

# Communicerende Agents en Avatars in Virtuele Werelden

Anton Nijholt & Hendri Hondorp<sup>1</sup>

Universiteit Twente, Postbus 217

7500 AE Enschede, Nederland

anijholt@cs.utwente.nl

**Abstract:** *We rapporteren over voortgaand onderzoek in een virtual reality omgeving waar bezoekers kunnen interacteren met agents die helpen bij het verkrijgen van informatie, bij het uitvoeren van transacties en bij het samenwerken om bepaalde taken te verrichten. In een multi-user versie van het systeem kunnen bezoekers bovendien met elkaar 'chatten'. Het onderzoek vindt plaats binnen het Parlevinkproject van het cluster TKI (Taal, Kennis & Interactie) van de Faculteit Informatica van de Universiteit Twente. De omgeving kan beschouwd worden als een laboratorium voor onderzoek en voor het doen van experimenten met gebruikers die op multimodale wijze met agents kunnen communiceren en daarbij kunnen verwijzen naar gevisualiseerde informatie en gebruik maken van kennis aanwezig bij domeinagents of bij andere bezoekers die wellicht ook als agent gerepresenteerd zijn in de omgeving. We bespreken standaarden die ten behoeve van het ontwerpen van dergelijke omgevingen ontwikkeld worden. Onze omgeving modelleert het Muziekcentrum van Enschede. De bedoeling is de omgeving te laten evolueren tot een theatergemeenschap waar we niet alleen doelgerichte bezoekers hebben die volstaan met het kopen van kaartjes, maar ook bezoekers die nog niet zeker weten of ze wel iets willen kopen, die wellicht alleen maar informatie willen of slechts rond willen kijken, willen praten met andere bezoekers, etc. We laten zien dat we een multi-user en multi-agent omgeving nodig hebben om onze doelstellingen te realiseren en dat we een unificerend raamwerk nodig hebben om in staat te zijn agents en gebruikersavatars met verschillende vaardigheden te introduceren en te beheren, inclusief intellectuele, interactie en animatie vaardigheden.*

## 1 Inleiding

We bespreken een virtual reality theateromgeving met ingebedde agents die de bezoeker kunnen helpen via, o.a., natuurlijke taal dialogen. De omgeving is gebouwd gebruikmakend van VRML (Virtual Reality Modeling Language) en ze is toegankelijk via WWW. Oorspronkelijk was de omgeving gebouwd rond een bestaand natuurlijke taal dialoogsysteem waarbij de dialogen enkel betrekking hadden op voorstellingen en reserveringen in het theater [6]. In de nieuwe omgeving is dit dialoogsysteem toegekend aan een gevisu-



Fig. 1 Luchtfoto van het Centrum van Enschede

aliseerde, menselijk uitziende agent aan wie gebruikers vragen kunnen stellen. Zodra we deze agent hadden en de omgeving gingen uitbreiden groeide de behoefte om andere agents toe te voegen, agents die in staat waren de bezoeker te helpen, die in staat waren met elkaar te communiceren en die in staat waren autonoom gedrag te vertonen. We bespreken hoe onze ideeën over deze omgeving in de tijd veranderden, in het bijzonder door meer aandacht te besteden aan potentiële bezoekerswensen. In plaats van een gebruikelijk doelgericht informatie- en transactiesysteem, vergelijkbaar met een telefonisch informatiesysteem, zien we dat de omgeving verandert in een virtuele gemeenschap waar verschillen tussen bezoekers en artificiële agents vaag worden en waar de onderzoeksonderwerpen variëren van het toekennen van persoonlijkheden en emoties aan artificiële agents, studies van gebruikersaspecten bij navigatie-agents tot formele specificaties van (interacties) in virtuele omgevingen en 'reinforcement' leren voor agents in deze virtuele, multimodale omgeving met als doel de autonomie van een agent te vergroten.

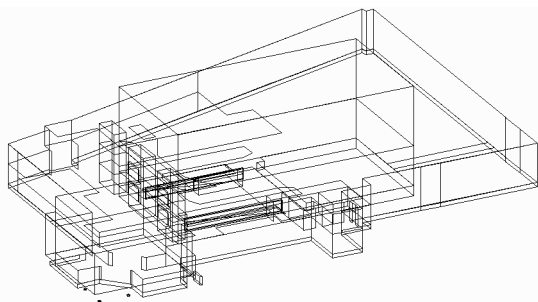
## 2 Bouwen van de Virtuele Omgeving

In Fig. 1 zien we een luchtfoto van het centrum van Enschede. Zichtbaar is het marktplein, de oude kerk (let op de donkere schaduw in het midden van de foto) en enkele theaters. Het voornaamste theatergebouw, het 'Muziekcentrum' is rechts te vinden. Het heeft enkele

<sup>1</sup> Het onderzoek waarover gerapporteerd wordt in dit artikel werd mogelijk gemaakt door de stichting "VR Valley Twente" te Enschede en door het U-Wish project van het *Telematica Instituut*, eveneens te Enschede. Een engelstalige versie van dit artikel is gepresenteerd op het Eurographics 2000 congres te Interlaken.

theaterzalen, kamers voor artiesten, recreatieruimten (zowel voor artiesten als bezoekend publiek), garderobes, etc. Deel van het gebouw bestaat uit een Conservatorium. Er zijn andere theaters in de stad. Informatie over voorstellingen kan verkregen worden uit brochures, advertenties, telefonisch (inclusief het bekende ‘toets 3 als u meer informatie wilt over ....’) of door simpel de fiets te nemen, er naar toe te gaan en een informatrice te vragen naar voorstellingen en vervolgens een reservering te maken voor een gekozen voorstelling.

Op dit moment zijn enkele gebouwen, inclusief het Muziekcentrum, straatjes (Bolwerkstraat) en de Oude Kerk gemodelleerd. In Fig. 2 is een draadmodel van het MuziekCentrum te zien. Het is gebouwd in overeenstemming met de constructietekeningen van de architecten van het gebouw. Kamers, trappen, podia, etc., ook volgens het werkelijke gebouw, werden toegevoegd en objecten, wanden, plafonds e.d. werden voorzien van ‘textures’ verkregen uit foto’s en videos van het gebouw om de virtuele omgeving er meer



**Fig. 2** Draadmodel van het MuziekCentrum

realistisch uit te laten zien. Een honderd procent realisme werd niet nagestreefd, al was het maar vanwege tijdsbeperkingen en de beperkte beschikbaarheid van studenten en programmeurs die ingeschakeld kunnen worden om aan de omgeving te werken. Er is echter ook niet altijd een noodzaak om volledig realisme na te streven. In een virtuele omgeving kunnen we toelaten dat bezoekers bepaalde taken uitvoeren zonder dat ze daarbij belemmerd hoeven te worden door fysieke of sociale beperkingen die in de echte wereld gelden. Het weghalen van alle beperkingen kan echter ook tot verwarring en niet-acceptatie leiden.

Fig. 3 geeft een ‘screenshot’ weer van het virtuele Muziekcentrum waarbij we door de deur van buiten naar binnen kijken. Binnen zien we een plattegrond van de zaal, enkele posters en een informatiebalie met daarachter een wachtende agent, Karin genoemd, die klaar staat om ons te vertellen over voorstellingen, artiesten en beschikbare kaartjes.

Bezoekers kunnen deze virtuele wereld exploreren, van de ene lokatie naar de andere lopen, vagen stellen aan zichtbare agents, objecten activeren door ze aan te klikken, etc. Karin, de informatrice van het theater, heeft een 3-D gezicht met enkele eenvoudige gezichtsuitdrukkingen en lipbewegingen die synchroon lopen met een tekst-naar-spraak systeem dat de uitingen van het systeem verwoordt. Vanwege beperkingen ons opgelegd door huidige webtechnologie is er nog niet echt sprake van verfijnde synchronisatie tussen de (inhoud van de) uitingen die worden geproduceerd door de dialoogmanager en corresponderende lipbewegingen en gezichtsuitdrukkingen van de Karin agent. Ontwerpoverwegingen die kunnen leiden tot het weergeven van verbaal en niet-verbaal agent-gedrag voor Karin zijn te vinden in [7].

Multimodaliteit is een ander onderwerp. Karin besluit informatie in een tabel op het scherm weer te geven als er teveel voorstellingen zijn om achter elkaar voor te lezen. Het mag duidelijk zijn dat als er teveel voorstellingen zijn die voldoen aan de wensen van de bezoeker we moeilijk kunnen verwachten dat na het geven van informatie over de derde of de vierde voorstelling de bezoeker nog weet heeft van de details van de eerste voorgelezen voorstelling. Vandaar dat we besloten hebben Karin en haar informatiebalie in te bedden in een windows omgeving waar we ook informatie kunnen presenteren in een aparte window met aanklikbare items en pop-up menu’s met veelgestelde vragen. Het dialoogsysteem laat toe bij de ingetikte vragen te verwijzen naar de items in de tabel. Dat wil zeggen, in plaats van een van de veelgestelde vragen aan te klikken kan men ook vragen stellen zoals:



**Fig. 3** Gezicht naar Binnen vanaf de Buitendeur

“Geeft u me alstublieft meer informatie over de derde voorstelling”, waarbij verwezen wordt naar het derde item in de table van voor de bezoeker relevante voorstellingen.

In dezelfde omgeving, noodgedwongen soms in een ander prototype van het virtuele Muziekcentrum, zijn andere agents geïntroduceerd. Zo is er bijvoorbeeld een navigatie-agent die afweert van de indeling van het gebouw en die zowel met spraak als in natuurlijke taal via het toetsenbord aangesproken kan worden. Van een echte dialoog is hierbij (nog) geen sprake. De bezoeker

kan vragen stellen over lokaties in het theater. Als de vraag wordt begrepen dan wordt een route berekend en het gezichtspunt van de bezoeker wordt als het ware geleid langs deze route naar de bestemming. De huidige navigatie-agent is niet gevisualiseerd als een avatar. Het gezichtspunt van de agent komt overeen met het gezichtspunt dat hoort bij de coördinaten van de bezoeker in het theater. Het zou vreemd zijn als de bezoeker zich zelf zou kunnen zien.

Een Java-gebaseerd agent-raamwerk is ontworpen om een protocol te definiëren voor de communicatie tussen agents. Het raamwerk laat toe andere, soortgelijke agents te introduceren. Bijvoorbeeld, waarom zou een theaterbezoeker niet tegen de zaalplattegrond mogen praten of tegen een poster met de aankondiging van een interessante voorstelling?

De versie van het virtuele theater met een spraakherkenningscomponent is niet toegankelijk gemaakt voor het grote publiek door ze op het web te zetten. Hoewel spraakherkenning op de server gebeurt (waarmee allerlei problemen rond downloaden, eigendom e.d. worden vermeden) blijven er desalniettemin te veel problemen aanwezig wat betreft herkeningsaccuraatheid en synchronisatie met de gebeurtenissen in het systeem. Het werk aan de navigatie-agent wordt echter voortgezet. Onderdelen van dit verdere werk zijn het modelleren en leren van bezoekersvoorkeuren voor navigatie in virtuele werelden, het modelleren van navigatiekennis en navigatiedialogen, het toekennen van instructiemodellen aan een navigatie-agent en natuurlijk een meer verfijnde visualisatie.

### 3 Bezoekers in Interactie met Avatars

In onze omgeving kunnen we verschillende mens-achtige agents introduceren. Sommigen worden gerepresenteerd als communicatieve 'humanoides', min of meer natuurlijk uitziende menselijke figuurtjes die zich bevinden en rondlopen in de virtuele wereld en interacties toestaan met bezoekers van de omgeving. In een browser die visualisatie van meerdere gebruikers toestaat kunnen ze elk zichtbaar gemaakt worden als (3D) avatars. Wij willen dat elke bezoeker kan communiceren met agents en met andere bezoekers, al of niet gevisualiseerd, in zijn of haar (zichtbare) omgeving. Dat betekent dat we conversaties kunnen hebben tussen agents, tussen bezoekers, en tussen bezoekers en agents. Daarbij kunnen meer dan twee bezoekers en/of agents betrokken zijn. Zoiets goed te laten verlopen is een tamelijk ambitieus doel waar de nodige tijd mee gemoeid zal zijn.

Een ander probleem dat we moeten noemen is dat communicatie gesitueerd in een zichtbare of op andere wijze waarneembare (virtuele) gedeelde omgeving toelaat dat de dialoogpartners hun communicatieve acties ondersteunen door andere manieren van aanduiden (bijvoorbeeld door te kijken of te wijzen) dan linguïstische verwijzingen. Introductie van dergelijke multimodale ondersteuning voor talige communicatie kan agents en bezoekers helpen elkaar te begrijpen, maar er worden ook nieuwe en uitdagende problemen geïntroduceerd. Een van die problemen is dat van co-referenties naar voor beide partners zichtbare objecten. De frase "die deur" is verbonden aan een zichtbaar object in de omgeving en veronderstelt dat de partners beiden dit object kunnen zien. De 'geometrische' virtuele omgeving (beschreven in VRML code of welke virtuele modelleertaal dan ook) moet eveneens beschreven worden op een abstract conceptueel niveau als op een linguïstisch niveau. Een agent moet op een of andere wijze weten naar welk object een bezoeker wijst (waarbij we wellicht ook rekening moeten houden met het feit of het object in het gezichtsvlak van de gevisualiseerde agent staat) en de agent moet dus ook in staat zijn deze wijze van aanduiden te matchen met de linguïstische verwijzing ("die deur").

In de voorgaande secties spraken we over agents die 'acteerden' in ons eigen virtuele theater. Karin werd geïntroduceerd als een 'visualisatie' van ons al bestaand dialoogsysteem. Ze heeft uitgebreide kennis over voorstellingen die opgevoerd worden in het theater. Ze



Fig. 4 Jacob 'praat' tegen Karin

heeft een repertoire van lipbewegingen (in functie van de dialoog) en enkele simpele hoofdbewegingen. Nadat Karin geïntroduceerd was bleek het wenselijk een agent platform te introduceren en we introduceerden een navigatie-agent met enige geografische kennis en vaardigheden op het terrein van spraaksynthese en spraakherkenning. De omgeving biedt ons de mogelijkheid tal van potentieel nuttige agents te introduceren, waarbij sommige van die agents slechts simpele, zich voortdurend herhalende animaties vertonen, anderen kunnen de muisbeweging van de bezoeker volgen (bijvoorbeeld een representatie van de gebruiker) of een berekende route lopen (nuttig voor bijvoorbeeld de navigatie-agent), terwijl anderen wat

meer verfijnde intelligentie hebben die toelaat zekere acties uit te voeren getriggerd door interacties met bezoekers.

#### 4. Multi-agents en Multi-users

We hebben onze omgeving ingebed in een multi-user 'shell' (DeepMatrix [8]), hetgeen inhoudt dat bezoekers voor andere bezoekers zichtbaar worden als avatars (VRML objects) waaraan animaties kunnen worden gekoppeld, maar ook intelligentie en interactie-vaardigheden die een afspiegeling zijn van die van de bezoeker, maar niet noodzakelijkerwijs daarmee overeenkomen omdat we afhankelijk van onze doelen (en toepassing) kunnen besluiten ze te modificeren (te versterken, af te zwakken of daartoe gedwongen te worden door de techniek, ze minder natuurlijk te laten verlopen). Om een voorbeeld te noemen, we kunnen gebruik maken van bezoekersprofielen (uit ervaring verkregen, door te veronderstellen of door vragen te stellen) die toegekend zijn aan een bezoeker's avatar die rondloopt in de virtuele omgeving. In een e-commerce omgeving, bijvoorbeeld, kunnen we verschillend koopgedrag gewaar worden en erop anticiperen (de omgeving aan te passen aan de voorkeuren van de gebruiker, in te zoomen op bepaalde producten, etc.), door rekening te houden met het profiel van de bezoeker [5].

In Fig. 4 zien we een bezoekers avatar in de buurt van Karin komen. De bezoeker heeft een keuze gemaakt uit een van onze geanimeerde avatars (in dit geval Jacob). Jacobs animaties laten hem wandelen langs de coördinaten van het bewegende gezichtspunt van de bezoeker die hij representeert. Jacob werd geïntroduceerd in een van onze andere projecten (zie [3]) waar het voorzien was van een instructie- en domeinmodel om aan studenten een bepaalde taak te onderwijzen (hoe de *Torens van Hanoi* op te lossen). Dat wil zeggen, Jacob heeft kennis van het Torens van Hanoi probleem, weet studenten te helpen dit probleem op te lossen en weet dit te doen op een interactieve manier. In deze theateromgeving is Jacob slechts een voorbeeld van hoe we VRML objecten en hun animaties kunnen hergebruiken om agents en bezoekers in andere omgevingen vorm te geven wanneer we ons houden aan standaarden. Jacobs intelligentie kan niet zonder meer toegepast worden in een andere VRML omgeving. Er dienen standaarden ontwikkeld te worden die betrekking hebben op intelligentie-aspecten van agents naast standaarden die betrekking hebben op afmeting, fysiek voorkomen en animaties in VRML werelden.

In onze omgeving bestaan meer mogelijkheden voor een gebruikers avatar om artificiële agents tegen te komen. Een pianist met enkele zich herhalende voorge-



Fig. 5 Jacob Maakt Ruzie met de Barok Danseres

definieerde animaties die suggereren dat de piano wordt gespeeld, is zichtbaar op het toneel. Aan de Università degli Studi di Milano in Italië wordt onderzoek verricht aan barokdansanimatie met virtuele dansers [1]. Met behulp van een editor kunnen dansen uit te voeren door virtuele dansers choreografeerd en gegenereerd worden. De animaties zijn verkregen door middel van motion capturing van de bewegingen van echte dansers. Aangezien de gegenereerde dansen en animaties beschreven worden in VRML wordt het ons mogelijk gastvoorstellingen van deze danseressen van de Scala van Milaan aan te bieden. Jammer genoeg is het huidige repertoire tamelijk beperkt en zouden we liever dansen automatisch vanuit muziek willen kunnen genereren. In Fig. 5 zien we dat een bezoeker gerepresenteerd door Jacob zo brutaal is geweest het podium op te klimmen en daar onze danseres lastig valt.

Het zal duidelijk zijn dat om een virtuele omgeving met een veeltal van domein- en gebruikersagents te beheren en te onderhouden enige uniformiteit bij het introduceren van agents noodzakelijk is. Binnen die uniformiteit kan divergentie plaats vinden in verschillende richtingen en combinaties van richtingen: agent intelligentie, interactievaardigheden van agents, visualisatie van agents en animatie van agents.

We kunnen een blik werpen op een aantal VRML gerelateerde standaarden die zijn voorgesteld of die in ontwikkeling zijn. Voor onze doeleinden zijn we geïnteresseerd in:

- De 'Humanoid Animation' (H-Anim) standaard [10]. Deze standaard definieert een (fysieke) structuur en een interface voor agents in VRML. Fig. 6 geeft een indruk van de opbouw van een menselijk lichaam volgens deze standaard. Een agent die volgens de standaard gebouwd is kan in elke VRML wereld geplugd worden en gecontroleerd worden via zijn interface. Animaties kunnen aan de H-Anim agents toegevoegd worden.
- De 'MPEG-4SNHC' standaard, een verzameling van definities, inclusief verzamelingen van parameters voor gezichtsanimatie en kenmerken voor gezichtsonderdelen [11].

- De 'Living Worlds' Standaard. Het doel van deze standaard is een conceptueel raamwerk te definiëren en interfaces vast te leggen om het creëren van multi-user en multi-developer toepassingen in VRML [12] te ondersteunen. Standaarden moeten toepassingen mogelijk maken die de virtuele aanwezigheid van vele mensen in een enkele scene op dezelfde tijd ondersteunen: mensen die in interactie zijn met objecten in de scene en met elkaar. Bovendien moeten ze toelaten dat toepassingen worden geassembleerd uit bibliotheken van componenten die onafhankelijk van elkaar door meerdere aanleveraars zijn ontwikkeld.

Jacob is gebouwd volgens de H-anim standaard. Zoals

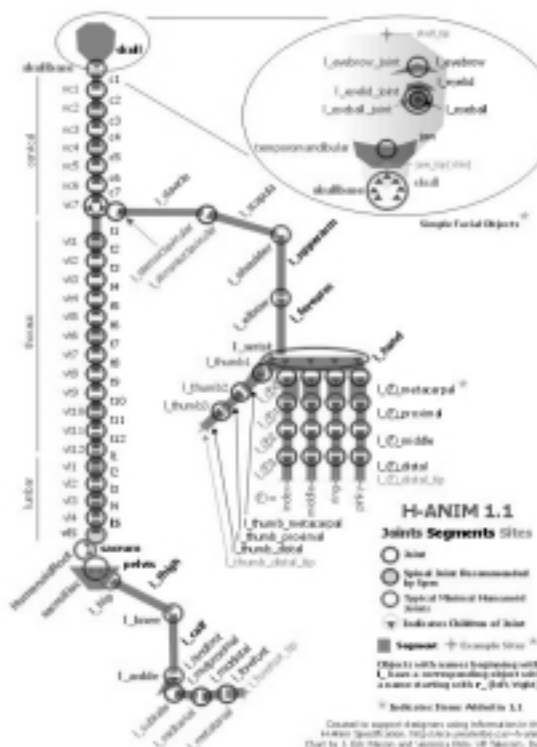


Fig. 6 Standaard voor het Menselijk Lichaam

vermeld, op dit moment gebruiken we het DeepMatrix systeem voor onze multi-user omgeving. Het systeem is overeenkomstig de Living Worlds specificatie. Men kan zeggen dat deze specificatie zorg draagt voor data distributie en scene (lokatie) synchronisatie. Een stap naar beneden krijgen we de standaarden die betrekking hebben op netwerken en applicatieprotocollen. Bij een stap naar boven krijgen we te maken met aspecten die spelen als we in de richting willen gaan van gestandaardiseerde platformen voor met elkaar communicerende agents in virtuele omgevingen.

## 5 Conclusies en Toekomstig Onderzoek

Tot slot, we denken dat voor onze omgeving de volgende drie onderzoekstrajecten simultaan moeten worden afgelopen:

- Het herontwerpen en uitbreiden van ons agent raamwerk zodanig dat individuele agents (menselijke) bezoekers kunnen representeren (inclusief bewegingen, lichamshoudingen, niet-verbaal gedrag, etc.) en eveneens kunnen staan voor artificiële, belichaamde domein agents die bezoekers in de virtuele wereld kunnen helpen (met gebruikmaking van multimodale interactie, inclusief spraak en taal).
- Het ontwerpen van H-Anim agents die beheerd worden overeenkomstig het protocol van het agent raamwerk en zich kunnen bewegen in de virtuele omgeving (daarbij acterend als een domeinagent, intelligent en autonoom gedrag vertonend, of staan voor een bezoeker en zijn of haar activiteiten in de omgeving).
- Het bestaande agent raamwerk omvormen tot een raamwerk dat ingepast kan worden in een theorie van multi-agent systemen en het leggen van relaties met aspecten zoals autonomie, reactief en pro-actief gedrag, sociale vaardigheden en lerend gedrag. Algemene raamwerken voor intelligente agents zijn ontwikkeld, waarbij de meest bekende wellicht die van de BDI (Belief, Desire, Intention) agents, eventueel nog uitgebreid met een emotioneel aspect. Bestaande raamwerken kijken echter niet naar uiterlijke kenmerken (fysiek voorkomen) van agents, natuurlijke interacties en bewegingen.

In een multi-user omgeving moet besloten worden welke delen van de omgeving gedeeld worden en welke delen voor de verschillende bezoekers privé zijn, dat wil zeggen, delen van de omgeving waarin gebeurtenissen en veranderingen alleen maar merkbaar zijn voor de individuele bezoeker, niet leidend tot updates van dat onderdeel op de computers bij andere bezoekers die dat onderdeel bezoeken.

Gedeelde omgevingen waarin we interacties hebben met de wereld en met agents en bezoekers in de wereld leiden tot tal van interessante problemen. Bijvoorbeeld, heeft iedere bezoeker van de omgeving zijn 'eigen' Karin? Of moeten ze in de rij staan om aan de beurt te komen om met haar te praten? In fig. 4 zien we dat Karin kijkt naar een andere (in de figuur niet zichtbare) bezoeker in plaats van naar Jacob. Niet-verbaal gedrag, inclusief het modelleren van kijkgedrag, wordt belangrijk, zeker wanneer we meer dan twee bezoekers hebben die verzeild raken in een conversatie met een agent. In [13] rapporteren we over experimenten en een

prototype versie van een omgeving waarin dergelijke conversaties plaats vinden.

Als een ander voorbeeld van problemen die we tegenkomen wanneer we een omgeving delen met anderen, bekijk de theaterzaal wanneer we besluiten voorstellingen op het podium te laten plaats vinden. Zoals genoemd in [8], wanneer we een grote menigte als publiek hebben, dan worden daarmee alle 'real world' logistieke problemen van evenement presentatie geïntroduceerd, inclusief zaken zoals het toekennen van plaatsen, zichtlijnen, etc. Dit is vergelijkbaar met een situatie beschreven in het boek *Snow Crash* van Neal Stephenson (1993), waar honderdduizenden avatars een groot amphitheater vullen om een voorstelling bij te wonen. Een alternatief zou zijn de zaal binnen te gaan zonder de omgeving te delen met anderen, dus in staat te zijn de beste stoel uit te kiezen (zoals ook vele anderen willen), het podium op te gaan zoals Jacob in Fig. 5, de plaats van een acteur in te nemen, etc.

Als eindconclusie willen we nogmaals de noodzaak van standaarden benadrukken. Dat begint echter een tamelijk geaccepteerde mening te worden. Voorbeelden van agent communicatie standaarden die de laatste jaren zijn geïntroduceerd voor het verkrijgen van uitwisselbaarheid zijn FIPA (Foundation of Intelligent Physical Agents: <http://www.fipa.org/>) and KQML (Knowledge Query and Manipulation Language). De FIPA specificatie staat de constructie en het management toe van een agentsysteem met verschillende agents en mogelijkere wijs gebouwd door verschillende ontwikkelaars. De standaard specificeert hoe agents kunnen interacteren met mensen, met andere agents, non-agent software en de fysieke wereld. De FIPA Agent Communication Language (ACL) is gebaseerd op 'speech act' theorie. Dit zou een nuttig startpunt kunnen zijn voor het koppelen van zulke formele communicatietalen met het gebruik van natuurlijke taal en spraak in dialoogsystemen tussen bezoekers en artificiële agents [2] of voor gemedieerde communicatie tussen bezoekers van virtuele omgevingen. Ook voor dialoogsystemen worden pogingen gedaan om standaarden in te voeren [4], bijvoorbeeld ook door naast speech acts 'dialogue acts' te onderscheiden. Recentelijk heeft DARPA binnen het eigen Communicator Program voor gesproken dialoogsystemen een dialoogstelsel architectuur (het Galaxy systeem van MIT) geadopteerd als referentie architectuur voor iedere onderzoeksgroep die mee wil doen met het onderzoek binnen DARPA (<http://fofoca.mitre.org/>). Men mag veronderstellen dat op deze wijze meer commerciële standaarden naar boven zullen komen binned de spraak- en taaltoepassingsgebieden.

Het is duidelijk dat om in de toekomst mee te kunnen doen met het ontwerpen en implementeren van virtuele

omgevingen - bewoond en bezocht door talloze agents en gebruikers, ontwikkeld en uitgebreid door verschillende ontwikkelaars - het nodig is aan standaarden zoals hier genoemd te voldoen.

## Literatuurverwijzingen

1. M. Bertolo, P. Maninetti & D. Marini. Baroque dance animation with virtual dancers. *Eurographics '99*, Short Papers and Demos, Milan, 1999, 117-120.
2. N. Dahlbäck, N. Reithinger & M.A. Walker. Standards for dialogue coding in natural language processing. Report on a Dagstuhl-Seminar, 1997.
3. M. Evers & A. Nijholt. Jacob – An animated instruction agent in virtual reality. Te verschijnen.
4. D. Gibbon, R. Moore & R. Winski. *Handbook of Standards and Resources for Spoken Language Systems*. Mouton de Gruyter, Berlin, 1997.
5. R.H. Guttman, A.G. Moukas & P. Maes. Agent-mediated electronic commerce: a survey. *Knowledge Engineering Review*, Juni 1998.
6. D. Lie, J. Hulstijn, R. op den Akker & A. Nijholt. A Transformational Approach to NL Understanding in Dialogue Systems. Proc. *NLP and Industrial Applications*, Moncton, 1998, 163-168.
7. A. Nijholt & J. Hulstijn. Multimodal Interactions with Agents in Virtual Worlds. In: *Future Directions for Intelligent Information Systems and Information Science*, N. Kasabov (ed.), Physica-Verlag: Studies in Fuzziness and Soft Computing, 2000.
8. G. Reitmayr, S. Carroll, A. Reitemeyer & M.G. Wagner. Deep Matrix: An open technology based virtual environment system. *The Visual Computer Journal* 15: 395-412, 1999.
9. B. van Schooten. Process- and agent-based modelling techniques for dialogue systems and virtual environments. CTIT Report No. 00-04, University of Twente, 2000.
10. VRML Humanoid Animation Working Group, <http://ece.uwaterloo.ca/~h-anim/>, 1998.
11. VRML-MPEG4 Working Group, <http://www.vrml.org/WorkingGroups/vrml-mpeg4/>.
12. Living Worlds Working Group: Making VRML 97 Applications Interpersonal and Interoperable. <http://www.vrml.org/living-worlds>, 1998.
13. R. Vertegaal, R. Slagter, G. van der Veer & A. Nijholt. Why conversational agents should catch the eye. Proceedings *CHI 2000*, 2000, 257-258.