogelijk om erachter te komen wat de gedachtengang was van de deelnemers als zij bepaalde acties uitvoerden.

4.4 Verwerking

De belangrijkste gegevens in dit onderzoek vormen de waarnemingen en de het materiaal van de modellerprocessen. Deze zijn na afloop van de onderzoeken aandachtig geanalyseerd en in de context van de theorie geplaatst. De belangrijkste vorm van de verwerking van de resultaten van dit onderzoek, begint al bij het afnemen van de experimenten. Door de ervaring die is opgebouwd na het afnemen van een aantal experimenten, kan er tijdens een proef beter worden doorgevraagd naar bepaalde fenomenen. De reactie van de deelnemer hierop kan weer leiden tot nieuwe inzichten.

De verwerking van de experimenten achteraf gebeurde vooral door het analyseren van de aantekeningen en het opnieuw afluisteren van de geluidsopnamen die tijdens de experimenten zijn gemaakt. Door de data van de verschillende experimenten te vergelijken konden ook nieuwe verbanden en fenomenen aan het licht komen die tijdens de proef zelf niet werden opgemerkt.

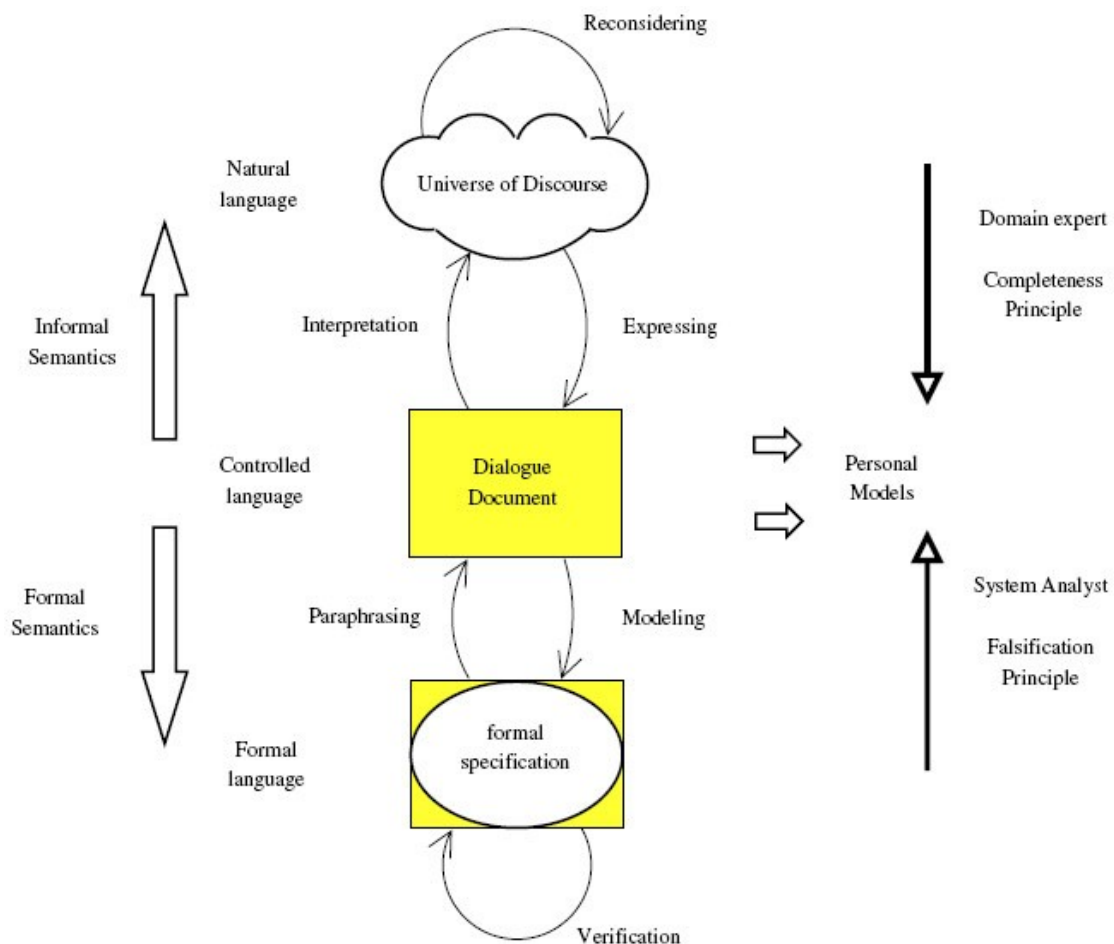
De informatieverwerking die nodig was om subvraag 4 te beantwoorden, was pas beschikbaar op het moment dat alle gegevens van de experimenten zelf waren verwerkt. Subvraag 4 is immers een evaluatie van het onderzoek en de gevolgde methode. Pas nadat alle resultaten van het onderzoek duidelijk waren, was het mogelijk om de sterke en zwakke kanten van het onderzoek in kaart te brengen en verbeteringen aan te dragen voor een eventueel vervolg.

5 Theorie

Hoewel de kern van dit afstudeeronderzoek wordt gevormd door de uitslagen van de experimenten, gaat dit hoofdstuk kort in op de theoretische achtergrond van het onderwerp. Met name wordt ingegaan op de theorie die op dit moment bestaat over het structureren en onderzoeken van modelleerprocessen.

5.1 Modelleerproces

Zoals bij de begripsbepaling is aangegeven, kan elk proces dat leidt tot een model worden beschouwd als een modelleerproces. Om toch een beter begrip te krijgen van de acties, rollen en onderdelen waar een modelleerproces uit kan bestaan, is in [8] een voorbeeld gegeven in de vorm van de volgende, schematische weergave.



Hoewel sterk versimpeld, geeft het diagram goed weer waaruit een modellerproces kan bestaan. Er worden twee partijen geschetst die samen tot een gezamenlijk document komen, te weten het dialoog document (dialogue document). De twee partijen hebben ieder hun eigen belangen en eigenschappen:

- Domein expert (DE). De domeinexpert levert correcte informatie over het domein en bepaalt welk gedeelte van de wereld het UoD (universe of discourse) is. De domein expert is gewend om in natuurlijke taal te communiceren en heeft er belang bij om zoveel mogelijk relevante informatie uit het domein in het dialogue document te krijgen. Bovendien moet hij in staat zijn om beweringen te doen en beweringen van de systeem analist te valideren die over het UoD gesteld kunnen worden.
- Systeem analist (SA). De systeem analist ontwerpt het conceptueel model en is verantwoordelijk voor de juiste keuzes van meta-concepten. Hij/zij stelt het model op in een formele taal en moet ervoor zorgen dat dit model geen tegenstrijdigheden of fouten bevat.

In het dialoog document wordt een zogenaamde gecontroleerde taal (controlled language) gebruikt. Deze taal bestaat uit een versimpelde versie van een natuurlijke taal (bijvoorbeeld het Nederlands) en bevindt zich op het punt waar formele talen en natuurlijke talen elkaar ontmoeten. De syntax van een dergelijke is relatief simpel en ligt vast, net als bij een formele taal. Anders dan bij een formele taal kunnen de verschillende partijen (en dan met name de domein expert) zich prima uitdrukken in deze taal omdat hij dicht genoeg bij natuurlijke taal ligt. De formele, gecontroleerde en natuurlijke talen waren tijdens de experimenten van dit onderzoek duidelijk te onderscheiden:

Formeel: ORM modellen

Gecontroleerd: Voorbeeld zinnen

Natuurlijk: Voorgelegde cases

Zoals de dunne pijlen in het bovenstaande figuur aangeven hebben de betrokken partijen een aantal mogelijkheden tijdens het modellerproces:

- Systeem analist:
 - Paraphrasing: Het verwoorden van de formalisatie naar het dialoog document.
 - Modelling: Het formaliseren van het dialoog document
 - Verification: Het nagaan of de formalisatie correct is en het verbeteren van de formalisatie.
- Domein expert
 - Interpretation: Het lezen van het dialoog document om na te gaan of deze klopt met het domein
 - Expressing: Het uitdrukken van het domein om tot een dialoog document te komen.
 - Reconsidering: Het heroverwegen van de eigenschappen van het domein.

Tijdens een modelleerproces staat de modelleur constant voor allerlei keuzes. De keuzes die hij maakt beïnvloeden behalve het uiteindelijke model, ook de stappen en de keuzes die hij vanaf dat moment moet maken. In [1] zijn deze keuzes onderverdeeld in een aantal categorieën

- Meta-model keuzes
Keuzes die gemaakt worden met betrekking tot het meta-model waarmee de modelleur werkt.
- Modelleer procedure keuzes
Keuzes die betrekken hebben op de procedure waarbinnen de concepten uit het meta-model worden gebruikt.
- Keuzes die betrekking hebben op beweegredenen
Deze keuzes hebben betrekking op de reden waarom het meta-model op een bepaalde manier wordt gebruikt.

5.2 CSDP

Terry Halpin e.a. hebben bij de ontwikkeling van ORM een procedure ontwikkeld die als leidraad kan worden gebruikt tijdens het modelleerproces. Deze procedure heet CSDP (Conceptual Schema Design Procedure) [3] en bestaat uit de volgende stappen.

1. Transformeer informatievoorbeelden naar elementaire feiten en voer kwaliteitscontroles uit.
2. Teken de feittypen en voer een populatie controle uit.
3. Controleer op entiteitstypen die gecombineerd kunnen worden en let op rekenkundige afleidingen.
4. Voeg uniciteit constraints toe en controleer de ariteit van de feittypen.
5. Voeg verplichte rol constraints toe en controleer op logische afleidingen.
6. Voeg waarde, set vergelijking en subtype constraints toe.
7. Voeg overige constraints toe en voer laatste controles uit.

Er bestaan verschillende versies van deze procedure maar ze leggen allemaal de nadruk op drie verschillende aspecten:

1. Het verbaliseren in de vorm van elementaire feiten
2. Populatie checks uitvoeren
3. Analyse van business rules

Deze procedure is opgesteld vanuit educatief oogpunt. Hij moet worden gezien als een manier waarop beginnende modelleurs met ORM aan de slag kunnen gaan. Het geeft op een redelijk hoog abstractieniveau weer hoe een modelleerproces het best kan worden ingericht. Het helpt de modelleur met het aanhouden van een volgorde om bepaalde stappen uit te voeren maar laat ook een hoop ruimte over.

5.3 Modelleragenda en strategiecatalogus

Een manier om een modelleerstrategie vast te leggen voor een meta-model geïntegreerde strategie, is door gebruik te maken van een modelleeragenda zoals in [1] is beschreven. Voor het creëren van een modelleeragenda worden een aantal regels en beperkingen gebruikt die richting geven aan een modelleerproces. Deze regels zijn een uitdrukking van het meta-model van de modelleertaal (ORM). Bijvoorbeeld:

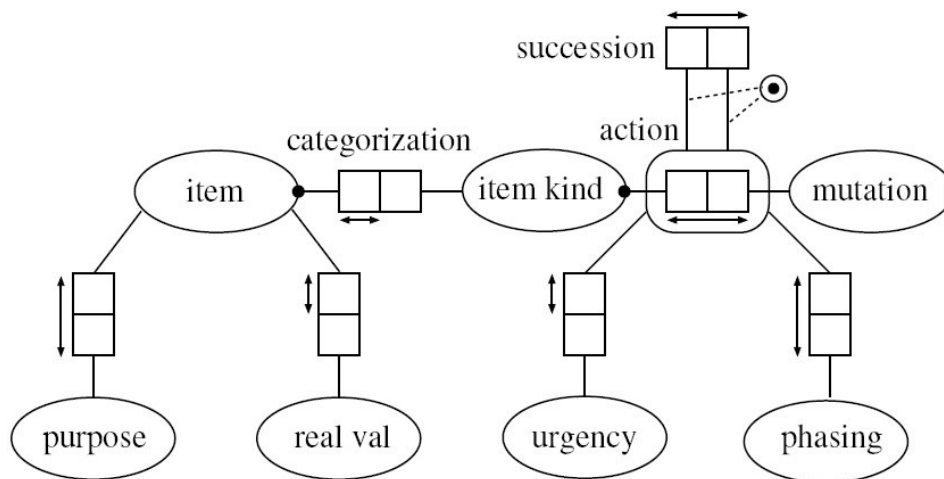
1. Elk objecttype is gepopuleerd door één of meerdere objecten
2. Elk object kent één of meerdere feiten.

Voordat deze regels richting kunnen geven aan een modelleerproces (en er dus sprake kan zijn van een strategie) moeten er beperkingen op de toepassing van dergelijke regels worden gelegd. Dergelijke regels stellen immers wel eisen aan een model, maar niet aan een proces. Hiervoor zijn zogenaamde “temporele beperkingen” nodig. Bijvoorbeeld:

1. Een actie die bestaat uit het toevoegen van een object wordt onmiddellijk gevolgd door een actie die bestaat uit het toevoegen van het objecttype dat wordt gepopuleerd door dat object
2. Geen actie die bestaat uit het toevoegen van een object vindt plaats voordat er een actie plaats vindt die bestaat uit het toevoegen van een feit dat dat object bevat.

Aan het begin van het modelleerproces kunnen uit dergelijke voorschriften mogelijke eerste stappen worden afgeleid. Bij “strict fact-based modelling” zou dit bijvoorbeeld in kunnen houden dat er een feit moet worden toegevoegd aan het model. Het toevoegen van een feit zou dan weer een ander voorschrift kunnen *triggeren*. Het zou ook kunnen zijn dat er een aantal voorschriften ontstaan waar de modelleur een keuze uit kan maken. Tijdens het modelleren wordt op deze manier een modelleeragenda gegenereerd die de modelleur sturing geeft op basis van zijn acties. Waar een bepaalde strategie één bepaalde actie *vereist*, wordt een dergelijke actie verplicht gesteld door de modelleeragenda. Op het moment dat een strategie geen eisen stelt aan een volgende actie of een aantal acties mogelijk zijn, zal de modelleeragenda een stap voorleggen zonder prioritering.

Om een modelleeragenda te kunnen genereren wordt in [1] gebruik gemaakt van een strategiecatalogus. In een strategiecatalogus kan een verzameling strategieën worden vastgelegd. Een voorbeeld van een strategiecatalogus die wordt voorgesteld in [1]:



Met behulp van een strategiecatalogus is het mogelijk om strategieën in te delen aan de hand van bepaalde kenmerken over bijvoorbeeld volgorde, onderdelen en doel. Waar mogelijk wil ik met behulp van een beknopte strategiecatalogus een aantal modellerestijlen proberen vast te leggen.

5.4 Eerdere experimenten

In de literatuur komen enkele voorbeelden [4] voor van experimenten die betrekking hebben op het modelleren van een bepaalde case. Hoewel deze experimenten andere doelen nastreven dan dit onderzoek noem ik ze hier vanwege de gebruikte methoden.

In [4] is een laboratorium experiment uitgevoerd om te bepalen welke van twee manieren om data te modelleren het meest effectief was voor het trainen van onervaren modellers:

1. Rule-based approach

De Rule-based approach voorziet in een aantal regels (rules) die gevolgd kunnen worden om tot een optimaal model te komen. Deze regels zijn generieker en abstracter dan de patterns

2. Pattern-based approach

Bij deze benadering wordt gebruik gemaakt van templates waarbij de omschrijving van de objecten, relaties en attributen al vastligt voor een bepaald domein.

In dit onderzoek werden in totaal 27 (redelijk onervaren) studenten opgedeeld in twee verschillende groepen. De eerste groep heeft drie verschillende cases opgelost met behulp van de rule-based approach en de tweede groep met behulp van de pattern-based approach. De cases verschilden in mate van complexiteit. De cases werden naarmate het experiment vorderde steeds complexer.

Aan de hand van deze experimenten werd bepaald welke aanpak tot de beste resultaten leidde. Om de kwaliteit van de modellen te bepalen werden alle uitslagen uitvoerig nagekeken met behulp van een schema. In dit schema werden strafpunten gegeven als er onderdelen ontbraken of onjuist waren gebruikt. Hierdoor was het mogelijk om redelijk objectief vast te stellen welke aanpak de hoogste kwaliteit schema's opleverde.

In dit onderzoek lag de nadruk duidelijk op het eindproduct van de experimenten. Het proces dat de studenten doorliepen om tot het eindproduct te komen werd niet bestudeerd. Aan het proces zelf werden, behalve door het voorschrijven van een bepaalde aanpak (rule-based of pattern-based), geen eisen gesteld. Door het stellen van bepaalde eisen aan het eindproduct werd objectief vastgesteld welke aanpak tot het beste resultaat leidde.

6 Resultaten

In dit hoofdstuk worden een aantal waarnemingen noemen die ik tijdens mijn experimenten ben tegengekomen. Anders dan bij de conclusie wil ik deze hier nog niet verbinden aan mijn onderzoeksvragen. In het kader van het exploratieve karakter van mijn onderzoek, lijkt het mij verstandig om alle waarnemingen te vermelden om op deze manier zoveel mogelijk informatie uit mijn experimenten vast te leggen.

6.1 Modellerstijlen

Het vaststellen van modellerstijlen kan op verschillende niveaus van detail worden gedaan. Vanzelfsprekend wordt één case door twee verschillende deelnemers, op twee verschillende manieren, opgelost en zijn de acties¹ die de deelnemers uitvoeren telkens iets anders. Als twee modellerprocessen erg gedetailleerd worden beschouwd, is het heel moeilijk om daar modellerstijlen in te herkennen omdat het modellerproces op dat niveau dusdanig verschilt dat er weinig overeenkomsten te herkennen zijn.

Als er echter naar een hoger niveau wordt gekeken zijn er wel degelijk modellerstijlen te herkennen. Het gaat op dat niveau niet om de exacte acties die de systeemanalist na elkaar uitvoert, maar om de gedachte achter die acties. Het is opgevallen dat veel deelnemers om ongeveer dezelfde redenen toch net iets andere acties ondernemen. Zo kunnen twee verschillende modellerprocessen volgens ongeveer dezelfde gedachte verlopen maar leveren ze uiteindelijk toch andere modellen op.

De gevonden modellerstijlen worden later meer concreet gemaakt door ze vast te leggen op de manier die in [1] wordt beschreven. Dat artikel beschrijft echter een techniek om modellerstrategieën te definiëren en niet modellerstijlen. Hierdoor ontstaat het praktische probleem dat een modellerstrategie in [1] onderdelen bevat die bij een modellerstijl moeilijk achteraf te bepalen zijn. Bij het beschouwen van een modellerstijl achteraf is het echter niet na te gaan wat de gedachte van de deelnemer was op het moment dat hij bijvoorbeeld een feittype aan een objecttype toevoegde. Met andere woorden: een structuur voor het vastleggen van een strategie (een vooropgesteld plan) kan niet één op één worden overgenomen voor het vastleggen van een stijl.

Bij het vastleggen van een stijl wordt daarom niet de structuur uit [1] gebruikt, maar alleen de onderdelen die relevant zijn voor de kenmerken die de stijl bevat.

1 Met een actie wordt bedoeld: het toevoegen, verwijderen of wijzigen van een ORM element uit het metamodel van ORM

6.1.1 Olievlek

Een belangrijke stijl die valt te onderscheiden, werd door één van de deelnemers “modelleren als een olievlek” genoemd. Tijdens het vrije modelleren is die stijl bij alle deelnemers te herkennen. Bij deze stijl probeert de systeemanalist na het bestuderen van de case één centraal objecttype te verzinnen. Vanuit dit object breidt hij het ORM diagram telkens uit doormiddel van het toevoegen van feittypen en meer objecten.

In het geval van de cases die bij de experimenten werden gebruikt, werd vaak begonnen met één van de volgende objecten:

Bij de formule 1 case [bijlage 1]:

- Coureur
- Grandprix

Bij de voetbalcase [bijlage 2]:

- Wedstrijd
- Team

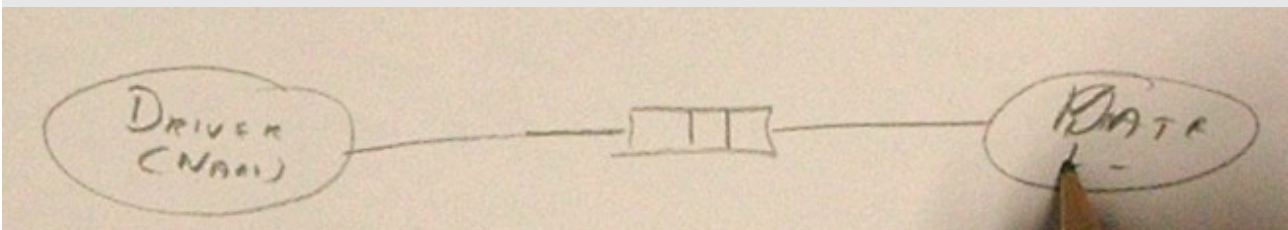
Bij navraag bleek de keuze voor het meest geschikte object erg subjectief te zijn en kwam vaak voort uit de voorkennis van het domein. Deelnemers gaven aan dat het object waarmee werd begonnen naar hun idee een centrale rol speelde in het domein.

Dat de deelnemer in het begin een centraal object probeert vast te stellen, betekent niet dat dit object zijn centrale rol in het ORM diagram gedurende het modelleren behoudt. Het kwam vaak voor dat naarmate het diagram vorderde, het eerste object niet de centrale rol speelde die de deelnemers aanvankelijk hadden verwacht.

Voorbeeld:

Bij de formule 1 case [bijlage 1] begonnen veel deelnemers met het objecttype “coureur”. De namen van de coureurs komen immers in beide tabellen voor. Als de opgeleverde modellen achteraf worden bekeken, is te zien dat de meeste rollen worden gespeeld door het objecttype “grandprix”.

Dit voorbeeld wordt nog duidelijker met behulp van de volgende foto's. De foto's zijn gemaakt tijdens een experiment met een ervaren deelnemer, zonder sturing. De eerste foto is gemaakt aan het begin van het modelleerproces.



De deelnemer begon hier met het centrale objecttype "driver", geïdentificeerd door zijn naam (name). Tussen het objecttype "driver" en het objecttype "date" bestaat blijkbaar een ternaire rol die hier nog geen naam heeft.

De volgende foto is genomen aan het einde van hetzelfde experiment.

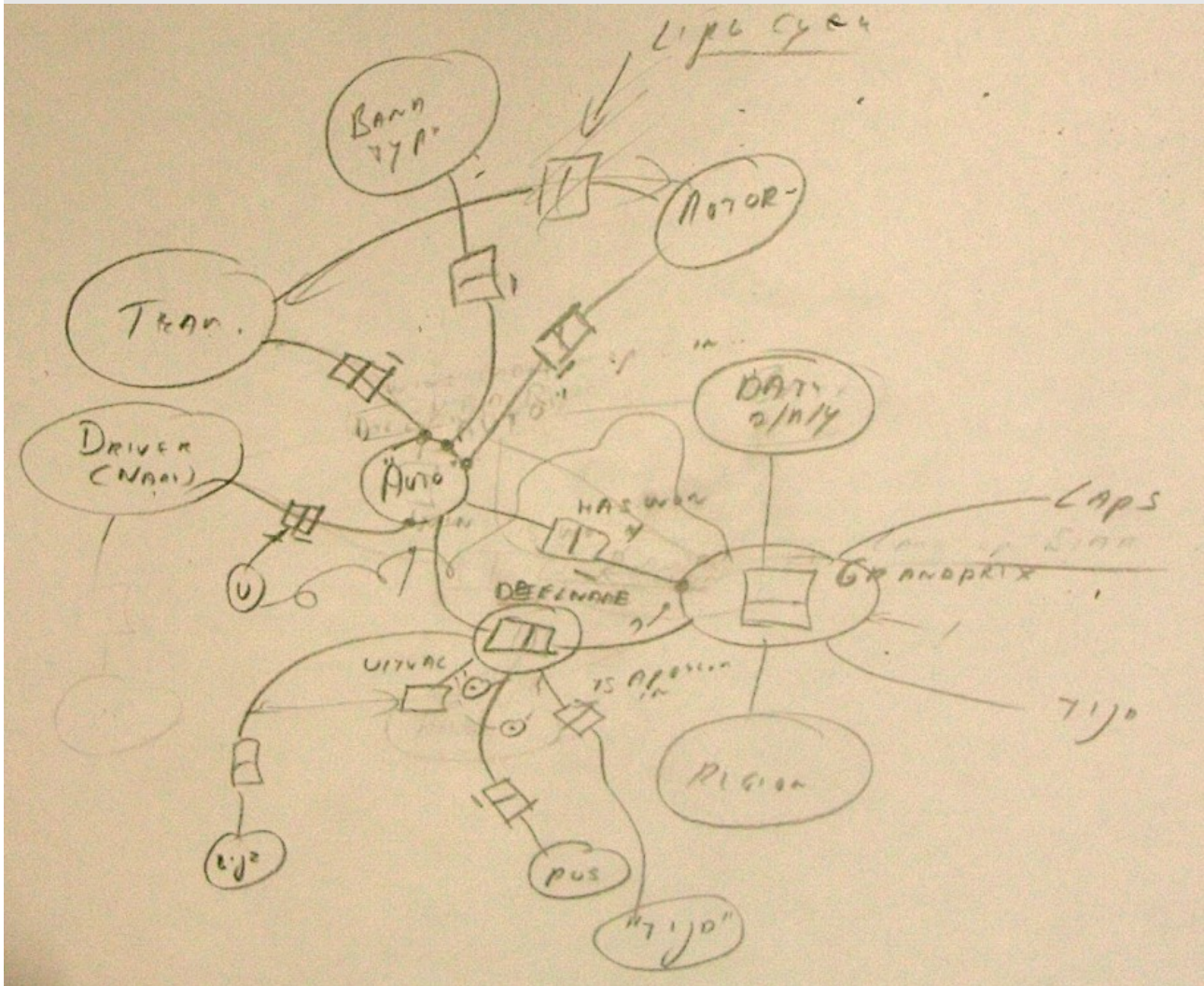
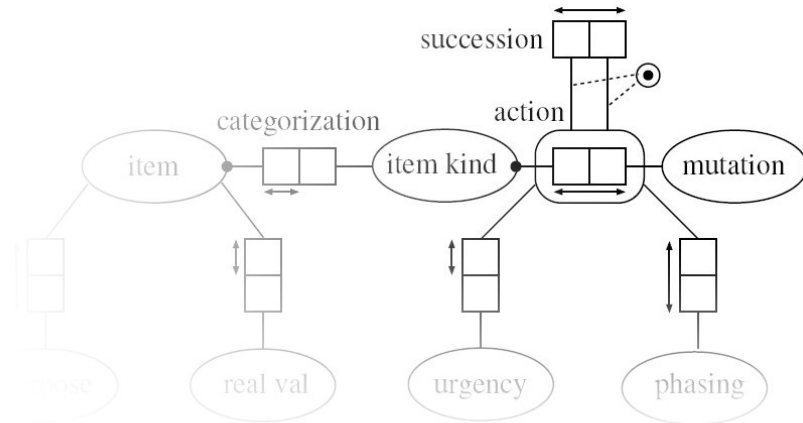


Foto 2

In deze situatie hebben de objecttypen "driver" en "date" geen gezamenlijke rol meer. "Date" speelt andere rollen en wordt geobjectificeerd tot "grandprix". Het objecttype "driver" speelt niet de centrale rol die de deelnemer op de vorige foto voor ogen had. Dit objecttype is niet meer rollen gaan spelen dan dat deze in de eerste situatie deed, namelijk één. De kernen van het bovenstaande model worden gevormd door "auto" en "grandprix", niet door "driver".

Een modelleerstijl zoals “de olievlek” omvat lang niet alle elementen en mutaties uit het meta-model van ORM. Het gaat slechts over de toepassing van objecttypen en feittypen. In de onderstaande afbeelding van de strategiecatalogus zijn de irrelevante onderdelen voor deze stijl daarom minder duidelijk aangegeven.



Het belangrijkste onderdeel van de strategiecatalogus voor deze stijl is de opvolging (succession) tussen het plaatsen van de objecttypen en de feittypen. Een mogelijke populatie van de strategiecatalogus door deze stijl is hieronder weergegeven. Doordat een “action” een geobjectificeerd feittype is, vormt de eerste tabel de acties en de tweede tabel de volgorde van deze acties (succession). De acties worden in de tweede tabel geïdentificeerd door het nummer dat ze meekrijgen in de eerste tabel.

action	item kind	mutation
1	objecttype	add
2	feittype	add

Omdat het label “succession” niet duidelijk aangeeft welke actie op welke andere actie volgt, wordt dit label vervangen door het meer duidelijke “is followed by”.

is followed by	
1	1
1	2
1	2
1	2

De uniciteit constraint van het feittype “succession” sluit een dergelijke populatie echter uit. Dit komt doordat de strategiecatalogus *alle mogelijke* opvolgingen bevat en niet *alle* opvolgingen. De kern van deze stijl is dat er na de fase zoals op foto 1 is weergegeven een aantal herhalingen plaats vinden. De eerste twee populaties van het bovenstaande feittype vormen de *mogelijke* opvolgingen van de stijl. De gehele populatie vormt *alle* opvolgingen die voorkomen in een modelleerproces waarbij een model met vier objecttypen en drie feittypen wordt gemodelleerd waarin de stijl “olievlek” te herkennen is.

Uit de bovenstaande populaties wordt het verschil tussen modelleerstijl en modelleerstrategie nog eens duidelijk. De bovenstaande populatie is absoluut onbruikbaar om richting te geven aan een modelleerproces, en dus om te gebruiken als een strategie. Als “de olievlek” zou worden beschouwd als een strategie en de bovenstaande populaties richting zouden moeten geven aan die strategie kan dit leiden tot een ORM diagram met alleen maar objecttypen. Het toevoegen van een objecttype (actie 1) kan immers gevolgd worden door het toevoegen van een objecttype (actie 1) dat weer gevolgd kan worden door toevoegen van een objecttype (actie 1) etc.. Andersom geven de populaties wel aan welke acties en welke opvolgingen er voorkomen in de beschreven stijl. Het populieren van de strategiecatalogus wordt hier dus gebruikt om een stijl vast te leggen en niet om een strategie te definiëren.

6.1.2 Constraints

Het aanbrengen van constraints (beperkingsregels) zorgt ervoor dat de populatie van een ORM diagram aan bepaalde eisen voldoet. Meestal (o.a. in het CSDP [3]) gebeurt dit nadat het diagram alle objecttypen en feittypen bevat. Tijdens de experimenten waren duidelijk twee verschillende momenten te onderscheiden waarop de deelnemers constraints aan het diagram toevoegden. Daarnaast viel op dat onervaren deelnemers anders met constraints omgingen dan ervaren deelnemers.

Tijdens het modelleren van de vrije case werden sommige constraints direct na het toevoegen van een item geplaatst. De deelnemers wachtten dus niet totdat ze het hele diagram hadden voltooid, maar modelleerden de constraints direct. Bij navraag bleek dat de deelnemers bepaalde constraints meteen plaatsten omdat deze heel makkelijk in het UoD te herkennen waren. Sommigen gaven zelfs aan dat de betreffende constraint zouden vergeten als ze die niet direct zouden plaatsen. Dit kwam voor bij zowel de ervaren als de onervaren deelnemers.

Voorbeeld

Het begrip “winnaar” was voor een aantal deelnemers een reden om direct over constraints na te denken. Dat komt doordat het algemeen bekend is dat een formule 1 race maar één winnaar heeft. Deze overduidelijke relatie tussen race en winnaar zorg ervoor dat veel deelnemers het gevoel hebben dat ze zo'n duidelijke beperking meteen in het diagram moesten opnemen.

Alle constraints die niet direct opvielen, kwamen aan bod op het moment dat alle objecttypen en feittypen waren geplaatst. Tijdens deze fase werd van elk feittype nagegaan welke relatie er bestond tussen de deelnemende objecttypen en werden de constraints geplaatst. Deze fase vormde vaak de afsluiting van het modellerenproces.

De volgende foto is gemaakt tijdens een experiment met een ervaren deelnemer, zonder sturing. Tijdens dit experiment waren de verschillende momenten voor het plaatsen van de constraints duidelijk te onderscheiden.

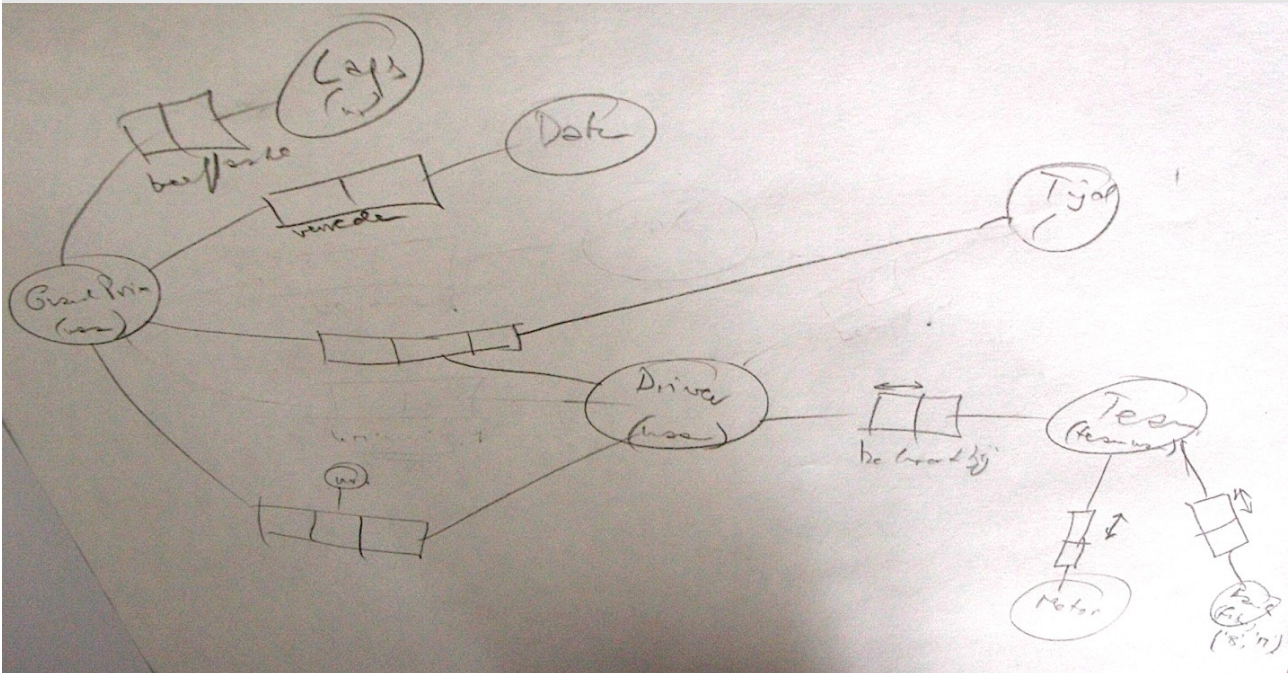


Foto 3

De bovenstaande foto is ongeveer halverwege het modellerenproces genomen. Behalve rechtsonder in de hoek staan verder nog nergens constraints. De deelnemer was in deze fase vooral bezig met het formuleren van objecttypen en feittypen. Een aantal constraints zijn in deze fase door de deelnemer wél geplaatst omdat hij aangaf deze belangrijk te vinden. In de case kwamen deze waarschijnlijk zó duidelijk naar voren dat hij deze meteen wilde plaatsen.

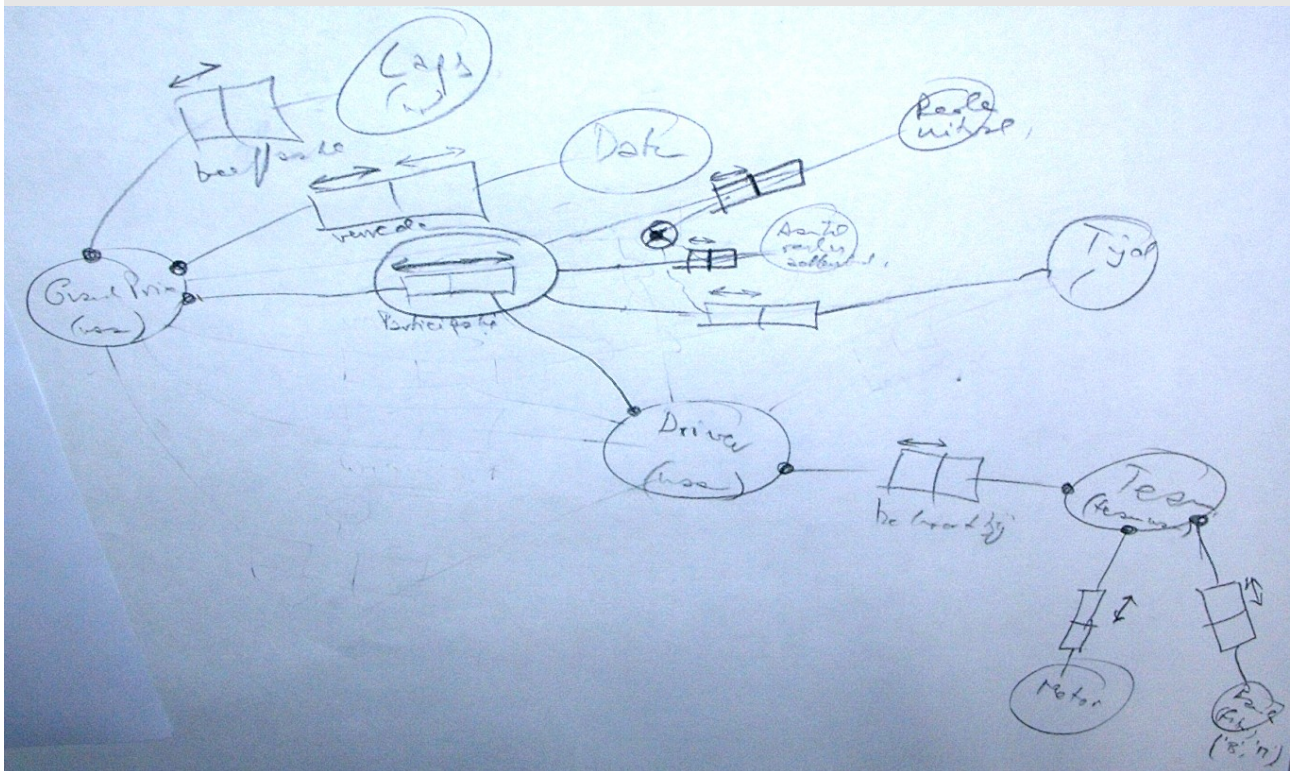


Foto 4

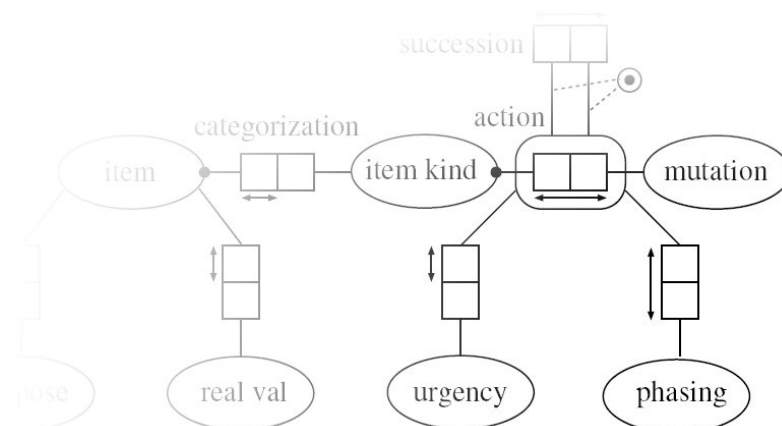
In de bovenstaande foto is het ORM diagram voltooid. In deze situatie kent elk feittypen minstens één constraint. Aan deze situatie is de laatste fase van dit modelleringsproces vooraf gegaan: het toevoegen van alle constraints. Het is opvallend om te zien dat er rechtsonder een aantal totale-rol constraints zijn toegevoegd. Ondanks dat de deelnemer in de fase van foto 3 al een aantal constraints had toegevoegd aan deze feittypen, heeft hij in de laatste fase van het modelleringsproces nog een aantal constraints toegevoegd aan deze zelfde feittypen. Blijkbaar komt de noodzaak van het plaatsen van de totale-rol constraints tijdens dit experiment op een ander moment dan de noodzaak van het plaatsen van de uniciteit constraints. Als gekeken wordt naar de context van de gebruikte case kan worden gesteld dat de deelnemer de constraint “elk team gebruikt precies één merk moter” eerder zag dan de constraint “elk team gebruikt een moter”. Op zich een enigszins vreemde volgorde.

De twee verschillende groepen deelnemers gingen anders om met constraints. Bij de onervaren groep was soms duidelijk te merken dat ze de betekenis van de verschillende constraints niet altijd goed begrepen. Ze konden (of wilde) de exacte betekenis van bepaalde constraints niet helemaal doorgronden en plaatste bepaalde constraints daardoor op gevoel. Sommige deelnemers plaatsten zo min mogelijk constraints omdat ze bang waren dat ze daarmee een correcte populatie van het ORM diagram uitsloten. Ook waren ze bang dat er populaties in de case voorkwamen die ze over het hoofd hadden gezien en die door het plaatsen van bepaalde constraints werden uitgesloten.

Anderen maakte zich daar minder druk om en plaatste juist veel constraints zonder precies na te gaan wat het effect hiervan was. Doordat de onervaren groep moeite had om de betekenis van bepaalde constraints te doorgronden, stelde ze ook aanzienlijk minder vragen over constraints aan de domeinexpert dan de ervaren groep.

De ervaren deelnemers konden veel sneller een vraag formuleren en deze voorleggen aan de domeinexpert als ze twijfelde over het toepassen van een bepaalde constraint. Doordat ze veel meer ervaring hadden met de betekenis van constraints, plaatsten ze deze niet op gevoel maar begrepen ze duidelijk de betekenis van de constraints.

Deze modelleerstijl heeft niet, zoals de “olievlek”, te maken met het exact uitvoeren van bepaalde stappen in een exacte volgorde. Er is ook geen harde grens te trekken tussen het type constraints die direct worden toegevoegd en type constraints die achteraf worden toegevoegd. Als strategiecatalogus uit [1] wordt gebruikt om deze stijl vast te leggen, wordt meteen duidelijk dat de stijl deze catalogus niet volledig kan populeren. Zo is bijvoorbeeld de volgorde (succession) bij deze stijl niet duidelijk. Een constraint kan op twee verschillende momenten in het modelleerproces worden geplaatst, maar dit zegt niets over de actie die daaraan vooraf ging of over de actie die daarop volgt. Ook geeft de stijl niet aan bij welk onderdeel van het ORM diagram nu precies de constraints worden geplaatst (item). In de onderstaande afbeelding van de strategiecatalogus zijn de irrelevante onderdelen voor deze stijl daarom minder duidelijk aangegeven.



Voor de twee verschillen momenten waarop er constraints worden toegevoegd, kan de catalogus ook op twee verschillende manieren worden gepopuleerd. Gedeelten van de populatie zullen hetzelfde blijven, maar anderen verschillen. Voor de populatie worden hieronder de voorbeeld populaties aangehouden zoals deze in [1] staan. De stijl omvat in elk geval het toevoegen van constraints dus deze populatie zal er als volgt uitzien:

mutation: add
item kind: Constraint

Het verschil tussen het direct plaatsen van constraints en het achteraf plaatsen van constraints heeft te maken met de urgentie (urgency) en fasering (phasing) van het modellerenproces. Deze populaties kunnen er als volgt uitzien:

Constraints direct plaatsen:

urgency: optional
phasing: middle

Constraints achteraf plaatsen:

urgency: required
phasing: last

Het verschil tussen de populaties van het objecttype “urgency” komt doordat het direct plaatsen van constraints niet verplicht is. Alle constraints aan het einde van het modellerenproces plaatsen is immers goed mogelijk. Om echter een volledig en correct model op te leveren zijn de constraints die aan het eind van het modellerenproces worden toegevoegd wel verplicht (required).

Bij het buiten beschouwing laten van opvolging (succession) bij deze stijl moet echter wel een kanttekening worden gemaakt. Het plaatsen van constraints aan het eind van het modellerenproces betekent dat er tijdens of na deze fase geen andere acties (dan het plaatsen van constraints) meer plaatsvinden. Dit houdt in dat het plaatsen van een constraint wordt gevolgd door het plaatsen van een constraint, net zolang totdat het ORM diagram af is. Hierbij is natuurlijk sprake van een heldere “opvolging relatie” tussen de acties. Deze populatie (van een stijl) kan echter niet in deze strategiecatalogus worden opgenomen.

6.1.3 Instantie- en typeniveau

Alle deelnemers aan het onderzoek begonnen tijdens het modelleren op typeniveau over het UoD te redeneren. Ook tijdens het gestuurd modelleren begonnen de meeste deelnemers feitzinnen op typeniveau te formuleren. Een belangrijke oorzaak hiervan zou de achtergrond van de deelnemers kunnen zijn; ze hebben vrijwel allemaal een ORM achtergrond. In modelleertechnieken als FCO-IM wordt sterk de nadruk gelegd op het beschouwen van het UoD op instantieniveau als het gaat om het formuleren van representatieve feitzinnen. Pas daarna worden uit deze zinnen de typen bepaald.

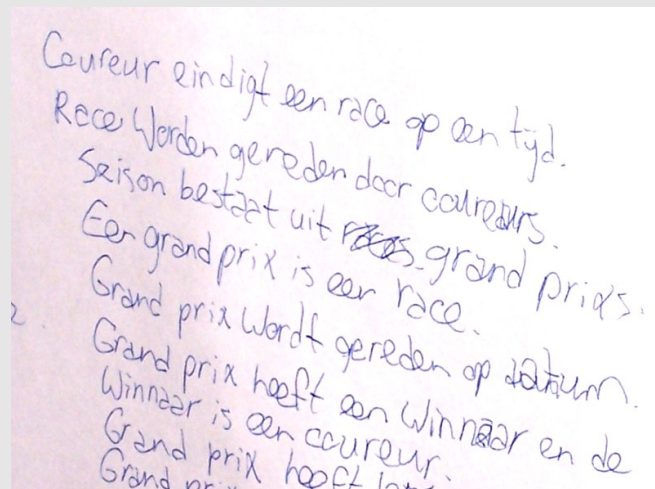
Voorbeeld:

Zin op instantieniveau:

Alonso rijdt bij Renault

Zin op typeniveau:

Coureur rijdt bij team



Coureur eindigt een race op een tijd.
Race worden gereden door coureurs.
Seizoen bestaat uit races.
Een grand prix is een race.
Grand prix wordt gereden op datum.
Grand prix heeft een winnaar en de winnaar is een coureur.
Grand prix heeft later

Foto 5

De bovenstaande foto is genomen tijdens het modelleerproces van een onervaren deelnemer die de formule 1 case met sturing oplost. Hier is duidelijk te zien dat de feitzinnen op typeniveau worden geformuleerd

Zelfs de deelnemers met een FCO-IM achtergrond, die tijdens de vrije case vanzelf feitzinnen begonnen te formuleren, deden dit echter op typeniveau.

Tijdens de gestuurde case werd aan alle deelnemers gevraagd om feitzinnen te formuleren. Hierbij werd niet aangegeven of deze op type- of op instantieniveau geformuleerd moesten worden. Zonder uitzondering werden alle zinnen geformuleerd op typeniveau zowel door de ervaren als door de onervaren groep. Doordat de objecttypen meteen duidelijk zijn op het moment dat een feitzin op typeniveau wordt geformuleerd, kregen alle zinnen ongeveer deze vorm:

<coureur> rijdt bij <team>

Een verklaring hiervoor is dat de cases dusdanig zijn gekozen dat de deelnemers zich eenvoudig een beeld van het UoD kunnen vormen. Bijna alle deelnemers waren behoorlijk bekend met voetbal en formule 1 en begrepen dus eenvoudig dat Alonso een coureur was. Deze extra stap hebben ze tijdens de gestuurde strategie niet expliciet uitgevoerd. Als het voor de deelnemers moeilijker was om een duidelijk beeld van het UoD te vormen, hadden ze deze stap wellicht wel expliciet uitgevoerd en waren ze wel begonnen met het formuleren van feitzinnen op instantieniveau. Zonder voorkennis van een domein is het lastiger om bijvoorbeeld objecten aan objecttypen te koppelen. Maar omdat dat niet onderzocht is, blijft het speculeren.

6.2 Veranderingen bij gestuurde strategie

Op het moment dat de deelnemers van het “vrij” modelleren omschakelde op het modelleren met behulp van de uitgereikte strategie traden er een aantal interessante fenomenen op. De twee verschillende soorten deelnemers (ervaren en onervaren) reageerden duidelijk anders als werd gevraagd om zich te houden aan een stappenplan, vandaar dat ik deze twee groepen hier ook apart wil behandelen.

6.2.1 Ervaren groep

De ervaren groep stond erg sceptisch tegenover een dergelijke strategie. Sommigen wilden het niet proberen en gaven aan dat ze door hun eigen, ingesleten werkwijze zich nooit aan een dergelijk stappenplan zouden kunnen houden.

Bij andere ervaren modelleers deed zich iets anders voor: Zij bleven in essentie aan hun eigen werkwijze vasthouden, maar deden het voorkomen alsof ze zich volledig aan de strategie overgaven. Het resultaat hiervan was dat ze hun eigen werkwijze niet op papier uitvoerden, maar helemaal in gedachte. Hetgeen ze vervolgens op papier zetten, voldeed volledige aan de strategie maar hier was wel een intensief denkproces aan vooraf gegaan.

Voorbeeld:

De feitzinnen werden geformuleerd nadat er uitvoerig over een mogelijk ORM diagram was geredeneerd. Die feitzinnen werden vervolgens opgeschreven en later gebruikt om objecttypen en objectlabels vast te stellen en om er uiteindelijk een ORM diagram van te kunnen maken. Een ORM diagram dat de deelnemer al vóór stap één van de strategie in gedachte had.

In de gevallen waar het bovenstaande voorbeeld op van toepassing is, wordt de strategie niet overgenomen door de betrokken deelnemer. De strategie die in dit onderzoek werd gebruikt, was bedoeld om volgens een gestructureerde manier richting te geven aan het modelleerproces. De strategie nam in feite een stuk van het modelleerproces over van de deelnemer waardoor deze zich beter op de inhoud van de case kon concentreren. Door de strategie niet over te nemen, maar als het ware “in te kapselen” in de eigen manier van modelleren, ontstond een nog veel complexer proces waarbij de nadruk minder lag op de inhoud van de case, en meer op het combineren van de eigen stijl met de strategie.

Als een modelleerstrategie is bedoeld om de modelleur te dienen (en dat is hij in mijn optiek), dan is de strategie die in dit onderzoek werd gebruikt daar absoluut niet in geslaagd (bij de ervaren groep deelnemers).

6.2.2 Onervaren groep

De onervaren groep deelnemers stond aanzienlijk minder sceptisch tegenover het modelleren met behulp van een strategie. Niemand gaf aan dat het gebruik van een strategie een onoverkomelijk probleem vormde bij het oplossen van de case en gaf de strategie in elk geval het voordeel van de twijfel. Het viel wel op dat het ook bij de onervaren groep een omschakeling vereiste voordat ze van vrij modelleren over konden stappen op gestuurd modelleren. Deze omschakeling bestond voornamelijk uit het verplaatsen van de aandacht van het toepassen van het meta model van ORM, naar het toepassen van de strategie. Veel onervaren deelnemers gebruikten de eerste (vrije) case om kennis van ORM weer op te rakelen. Tijdens de tweede (gestuurde) case was dat daardoor minder nodig en kon men zich meer op de strategie richten.

Anders dan bij de ervaren groep deelnemers, gaven alle onervaren deelnemers aan dat de strategie ze “houvast” bood bij het oplossen van de case. Vooral het verbaliseren van de cases hielp de deelnemers. Het formuleren van feitzinnen dwong de deelnemers om de case aandachtig te bestuderen. Deze “dwang” resulteerde erin dat de case beter en langer werd bestudeerd dan tijdens het modelleren zonder strategie. Het analyseren van de case gebeurde ook op een ander moment in het modelleerproces. Ondanks dat de deelnemers de case van tevoren bestudeerden, werd de case pas echt geanalyseerd *tijdens* het opstellen van het ORM diagram. Bij het gestuurd modelleren werd de case geanalyseerd tijdens het formuleren van feitzinnen. Omdat het formuleren van feitzinnen zelf minder aandacht kost dan het opstellen van feitzinnen (de Nederlandse taal is immers beter bekend dan het meta-model van ORM), leken de deelnemers meer aandacht te hebben voor de inhoud van de case. Bij de ervaren groep speelde dit geen rol omdat zij zich minder in hoefden te spannen om zich in ORM uit te drukken.

Ook konden de feitzinnen worden gebruikt om na te gaan of het ORM diagram compleet was zonder de hele case opnieuw te bestuderen. Feitzinnen konden eenvoudig worden afgevinkt als ze in het diagram voorkwamen. Dit kwam de compleetheid van het model ten goede en werd ook door meerdere onervaren deelnemers gebruikt.

Eén enkele deelnemer had totaal geen moeite om te schakelen van het vrije naar het gestuurde modelleren. Zijn eigen manier van modelleren kwam namelijk volledig overeen met de strategie. Dit kwam doordat hij behalve een ORM achtergrond, ook ervaring had met FCO-IM. Omdat de gebruikte strategie is gebaseerd op de voorgeschreven werkwijze van FCO-IM [5], is het niet verwonderlijk dat zijn modelleerproces aansloot bij de strategie.

Toch hadden deelnemers uit de onervaren groep soms ook moeite met het oplossen van de case met behulp van de strategie. Soms waren ze niet bekend met het jargon dat in de strategie werd gebruikt, maar wisten ze wel hoe het meta model van ORM ongeveer in elkaar zat. Zo waren de termen “objecttype” en “objectlabel” lang niet bij alle onervaren deelnemers bekend. Toch kwamen de problemen bij de onervaren groep meer voort uit onkunde dan uit onwil, wat bij de ervaren groep het geval was.

6.3 Mate van ervaring

Behalve de eerder genoemde verschillen tussen de ervaren en de onervaren groep, zijn er nog een aantal zaken aan het licht gekomen. Tijdens het gehele modelleerproces viel op dat de onervaren groep veel meer op gevoel modelleerden dan de ervaren groep. Waar ervaren deelnemers de betekenis van hun acties meteen doorzagen, waren de onervaren deelnemers hier vaak niet tot in staat of kostte het zeer veel tijd.

Voorbeeld

Tijdens één van de experimenten ontstond er een feittype met vijf rollen. De betrokken (onervaren) deelnemer wist dat feittypen met vijf rollen vrij zeldzaam waren in verband met splitsbaarheid, maar wist niet wáár het probleem precies zat. Vervolgens probeerde hij het feittype de splitsen, zonder dat hij wist of hij het echte probleem daarmee oploste.

Ervaren deelnemers konden sneller en beter bepalen wat de betekenis was van hun acties. Ze gingen niet op hun gevoel af maar wisten precies wat een verandering in het model betekende en welke effecten dit had op de semantiek van het model.

Ervaren deelnemers hadden ook veel minder tijd nodig (en beschikbaar) om tot een model te komen. Waar onervaren deelnemers soms een uur bezig waren om hun schema's te controleren, deden ervaren deelnemers dit in enkele minuten.

6.4 Voorkennis

Hoewel de cases redelijk eenvoudig en inzichtelijk waren, speelde voorkennis van het domein toch soms een rol. Een deelnemer die aangaf een grote fan te zijn van formule 1, modelleerde de formule 1 case totaal anders dan de voetbal case. Behalve dat één van de cases natuurlijk gestuurd werd uitgevoerd, ging hij ook anders om met de kennis van het domein en de interactie met de domeinexpert.

Veel voorkennis van het domein had tot resultaat dat de deelnemer ook erg op die voorkennis ging vertrouwen. De case werd vooral gebruikt om een beeld te vormen bij aanvang van het modellerenproces. Naarmate het modellerenproces vorderde, verdween de case naar de achtergrond en speelde de voorkennis een steeds grotere rol. Dit had tot resultaat dat er zaken werden gemodelleerd die niet in de case voorkwamen en waar niet over gecommuniceerd werd met de domeinexpert. Veel voorkennis leidde in dit geval tot “roekeloos” modelleren en tot weinig verdieping in de case.

Deelnemers met weinig voorkennis verdiepte zich vaak meer in de case en stelden vooral meer vragen aan de domeinexpert. Doordat ze onzeker waren over hun kennis van het domein, wilde ze zeker weten dat de acties die ze uitvoerde, overeen kwamen met de kennis van de domeinexpert.

6.5 Afleidbare informatie

Onder afleidbare informatie wordt informatie verstaan die *niet* in het model kan worden vastgelegd, maar die wel afgeleid kan worden van informatie die *wel* in het model kan worden vastgelegd.

Voorbeeld

- Van de finishtijden van de coureurs kan worden afgeleid op welke plek ze eindigden.
- Van de uitslag van een voetbalwedstrijd kan worden afgeleid welk team won.

Bijna alle deelnemers houden rekening met afleidbare informatie, en modelleren geen informatie uit het UoD die kan worden afgeleid uit informatie die al aanwezig is in het model. Ze gaan er blijkbaar vanuit dat die informatie later gegenereerd kan worden uit de informatie die wel in het model word meegenomen. Het nadenken over afleidbare informatie geeft aan dat de proefpersonen verder dachten dan alleen het model. Ze dachten na over complexe constraints en rules van het model. Onderdelen die niet in het model kunnen worden vastgelegd met behulp van ORM.

7 Evaluatie onderzoek

In dit hoofdstuk ga ik in op de laatste subvraag (subvraag 4) van mijn onderzoek:

“Hoe kan een theorie die betrekking heeft op modelleerstijlen empirisch getest worden?”

Omdat deze vraag redelijk los staat van de inhoud van mijn onderzoek wil ik hier een apart hoofdstuk aan wijden. De overige subvragen worden beantwoord in het laatste hoofdstuk.

7.1 Methode

De manier waarop dit onderzoek is uitgevoerd, is in mijn ogen een geschikte methode om een dergelijk kwalitatief en exploratief onderzoek uit te voeren. Het observeren en ondervragen van deelnemers die een case modelleren, is een effectieve manier om te bepalen hoe modelleerprocessen in elkaar zitten. De methode is echter niet geschikt om “harde” uitspraken te doen over effectiviteit of correctheid. Daarvoor zouden de cases, de gebruikte strategieën en de selectie van de deelnemers verder moeten worden aangescherpt. Juist het “open” karakter van het onderzoek zorgde ervoor dat de deelnemers konden modelleren zoals ze gewend zijn en dat er fenomenen naar voren kwamen die waarschijnlijk ook in de praktijk voorkomen.

De technieken om de experimenten vast te leggen, waren achteraf gezien niet allemaal even effectief en nuttig. Hieronder geef ik een korte motivatie en een ranglijst van de technieken (1 is nuttigst, 4 is het minst nuttig)

1. Aantekeningen De aantekeningen die ik tijdens de experimenten heb gemaakt, hebben een belangrijke rol gespeeld bij het analyseren van de modelleerprocessen achteraf. In de aantekeningen heb ik alles vastgelegd wat ik tijdens de experimenten tegenkwam. Veel van deze gegevens bleek achteraf irrelevant te zijn en werden later niet meegenomen in de analyses. Of een bepaald fenomeen relevant is, is tijdens het experiment zelf vaak lastig vast te stellen. Daarom is het belangrijk om zoveel mogelijk aantekeningen te maken. Na de experimenten heb ik de aantekeningen uitgewerkt zodat er een duidelijk beeld ontstond van het betreffende experiment. Tijdens het uitwerken van de aantekeningen worden de relevante onderdelen overzichtelijk weergegeven zodat deze eenvoudiger te analyseren zijn. Bij de analyses heb ik deze uitwerkingen vervolgens met elkaar vergeleken. De rol van de aantekeningen in mijn onderzoek is zó groot dat een dergelijk onderzoek in mijn ogen onmogelijk is *zonder* aantekeningen en nog erg relevante informatie kan opleveren met *alleen* aantekeningen. In bijlage 4 zijn de uitgewerkte aantekeningen te zien van een experiment.

2. Opgeleverde modellen Hoewel de opgeleverde modellen vaak slechts een momentopname vormen van het einde van het modellerproces, gaven ze een duidelijk beeld over hoe een deelnemer gewend is te modelleren. Vooral de vergelijking tussen verschillende modellen leverde soms interessante informatie op. Zo gaven de modellen bijvoorbeeld aan hoe belangrijk de deelnemers constraints vonden. In sommige opgeleverde modellen kwamen constraints nauwelijks voor terwijl in andere modellen op veel meer plaatsen constraints staan.
3. Foto's Voor aanvang van het eerste experiment verwachtte ik dat foto's mij een belangrijke houvast zouden geven tijdens het analyseren van de modellerprocessen. Met behulp van een aantal momentopnamen van het modellerproces, zou het mogelijk zijn om dit proces achteraf te analyseren. Tijdens de analyse achteraf viel dit echter nogal tegen. De rol die foto's spelen bij het vastleggen van het proces, wordt eigenlijk helemaal overgenomen door het maken van aantekeningen. Het maken van aantekeningen heeft dan nog als voordeel dat er bepaalde nuances in vastgelegd kunnen worden. De foto's zijn eigenlijk alleen van pas gekomen ter illustratie in deze scriptie. Daarnaast moeten er behoorlijk veel foto's worden gemaakt om elke stap van het proces duidelijk in beeld te brengen. Het maken van veel foto's kan bovendien het proces zelf verstoren.
4. Audio opnamen De audio opnamen hebben geen enkele rol gespeeld tijdens de analyse van de modellerprocessen. Zaken die via het "think aloud protocol" aan het licht kwamen, werden veelal met behulp van aantekeningen vastgelegd. Het probleem van audio opnamen was dat de context van bepaalde uitspraken bijna niet te herleiden was. Antwoorden op vragen als "Waar was de deelnemer mee bezig toen hij deze uitspraak deed?" zijn moeilijk na te gaan aan de hand van alleen geluid.

Hoewel in mijn ogen aantekeningen absoluut noodzakelijk zijn voor het vastleggen van dergelijke experimenten, valt het te overwegen om bij een volgend onderzoek video opnamen te maken in plaats van audio opnamen en foto's. Met behulp van video opnamen is het mogelijk om een modellerproces echt helemaal vast te leggen. Bovendien kan een videocamera gewoon aanstaan, zonder dat de onderzoeker of de deelnemer daar verder last van heeft. Daarnaast hebben video opnamen als voordeel dat het beeld en geluid gelijk loopt waardoor er een beter beeld ontstaat van het modellerproces.

7.2 Eigen rol

In de loop van het onderzoek zat er een duidelijke ontwikkeling in mijn eigen rol. Vóór de eerste proef had ik geen duidelijk beeld van wat ik tijdens een proef kon verwachten en wat mijn eigen rol hierin zou zijn. Naarmate ik meer experimenten had afgenomen, werd ik meer bedreven in het ontdekken van verbanden en het opmerken van bepaalde fenomenen. Hierdoor kon ik ook betere feedback geven aan de proefpersonen en gerichtere vragen stellen. Je zou kunnen stellen dat de latere experimenten daardoor meer informatie opleverde dan de eerder afgenomen experimenten.

Ik heb echter ook gemerkt deze toenemende ervaring een obstakel kan vormen. Naarmate ik meer experimenten had afgenomen, ontstond er de neiging om sommige verbanden “te snel” te leggen. Ik dacht dan dat een actie van een deelnemer leek op iets dat ik al eerder was tegengekomen. Na nadere bestudering bleek dat soms niet zo te zijn. Het werkt erg prettig als je door je ervaring sneller verbanden kan leggen en meer informatie uit een proef kan halen, maar je moet wel scherp en kritisch blijven.

Een manier om meer informatie uit een proef te halen is doormiddel van het gericht vragen stellen aan de deelnemer. Het is vaak interessant om door te vragen naar -bijvoorbeeld- de achtergrond of motivatie van de deelnemer. Als degene die de experimenten afneemt alleen maar “aanschouwt” hoe de cases worden opgelost gaat er veel informatie verloren. De cases bieden juist een kans om vragen te stellen over de acties die een deelnemer tijdens de experimenten onderneemt. Het formuleren van de juiste vragen ging, naarmate ik meer experimenten had afgenomen, steeds beter. Ik heb voor het stellen van deze vragen nooit van tevoren een lijst opgesteld.

Een groot nadeel van het uitvoeren van het onderzoek met slechts één onderzoeker is de grote mate van subjectiviteit. Bij een exploratief onderzoek als dit, vormen de indrukken, ervaringen en observaties die de onderzoeker opdoet tijdens de experimenten de kern van het onderzoek. Alle experimenten en analyses zijn tot stand gekomen vanuit mijn eigen oogpunt en worden verwerkt met behulp van mijn ervaring, vooroordelen en expertise. Op die manier ontstaat er een beeld van de experimenten die erg wordt beïnvloed door één persoon.

7.3 Voorleggen van strategieën

Ik heb gemerkt dat het erg lastig is om een geschikte strategie voor te leggen aan de deelnemers tijdens een onderzoek als dit. Omdat het onderzoek er niet op was gericht om één bepaalde strategie te toetsen, maar zich richtte op het “sturen van een modelleerproces” in het algemeen, was het lastig om een geschikte strategie te formuleren die bruikbaar was tijdens de experimenten.

De strategie die tijdens de pilot proef was gebruikt, bleek achteraf veel te ingewikkeld en star te zijn. Tijdens de pilot was het nauwelijks mogelijk om de strategie te volgen omdat hij te weinig ruimte liet aan de deelnemer. Dit had tot gevolg dat het praktisch onmogelijk was om met “behulp” van de strategie tot een goed conceptueel schema te komen.

De uitkomsten van de pilot zorgden enigszins voor een dilemma: een simpelere strategie zorgde ervoor dat het modelleerproces vlotter verliep, maar de nadruk lag ook minder op het gestuurde modelleren. Een strategie die de deelnemers strikt moesten volgen en die weinig ruimte liet voor hun eigen expertise, zou onherroepelijk tot een moeizaam modelleerproces leiden. Omdat ik wilde weten hoe de deelnemers zouden reageren op een gestuurde strategie was het belangrijk dat de gebruikte strategie niet de strikt was, met als gevolg dat het modelleerproces tot geen enkel resultaat zou leiden. Een te “vrije” strategie zou zeker tot een conceptueel schema leiden, maar hierbij zou niet duidelijk worden welke fenomenen er optreden bij het gebruik van sturing tijdens een modelleerproces omdat de sturing eenvoudigweg te weinig aanwezig zou zijn.

Bij de strategie die uiteindelijk is gebruikt, lag de nadruk op overzichtelijkheid. De deelnemers waren in staat de strategie in relatief weinig tijd te doorgronden doordat hij slechts uit zes regels bestond. Het sturende karakter bleef behouden doordat de strategie was gebaseerd op een modelleerproces die weinig deelnemers tijdens de vrije strategie volgden. Met behulp van een gesprek met één van de grondleggers van FCO-IM, Guido Bakema, heb ik de strategie verbeterd. Dit heeft geleid tot de strategie die ik uiteindelijk heb gebruikt.

7.4 Think aloud protocol

Tijdens het afnemen van de experimenten is het mij opgevallen hoe goed de deelnemers zich houden aan het “think aloud protocol” (hardop denken). Aan het begin van de proef wordt aangegeven dat de deelnemers hardop moeten denken en dit houden ze meestal tot het einde vol.

Bij het hard op denken kan het bijvoorbeeld gaan over:

- De motivatie van de deelnemer om een bepaalde actie uit te voeren
- Twijfels
- De werkwijze die de deelnemer normaal volgt
- Acties waar de deelnemer extra lang over na moet denken en waarom

Hoewel de deelnemers constant vertellen waar ze mee bezig zijn, vertellen ze vaak over oppervlakkige zaken. Ze vertellen erg snel een eenvoudig *wat* ze doen, maar minder eenvoudig *waarom*. Om ook daar achter te komen is het belangrijk om vervolgvragen te stellen aan de hand van de informatie die ze vanzelf geven. Ook worden de deelnemers stiller op het moment dat ze ergens diep over na moeten denken. Vaak zijn de motivaties en gedachtegangen juist op die momenten erg interessant, en is het belangrijk dat ze aangeven waar ze over nadenken of twijfelen. Op die moment is het belangrijk dat de onderzoeker doorvraagt of de deelnemers eraan herinnert dat ze hard op moeten denken.

7.5 Populatie

Doordat alle deelnemers studeerden of werkten aan de Radboud Universiteit Nijmegen, ging ik er bij de selectie van de populatie vanuit dat iedereen ongeveer dezelfde achtergrond had. Voor de ervaren deelnemers klopte dit, maar de achtergrond van de onervaren groep verschilde sterk. Een aantal onervaren deelnemers was vanuit vorige (HBO-) opleidingen bijvoorbeeld beter bekend met UML dan met ORM. Ook kwam het voor dat onervaren deelnemers ORM niet goed genoeg beheerste om vlot te modelleren. De onervaren groep was wat betreft modelleerkwaliteiten te divers.

Achteraf gezien had ik meer tijd moeten besteden aan het selecteren van de deelnemers. Een goede selectie zou meer deelnemers met dezelfde beheersing van ORM opleveren waardoor de experimenten beter vergelijkbaar zouden worden en soepeler zouden verlopen.

7.6 Vervolgonderzoek

Een exploratief onderzoek als dit kan prima op deze manier worden uitgevoerd en zou in de toekomst eventueel groter van opzet kunnen zijn. Hierbij zou gedacht kunnen worden aan:

- Meer deelnemers. Door de experimenten bij meer deelnemers af te nemen ontstaat een beter beeld van de verschillende modelleerstijlen en hoe deze voorkomen.
- Meer onderzoekers. Door het onderzoek bijvoorbeeld met twee mensen uit te voeren, wordt een gedeelte van de subjectiviteit van één onderzoeker verkleind. Ook zijn twee onderzoekers in staat om een grotere populatie deelnemers te onderzoeken.

Dit onderzoek is voortgekomen uit de onderzoeksgroep IRIS. Naarmate de theorie over modelleerstrategieën zich verder ontwikkeld, zou er ook behoefte kunnen ontstaan om opnieuw empirische gegevens te vergaren over dit onderwerp. De doelstellingen van een dergelijk onderzoek zullen waarschijnlijk verschillen van de doelstelling van dit onderzoek. Het is bijvoorbeeld mogelijk om de effectiviteit van één bepaalde strategie te toetsen bij één specifieke doelgroep. Een andere mogelijkheid is om dergelijke experimenten te herhalen in een bedrijfssituatie zoals in [9] is gedaan.

Dit onderzoek zou eventueel op een grotere schaal kunnen worden herhaald om meer gegevens te verkrijgen, maar verder zie ik geen redenen om meer onderzoeken uit te voeren met een exploratief karakter als deze. Nadat duidelijk is welke zaken een rol spelen tijdens de modelleerprocessen bij verschillende deelnemers en doelgroepen, wordt het interessant om meer strikte onderzoeksvragen te formuleren. De nadruk ligt dan niet op het observeren en analyseren van situaties en acties maar bijvoorbeeld op het meten van effectiviteit. Op die manier kan de theorie omtrent modelleerstrategieën echt aangescherpt worden.

Zoals al eerder is aangegeven, is het aan te bevelen om een dergelijk onderzoek in de toekomst met twee onderzoekers uit te voeren. Daarnaast zou ik voor het vastleggen van de modellerprocessen uitgaan van video opnamen en uitgebreide aantekeningen door beide onderzoekers. Op die manier gaat er zo min mogelijk informatie over het modellerproces verloren. Afhankelijk van het onderzoeksdoel zouden er een aantal vragen kunnen worden geformuleerd die tijdens het experiment aan de deelnemer kunnen worden gesteld.

8 Conclusie

In dit hoofdstuk worden de antwoorden gegeven op de eerder gestelde vragen. Als eerste worden de subvragen beantwoord omdat deze leiden naar het antwoord op de hoofdvraag. Hoewel er in de eerdere hoofdstukken al een aantal voorzetten zijn gegeven met betrekking tot de beantwoording van de vragen, zullen de resultaten en de onderzoeksvragen in dit hoofdstuk pas echt aan elkaar worden gekoppeld en leiden tot een antwoord. Zoals al eerder aangegeven is, zal subvraag 4 niet in dit hoofdstuk worden behandeld (Zie hoofdstuk 7).

8.1 Subvraag 1

Welke modelleerstijlen zijn er te onderscheiden tijdens de experimenten?

Zoals al eerder is aangegeven is een modelleerstijl een groep modelleerprocessen met eigen kenmerken. Tijdens mijn experimenten heb ik hiervan drie duidelijke voorbeelden kunnen vinden:

- De “olievlek”: Een stijl die zich kenmerkt doordat de modelleur begint met één centraal object en het model vervolgens uitbreidt met telkens een object en een feittype. Deze stijl is in 6.1.1 uitgebreid omschreven.
- Het plaatsen van constraints zoals in 6.1.2 is beschreven.: Deze stijl kenmerkt zich doordat de deelnemers constraints aanbrengen op twee specifieke momenten: tijdens het plaatsen van de bij de constraint behorende objecttype of feittype, of achteraf als het model alle objecttypen en feittypen bevat.
- Modelleren op typeniveau.: Bijna alle deelnemers beschouwden het UoD op typeniveau en formuleerde de feitzinnen daardoor ook op typeniveau. Deze stijl is in 6.1.3 uitgebreid omschreven.

8.2 Subvraag 2

In welke mate heeft de ervaring van de deelnemer invloed op het modelleerproces?

Het belangrijkste verschil tussen ervaren en onervaren deelnemers is de snelheid en de kwaliteit waarmee gemodelleerd wordt. Als ervaren deelnemers tijdens de experimenten twijfelden, had deze twijfel vaak te maken met de case of het UoD. Bij onervaren deelnemers had deze twijfel vaak betrekking op ORM, het meta-model van ORM of de manier van modelleren zelf. Ervaren deelnemers konden zich daardoor beter verdiepen in de case omdat ze minder tijd nodig hadden om over het modelleren zelf na te denken. Dit betekende dat ervaren deelnemers bepaalde verbanden sneller en eerder zagen dan onervaren deelnemers.

8.3 Subvraag 3

Wat gebeurt er als er van vrij modelleren wordt overgeschakeld op een strategie?

Doordat de twee verschillende groepen deelnemers ook ieder verschillend omgingen met de strategie, wordt de beantwoording van deze vraag ook toespitsen in de twee groepen.

Ervaren deelnemers

De groep van de ervaren deelnemers had duidelijk het meeste moeite met de overschakeling. Zij kwamen in veel gevallen niet meer aan modelleren toe of gaven aan dat ze niet volgens een strategie wilden werken. Voorzover deze aanpak bij deze groep nog leidde tot modellen, waren deze vaak erg beknopt en kwamen ze niet zo makkelijk tot stand als tijdens het vrije modelleren.

De ervaren deelnemers die de strategie wel probeerden te volgen, bleven in essentie vaak toch aan hun eigen manier van modelleren vasthouden. Op papier volgden ze de strategie maar bij navraag bleek dat ze in gedachten hun eigen werkwijze bleven volgen. Het resultaat hiervan was dat het heel veel moeite kostte om deze manier van modelleren vol te houden.

Onervaren deelnemers

Doordat deze groep minder “ingesleten” gedrag had op het gebied van modelleren, was de overstap naar modelleren met behulp van een strategie een stuk minder groot. In een aantal gevallen bood de strategie zelfs houvast bij het modelleren. Het probleem bij deze groep was dat ze niet altijd bekend waren met de termen die in de strategie werden gebruikt.

Behalve de verschillen tussen de twee verschillende groepen deelnemers, reageerden beide groepen soms ook op dezelfde manier. Geen van beide groepen volgde de strategie bijvoorbeeld blindelings. Alle deelnemers keken “voorbij” de strategie en dachten na over het uiteindelijke model. Er was niemand die zich volledig overgaf aan de strategie.

8.4 Hoofdvraag

De hoofdvraag zoals deze eerder is geformuleerd luidt:

Hoe verhoudt de theorie over modelleerstrategieën, zoals gedefinieerd binnen de onderzoeksgroep IRIS, zich met de modelleerstijlen die in de praktijk worden toegepast?

Om deze vraag te kunnen beantwoorden, is het niet alleen belangrijk om de modelleerstijlen te bepalen zoals deze in de praktijk voorkomen (subvraag 1) maar ook om helder te krijgen welk gedeelte van de theorie binnen de onderzoeksgroep IRIS hier duidelijk mee te maken heeft. Niet alle aspecten van de theorie zijn relevant voor de beantwoording van deze vraag; de aspecten die betrekking hebben op het sturen van een modelleerproces uiteraard wel. Hierbij valt te denken aan aspecten die in [1] worden behandeld. Om de verhouding tussen theorie en praktijk weer te geven, wil ik uitgaan van de praktijk. Met andere woorden: de keuze voor relevante aspecten uit de theorie hangt af van de modelleerstijlen die in de praktijk zijn waargenomen.

De theorie die tot op dit moment binnen IRIS is ontwikkeld, doet nog geen harde uitspraken over hoe een modelleerproces het beste kan worden ingericht. De theorie gaat in op de manieren waarop modelleerstrategieën in de toekomst kunnen worden gedefinieerd en welke attributen hierbij een rol spelen. De modelleeragenda in combinatie met de strategiecatalogus zijn hier een goed voorbeeld van. Mijn onderzoek levert empirische informatie die gebruikt kan worden bij de ontwikkeling van concrete modelleerstrategieën.

De modelleerstijlen zoals ik deze tijdens mijn experimenten ben tegengekomen, zouden in de toekomst kunnen worden gebruikt om een strategiecatalogus te populieren. Eén van de eigenschappen van een bruikbare strategie is dat hij aansluit op het “natuurlijke” modelleerproces van de modelleur. De modelleerstijlen die in mijn onderzoek naar voren zijn gekomen zijn hier voorbeelden van.

De modelleeragenda moet in de toekomst op een dynamische en voorschrijvende manier richting geven aan een modelleerproces. Hoewel ik de principes achter de modelleeragenda niet aan de deelnemers van mijn experimenten heb voorgelegd, is uit dit onderzoek wel gebleken hoe een modelleerstrategie een soepel modelleerproces in de weg kan staan. Behalve als input voor de populatie van een strategiecatalogus, kan dit onderzoek daarom ook worden gebruikt om te bepalen hoe bijvoorbeeld een modelleeragenda zich moet “gedragen” bij het sturen van een modelleerproces.

Voor de ervaren deelnemers houdt dit in dat de modelleeragenda in elk geval ruimte moet laten voor hun eigen expertise en vakmanschap. Het is belangrijk dat er een goede balans gevonden moet worden tussen deze vrijheid en de sturing van het modelleerproces. Gebeurt dit niet dan zullen ervaren modelleurs de grootste moeite hebben om met dergelijke sturing om te gaan. Voor onervaren modelleurs moet de modelleeragenda daarentegen genoeg houvast bieden zodat deze groep de juiste keuzes voorgelegd kan worden. Om deze twee groepen zo goed mogelijk te ondersteunen, is het belangrijk dat een modelleeragenda flexibel genoeg is om hiermee om te gaan.

9 Referenties

1. P.van Bommel, S.J.B.A.Hoppenbrouwers, A.(Erik)Proper and Th.P.van der Weide - "Exploring modelling strategies in a meta-modelling context"
2. A.(Erik)Proper - "Information Intensive Organizations"
3. Terry Halpin – "Object role modeling (ORM/NIAM)"
4. Dinesh Batra, Nicole A. Wishart - "Comparing a rule-based approach with a pattern-based approach at different levels of complexity of conceptual data modelling tasks"
5. http://en.wikipedia.org/wiki/Think_aloud_protocol
6. Guido Bakema, Jan Pieter Zwart, Harm van der Lek - "Volledig Communicatiegeörienteerde Informatiemodellering"
7. H.A. Proper, A.A. Verrijn-Stuart and S.J.B.A.Hoppenbrouwers - "On Utility-based Selection of Architecture-Modelling Concepts"
8. S.J.B.A.Hoppenbrouwers, A.(Erik)Proper and Th.P.van der Weide - "A Fundamental View on the Process of Conceptual Modeling"
9. Jinwoo Kim, Jungpil Hahn, Hyoungmee Hahn - "How do we understand a system with (so) many diagrams? Cognitive integration processes in diagrammatic reasoning"
10. Gesprek met Guido Bakema op 12 april 2006 in het HAN kenniscentrum.

10 Woordenlijst

Actie: Kleinste onderdeel van een modelleerproces. Bijvoorbeeld:

- Objecttype toevoegen
- Constraints aanbrengen

Controlled language: Een vastgelegde, helder omschreven en versimpelde versie van een natuurlijke taal (bijvoorbeeld het Nederlands). Een voorbeeld hiervan zijn elementaire voorbeeldzinnen.

Dialogue document: Document waarin met behulp van een controlled language een modelleerproces wordt vastgelegd.

Gestuurd modelleren: Modelleren volgens een vooropgesteld plan of strategie. De deelnemer moet zich hieraan houden.

IRIS: Information Retrieval and Information Systems

Item: onderdeel van het meta model van ORM. Bijvoorbeeld een objecttype. In de strategiecatalogus wordt hiervoor de term “item kind” gebruikt.

ORM: Object Role Modelling. De modelleer methodologie die tijdens de experimenten is gebruikt.

UoD: Universe of discourse. Het gedeelte van de werkelijkheid die in beschouwing wordt genomen tijdens het modelleren.

Vrij modelleren: Modelleren zonder dat de deelnemer gebruik hoefde te maken van een strategie. De deelnemer is helemaal vrij om zijn eigen aanpak en werkwijze te volgen.

11 Bijlagen

In de scriptie wordt naar bijlagen 1, 2, 3 en 4 verwezen. Bijlage 5 geeft een modelleerproces weer zoals deze door een ervaren deelnemer, zonder sturen werd uitgevoerd. Foto's van dit modelleerproces zijn in eerder gebruikt in paragraaf 6.1.2. Deze bijlage is bijgevoegd om een idee te krijgen hoe een dergelijk modelleerproces eruit ziet, uitgedrukt in tien foto's. In de scriptie wordt verder niet verwezen naar deze bijlage.

11.1 Bijlage 1: Formule 1 case

Opdracht:

Bestudeer de volgende voorbeelden van het domein en ontwerp een ORM model

Voorbeeld 1:

2005 Season Review					
Grand Prix	Date	Winning Driver	Team	Laps	Time
Australian	06/03/2005	Giancarlo Fisichella	Renault	57	1:24:17.336
Malaysian	20/03/2005	Fernando Alonso	Renault	56	1:31:33.736
Bahrain	03/04/2005	Fernando Alonso	Renault	57	1:29:18.531
San Marino	24/04/2005	Fernando Alonso	Renault	62	1:27:41.921
Spanish	08/05/2005	Kimi Räikkönen	McLaren-Mercedes	66	1:27:16.830
Monaco	22/05/2005	Kimi Räikkönen	McLaren-Mercedes	78	1:45:15.556
European	29/05/2005	Fernando Alonso	Renault	59	1:31:46.648
Canadian	12/06/2005	Kimi Räikkönen	McLaren-Mercedes	70	1:32:09.290
United States	19/06/2005	Michael Schumacher	Ferrari	73	1:29:43.181
French	03/07/2005	Fernando Alonso	Renault	70	1:31:22.233
British	10/07/2005	Juan Pablo Montoya	McLaren-Mercedes	60	1:24:29.588
German	24/07/2005	Fernando Alonso	Renault	67	1:26:28.599
Hungarian	31/07/2005	Kimi Räikkönen	McLaren-Mercedes	70	1:37:25.552
Turkish	21/08/2005	Kimi Räikkönen	McLaren-Mercedes	58	1:24:34.454
Italian	04/09/2005	Juan Pablo Montoya	McLaren-Mercedes	53	1:14:28.659
Belgian	11/09/2005	Kimi Räikkönen	McLaren-Mercedes	44	1:30:01.295
Brazilian	25/09/2005	Juan Pablo Montoya	McLaren-Mercedes	71	1:29:20.574
Japanese	09/10/2005	Kimi Räikkönen	McLaren-Mercedes	53	1:29:02.212
China	16/10/2005	Fernando Alonso	Renault	56	1:39:53.618

Voorbeeld 2:

Zondag 16 oktober: Uitslag race

nr.	coureur	team	motor	bd	tijd
01.	Alonso	Renault	Renault	M	1:39:53,618
02.	Raikkonen	McLaren	Mercedes	M	+ 4,015
03.	R.Schumacher	Toyota	Toyota	M	+ 25,376
04.	Fisichella	Renault	Renault	M	+ 26,114
05.	Klien	Red Bull	Cosworth	M	+ 31,839
06.	Massa	Sauber	Petronas	M	+ 36,400
07.	Webber	Williams	BMW	M	+ 36,842
08.	Button	BAR	Honda	M	+ 41,249
09.	Coulthard	Red Bull	Cosworth	M	+ 44,247
10.	Villeneuve	Sauber	Petronas	M	+ 59,977
11.	Monteiro	Jordan	Toyota	B	+ 1:24,648
12.	Barrichello	Ferrari	Ferrari	B	+ 1:32,812
13u	Pizzonia	Williams	BMW	M	+ 1 ronde
14u	Doornbos	Minardi	Cosworth	B	+ 1 ronde
15.	Trulli	Toyota	Toyota	M	+ 1 ronde
16u	Albers	Minardi	Cosworth	B	+ 5 ronden
u.	Sato	BAR	Honda	M	34 versnelling
u.	Karthikeyan	Jordan	Toyota	B	28 crash
u.	Montoya	McLaren	Mercedes	M	24 radiator
u.	M.Schumacher	Ferrari	Ferrari	B	22 spin

Bd: Banden

B:Bridgestone

M: Michelin

11.2 Bijlage 2: Voetbal case

Opdracht:

Bestudeer de volgende voorbeelden van het domein en ontwerp een ORM model

13 februari 2006

Programma			
Datum	Tijd	Wedstrijd	Accommodatie
17-02-2006	2030	NEC - FC Twente	DE GOFFERT
17-02-2006	2030	Vitesse - PSV	GELREDOME
18-02-2006	1930	Roda JC - SC Heerenveen	PARKSTAD LIMBURG STADION
18-02-2006	1930	FC Groningen - RKC Waalwijk	EUROBORG STADION
19-02-2006	1230	Feyenoord - Sparta	DE KUIP
19-02-2006	1230	AZ - FC Utrecht	ALKMAARDER HOUT
19-02-2006	1430	Ajax - RBC Roosendaal	AMSTERDAM ARENA
19-02-2006	1430	Willem II - NAC Breda	WILLEM II STADION
19-02-2006	1430	Heracles Almelo - ADO Den Haag	POLMAN STADION

Uitslagen			
Datum	Thuis	Uit	Uitslag
12-02-2006	ADO Den Haag	Ajax	1 - 2
12-02-2006	FC Utrecht	NEC	1 - 1
12-02-2006	RBC Roosendaal	FC Groningen	1 - 2
12-02-2006	NAC Breda	Feyenoord	3 - 3
11-02-2006	SC Heerenveen	Vitesse	4 - 1
11-02-2006	PSV	Heracles Almelo	1 - 0
11-02-2006	FC Twente	Roda JC	1 - 0
11-02-2006	RKC Waalwijk	Willem II	1 - 0
10-02-2006	Sparta	AZ	0 - 1

11.3 Bijlage 3: Voorgelegde strategie

- 1. Formuleer feitzinnen**
- 2. Maak elementaire zinnen**
- 3. Bepaal feittypen en voor elk feittype:**
 - 1. Bepaal objecttypen**
 - 2. Bepaal objectlabel**
- 4. Breng beperkingsregels (constraints) aan.**

11.4 Bijlage 4: Aantekeningen

Vrij: voetbal case

Bekend met domein 8

5. Bestudeert de case
6. Tekent heel snel een diagram en neemt een aantal zaken aan die niet uit de case te halen zijn (uit/ thuis staat bijvoorbeeld niet bij programma)
7. Redeneert dan verder over de diagram icm de case (zit dit wel goed, zit dat wel goed... etc)
8. Past nog wat kleine dingen aan maar laat alle objecttypen en feittypen op hun plek.
9. Constraints aanbrengen gaat vanuit voorkennis, niet vanuit de case ("Ik neem aan dat een wedstrijd één datum heeft")

Modelleert verder niet erg uitgebreid. Geen labels, weinig constraints.

Gestuurd: F1 case

Bekend met domein: 4

Gaf meteen al aan dat hij gestuurd modelleren niets vond.

Begon toch met het opschrijven van een zin, maar week meteen van de strategie af. Kon er niets mee.

Problemen:

F1 case vindt hij complexer

F1 case vindt hij onbekender

De strategie vindt hij vervelend. Niet zozeer deze strategie, maar meer strategieën in het algemeen. Hij gelooft niet erg in natuurlijke taal analyse omdat dat "in de weg staat". Hij heeft er vroeger wel "les" in gehad, maar dat is weggezaakt omdat hij het zonder beter kan. Vertrouwt meer op z'n eigen jbf aanpak.

Als het met (feit) zinnen moet werken zou hij van het domein eerst een ORM diagram maken en dan vervolgens daar zinnen uit proberen te halen of het diagram toetsen aan het domein. Eerst zinnen maken zou hij niet snel doen.

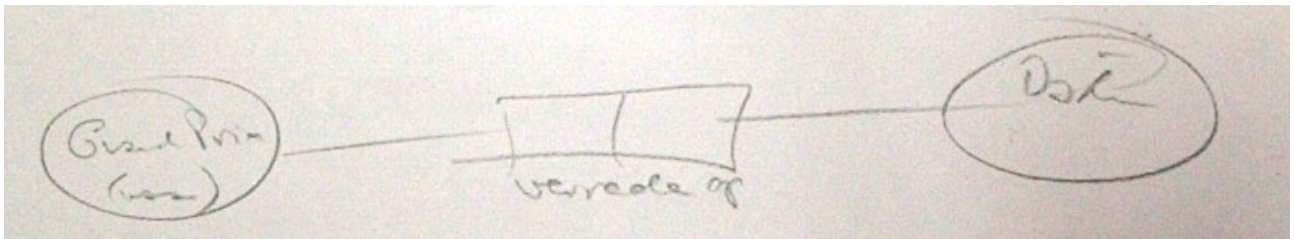
Hij is niet erg bekend met de termen feitzinnen en elementaire zinnen. Maar toen hij zelf aangaf wat ze ongeveer moesten betekenen had hij het wel goed. Houd zich er verder helemaal niet mee bezig.

Algemeen

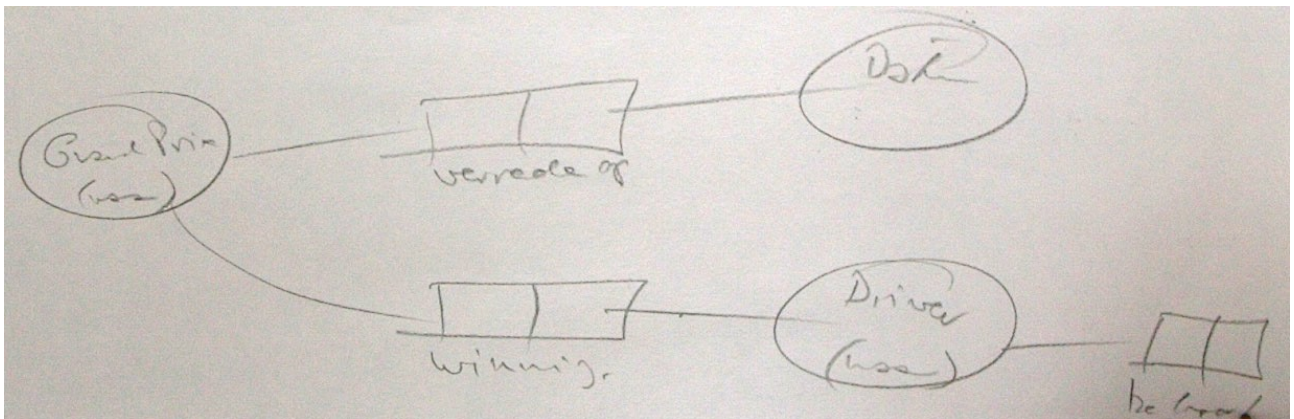
Beschrijft eigen manier van modelleren als een olievlek: Eerst twee objecten en dan telkens uitbreiden met een (vaak binair) feit. Hij zou nooit aan twee kanten van het domein beginnen, hij begint ergens, en breidt dat uit.

Totaal 30 minuten bezig geweest.

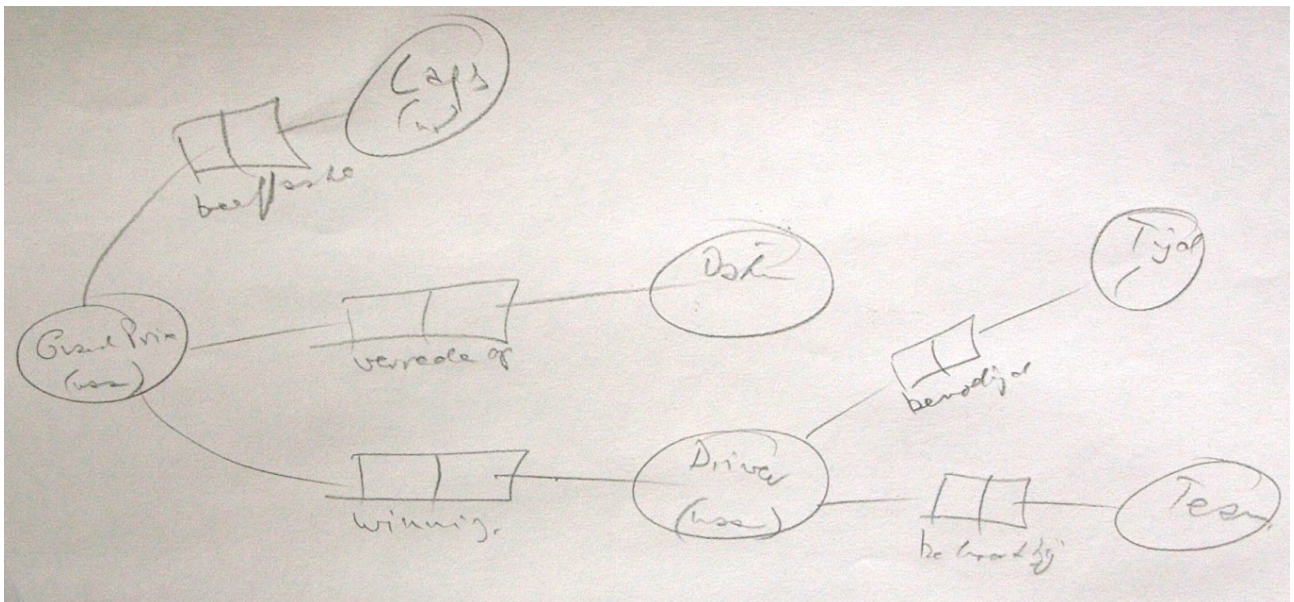
11.5 Bijlage 5: Foto's van een modelleerproces



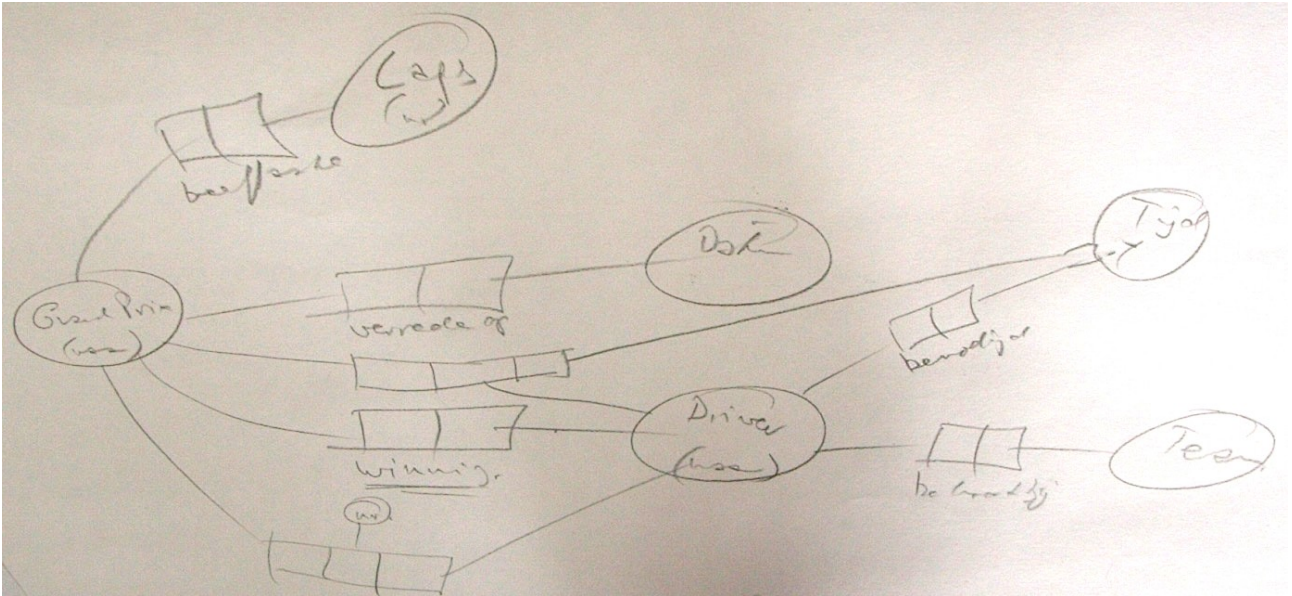
1



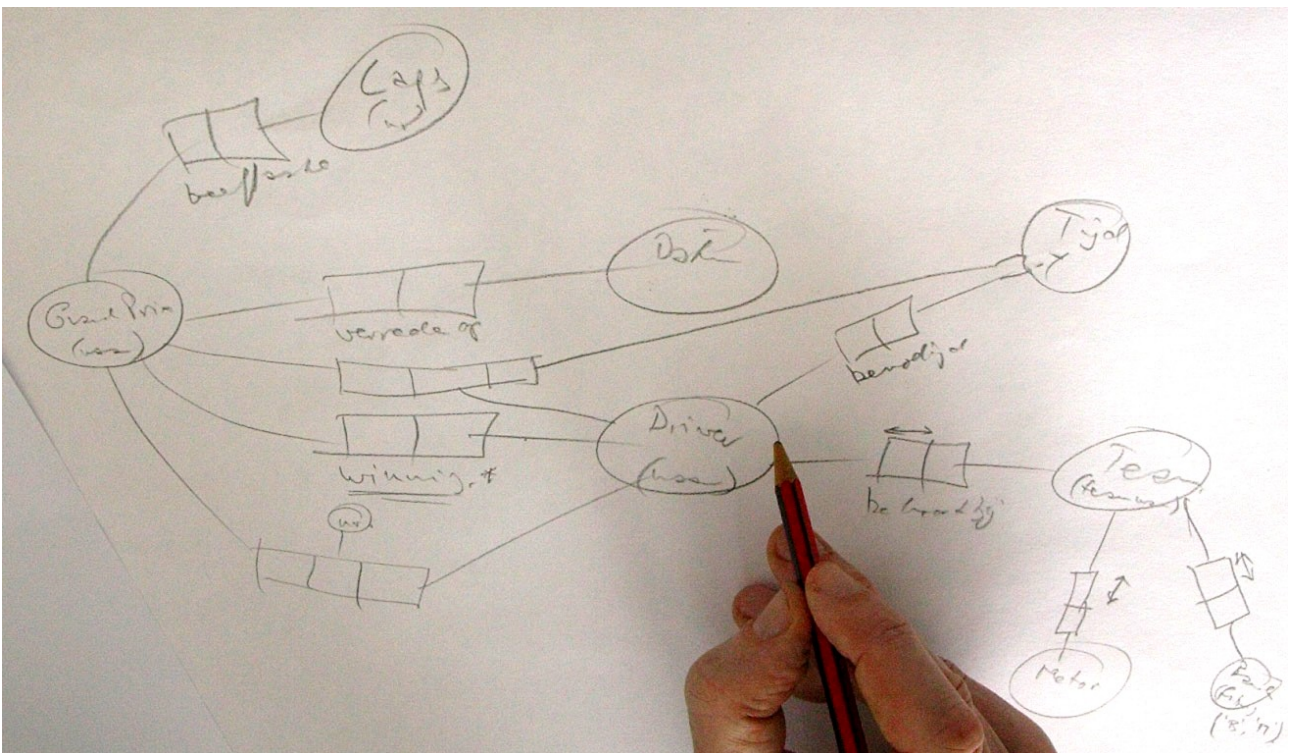
2



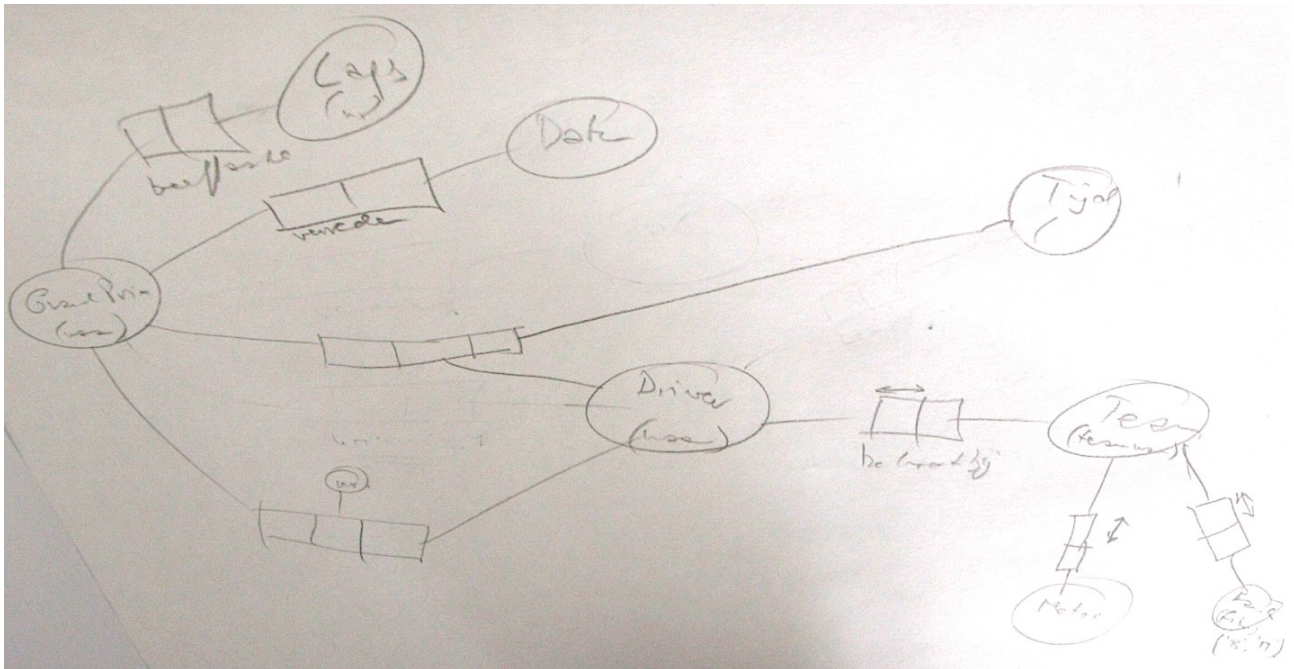
3



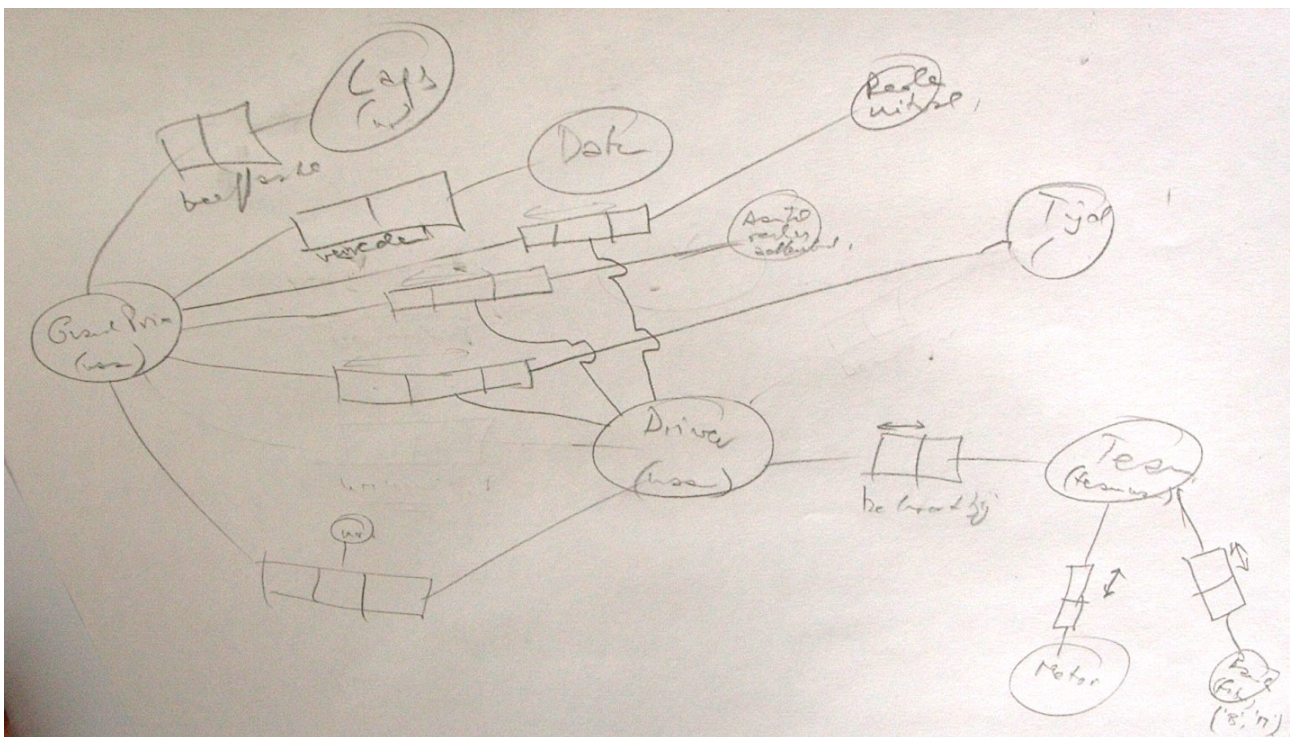
4



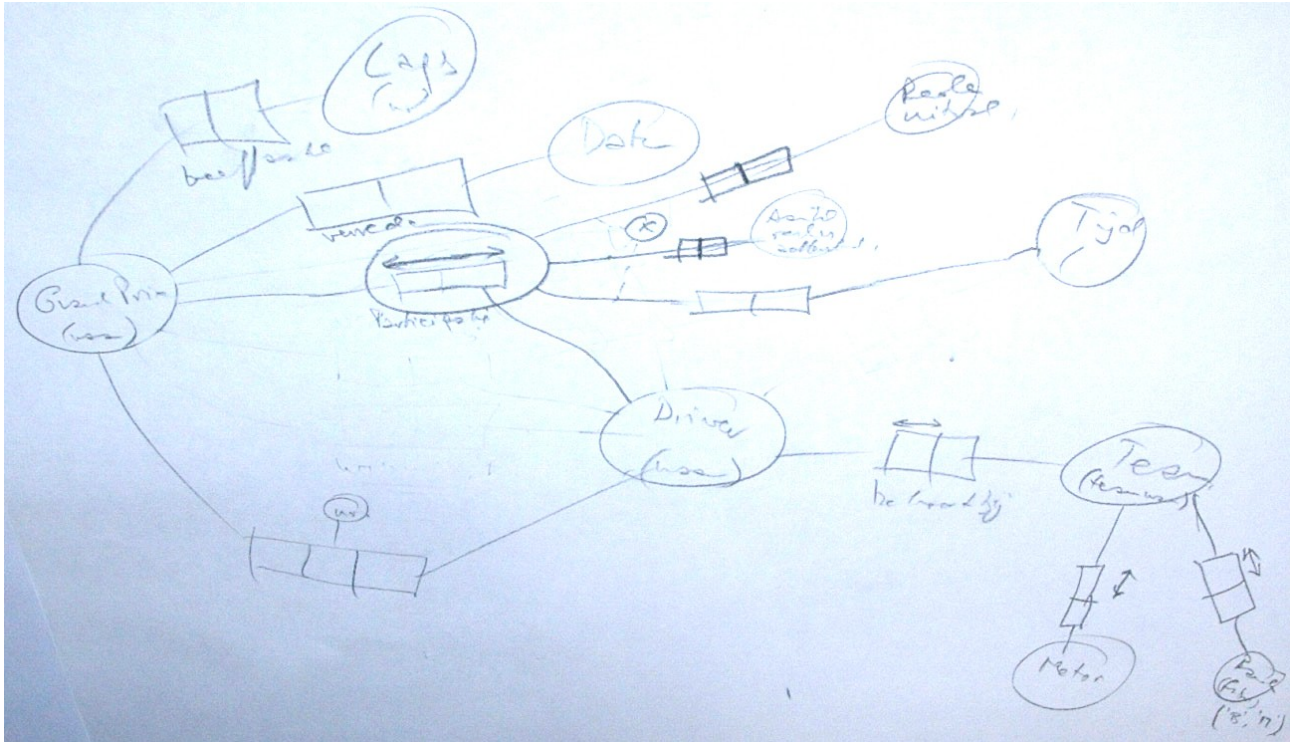
5



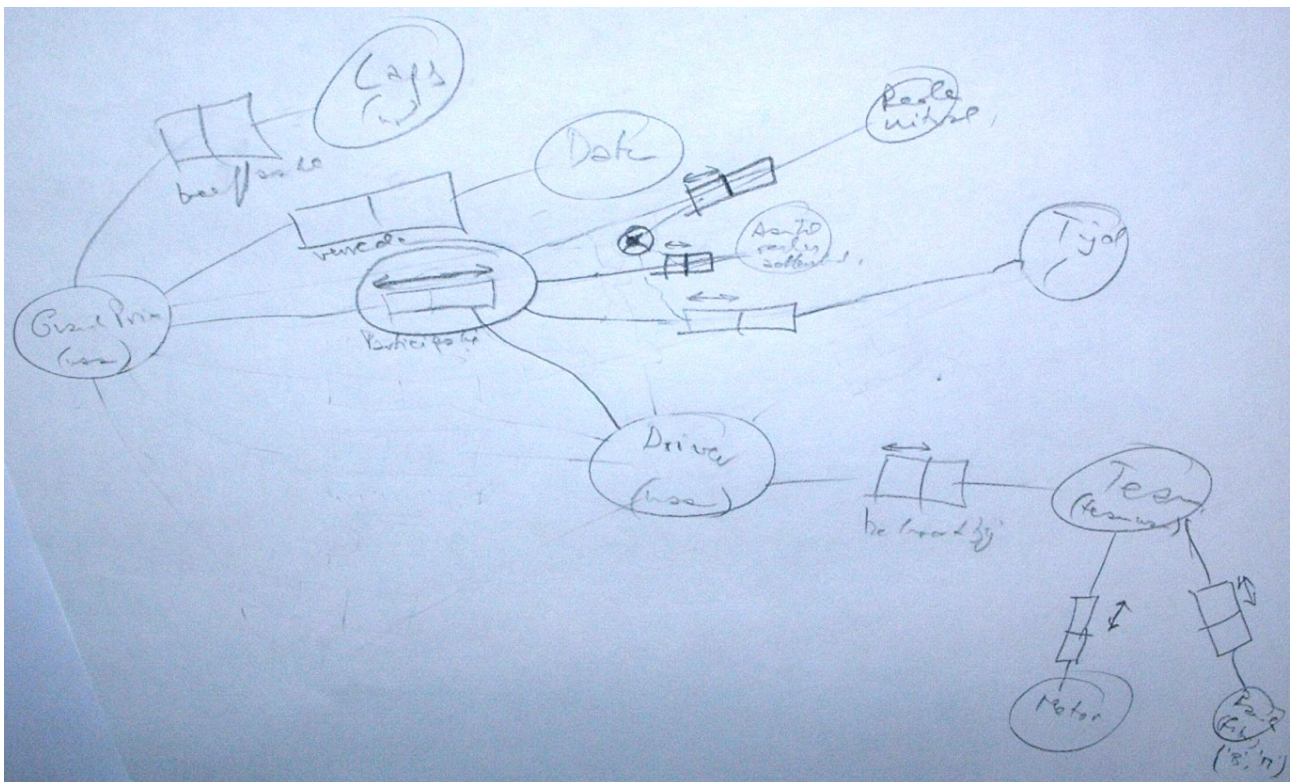
6



7



8



9

