Semantics and Agents for Intelligent Simulation and Collaboration in the 3D Internet

Matthias Klusch #1, Xiaoqi Cao #1, Patrick Kapahnke #1, Stefan Warwas #1

#*German Research Center for Artifical Intelligence (DFKI),*
Stuhlsatzenhausweg 3, Saarland University Campus D3.2, Saarbruecken, Germany
1{klusch, xiaoqi.cao, patrick.kapahnke, stefan.warwas}@dfki.de

***Abstract*—We present the open 3D Internet research platform ISReal for semantic-enabled intelligent 3D simulations of realities in a XML3D web browser. For this purpose, the platform integrates semantic Web technologies, semantic services, intelligent agents, verification and web-based 3D graphics. In addition, we describe our vision of a scalable web-based multi-user 3DI platform for intelligent semantic-enabled collaboration between multiple users and outline selected research challenges of realizing this vision.**

I. Introduction

The 3D Internet (3DI) is the set of 3D virtual and mixed reality worlds in the Internet that users can immersively experience, use and share with others for various applications [21], [6], [1]. For example, the 3DI offers various alternative worlds like SecondLife[1], questville, Croquet and World of Warcraft as well as mirror worlds like Twinity[2]. These worlds are mainly created for the purposes of edutainment, socializing and business collaboration in 3D meeting spaces, the 3D exploration of virtual cities, the participation in cross-media edutainment events like concerts and lectures, the trading of real and virtual assets, the functional 3D simulation of production lines and architecture at design time. Other applications include the advanced visual 3D information search with 3D Web browsers such as SpaceTime and ExitReality. In most cases, the user is represented in such a virtual world by an avatar.

One main challenge of the 3DI is to make avatars more intelligent. For example, the intelligence of avatars in every virtual 3D world today is either restricted to the direct execution of non-verbal user commands or to rather simple event rule-based but resource-optimized means of AI planning with massive volumes of action scripts in online games. Moreover, most avatars are not even capable of understanding the semantics of their 3D environment or, if the virtual world scenes are semantically annotated, do not exploit these semantics for intelligent action planning. The only approach that addresses this challenge today is the open research platform ISReal [11] which integrates semantic technologies, intelligent agents and web-based 3D graphics for intelligent 3D simulation of realities in the 3DI. However, it does not allow multiple users to collaborate in a shared 3D working space online yet while other tools that provide this functionality do not offer any support of semantic-enabled intelligent collaboration.

The remainder of the paper is structured as follows. Section 2 provides an overview of the ISReal platform while in Section 3 we demonstrate its use for a small virtual world example. Furthermore, in Section 4 we briefly describe our vision of intelligent collaboration in the 3D Internet through an example of ad hoc collaborative 3D engineering together with selected research challenges of realizing this vision. Section 5 summarizes related work for both the ISReal platform for intelligent 3D simulations and the envisioned system for intelligent 3D collaboration, and we conclude the paper in Section 6.

II. The 3DI Research Platform ISReal

The ISReal platform enables the intelligent and web-based simulation of realities in virtual 3D worlds which are described in XML3D[3][24]. In contrast to the actual 3D standard X3D, XML3D is a 3D graphics-oriented extension of HTML4. As a consequence, any virtual world scene description in XML3D can be directly embedded into a standard XHTML page. Every scene object becomes part of and accessible in the standard HTML document object and event model by any XML3D-compliant web browser which can render the scene without using any additional viewer plug-in. The kind of scenarios that can be deployed with the ISReal platform cover a wide range of applications such as demonstrators, decision support systems, and virtual training environments.

A. Architecture

The component architecture of the ISReal platform encompasses the XML3D web browser as an user interface, the global semantic environment, the 3D graphics environment, the intelligent agent environment and the verification environment (cf. figure 1). The 3D graphics environment maintains and renders XML3D scene graphs of virtual 3D worlds with the real-time scene graph system RTSG and a selected 3D rendering engine. The global semantics environment maintains not only the global scene ontologies each of which describing the semantics of a virtual 3D world in its application domain but handles the execution of registered semantic services which groundings in the animation of scene objects will have an effect on the respective scene ontology.

[1] http://secondlife.com
[2] http://www.twinity.com
[3] http://www.xml3d.org

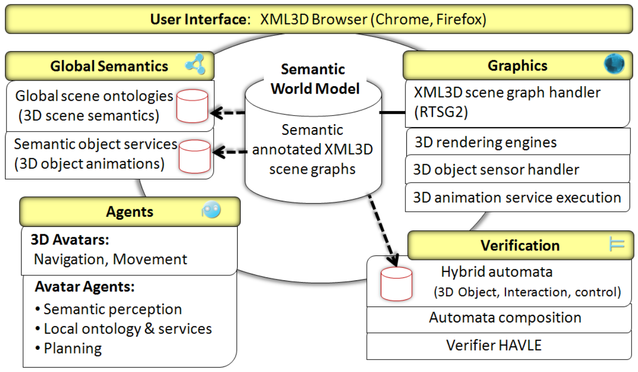

Fig. 1. ISReal platform components

The agent environment is responsible for the management and execution of intelligent agents each of which is controlling the behavior of an uniquely assigned avatar in a given virtual world. The verification environment manages and composes hybrid automata that describe spatial and temporal properties of scene objects and their interactions and verifies them against given safety requirements at design time; for reasons of space, we omit a description of this platform component. The semantic world model of the ISReal platform is the set of semantically annotated 3D scene graphs with references to the global semantics and the verification component.

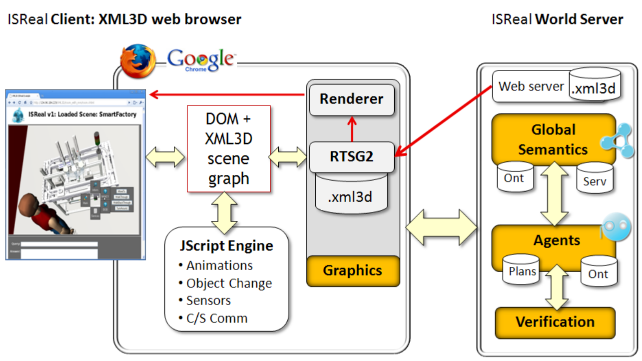

Fig. 2. ISReal platform communication architecture

The client-server-based communication architecture of the ISReal platform allows for single-user online simulations of virtual 3D world scenes only (cf. figure 2). While the environments for global semantics and agents are hosted by the ISReal world server, the XML3D-based 3D graphics environment is embedded into the XML3D web browser client. In contrast to multi-user online architectures like Sirikata, the thin ISReal client is exclusively responsible for maintaining and rendering the complete virtual world scene and communicates with the ISReal server in an asynchronous and bidirectional mode that is implemented by use of web sockets. Once the ISReal client has loaded the initial world scene page in XML3D from the web server of the ISReal server it connects to all server-sided components, initiates their scene related configurations and then starts to continously render and update the XML3D scene graph based on the user interaction with the respective virtual world scene. The ISReal platform has been fully implemented in Java, PHP and JavaScript.

B. Semantic 3D Annotation and Processing

Semantic 3D data annotation. The semantic world model of the platform is the set of all semantically annotated XML3D scene graphs for simulated virtual worlds. Any 3D scene object in a virtual world is represented as a node of the XML3D scene graph that graphically describes this world. The semantics of a 3D scene object is described in the annotation record of the respective scene graph object node by use of standard RDFa with references to the corresponding object and concept of the scene ontology in OWL2. The ISReal platform provides the user with a special tool xml3dsat for the semantic annotation of XML3D scene objects.

```
<html> <head> ...
    <script src = "ISRealServerCon.js" type="text/javascript"></script>
    <script src = "ObjIDAnnotationTable.js" type="text/javascript"></script>
    <script src = "AgentA1Sensor.js" type="text/javascript"></script>
... </head> <body> ...
<xml3d id = "scene" style = "..." about ="scene01" typeof ="3DScene" ...> ...
<group id = "doorAB" about = "http://134.96.184.229/isreal/room_abox.owl#doorAB"
            typeof = "http://134.96.184.229/isreal/room.owl#Door" ...>
   <div rel = "http://.../annotation.owl#hasService"
            resource = "http://.../open_door.owl#openDoor"/>
   <div rel = "http://.../annotation.owl#hasService"
            resource = "http://.../close_door.owl#closeDoor"/>
   <div rel = "http://.../annotation.owl#hasAutomata"
            resource = "http://.../door.ha"/>
<group id="DOOR_vertical" transform="#t_DOOR_vertical">
   <group id="shape_d1e148" shader="#s_d1e149"> <mesh type="triangles" id="m_d1e151">
...  </mesh> </group> </group> </group> ...
</xml3d> ... </body>
</html>
```

Fig. 3. Semantic annotation of XML3D scene object "door" with RDFa

For example, the semantic annotation of the "doorAB" node of the XML3D scene as sketched in figure 3 is in RDFa with references to (i) an uniquely assigned object "doorAB" of the global scene ontology in OWL2, (ii) a set of semantic services in OWL-S which describe the functionality of the "doorAB" in terms of its services for opening and closing the door and (iii) a hybrid automaton which describes the temporal-spatial property that this door can be opened and closed with angular speed of 10 degrees per second which is not possible to encode in OWL2. Both the given global ontology and semantic object services are maintained in the global semantics environment of the ISReal platform.

Global semantic environment. The global semantic environment (GSE) of the ISReal platform consists of (i) the global ontology management system (OMS) and (ii) the semantic service handler (SemSH) While the OMS maintains the global scene ontologies and the processing of semantic queries about the scene, the SemSH handles the execution of semantic scene object services which either have preconditions to be checked against the fact base of the considered scene ontology or have semantic effects on the ontology. Depending on the type of semantic query, the query processing decider of the OMS is passing the query to either the RDF triple store of the OMS or the appropriate reasoner for its processing. In particular, the OMS has an open plug-in architecture that allows to use any RDF/S store such as SwiftOLIM with materialization of OWL2 in RDF under OWL-Horst semantics and semantic reasoners like the OWL-DL reasoner Pellet[4] with internal Jena RDF store and the RDF relational reasoner STAR [12] as appropriate.

Semantic queries and service execution. In its current version, the GSE of the ISReal platform is capable of answering three different types of semantic queries about a vortual 3D world scene: (i) Object queries like ”Is the scene object doorAB locked?”, (ii) conceptual queries like ”Are all doors unlocked in general, i.e. is the class of doors a subclass of the class of unlocked objects?”' and (iii) relational object queries like ”How are the scene objects doorAB, doorBC and roomC related to each other?”'

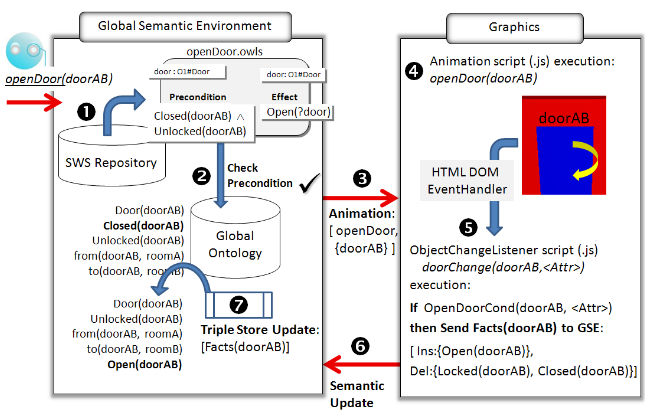

Fig. 4. Semantic-based execution of 3D scene object animation services

Figure 4 illustrates the execution of a semantic service in OWL-S with a logical precondition and a grounding in a 3D animation script which execution has a semantic effect on the fact base of the global scene ontology. In this example, an intelligent avatar wants to open a door named ”doorAB”' in a scene. For this purpose, its avatar agent does request the GSE to execute the corresponding semantic service ”openDoor”' of this scene object. In case the precondition of this services holds in the actual global fact base, the SemSH of the GSE then triggers the execution of the respective 3D animation script by the XML3D browser. If this animation, in turn, changes any of the (DOM) attribute values of the object ”doorAB”' in a way that its semantics changes like the fact that the ”doorAB”' is now open, the object script doorChange() of the ISReal client triggers the respective update of the global fact base by the OMS of the GSE in the ISReal world server.

[4]http://clarkparsia.com/pellet

C. Intelligent Avatars

An ISReal avatar is uniquely associated with an intelligent agent which is implementing the behavior of the avatar in simulated virtual 3D worlds it is involved in. The appearance of an intelligent avatar to the user is described as just another 3D scene object in the corresponding XML3D scene graph, hence the user cannot distinguish between the avatar and its intelligent agent or avatar agent, that is between the body and soul of an intelligent avatar. While the avatar is running in the client, that is the XML3D web browser, it is connected with and driven by its intelligent agent in the agent environment of the ISReal world server.

Semantic perception of the environment. The local knowledge of an intelligent avatar about the virtual world it is part of is acquired by means of semantic perception. For this purpose, an avatar is equipped with a configurable sensor to perceive any scene object within a given range and with certain frequency and resolution. The semantics of perceived objects are requested by the avatar agent from the GSE which returns the set of all object related facts from the actual global fact base of the scene ontology and the set of the object related semantic service descriptions. Different sensor configurations and local update strategies may cause agents to have different views on the simulated world scene. As a result, the local fact base of an avatar agent may be inconsistent with the global one while the conceptual knowledge about the considered virtual world is supposed to be static and common knowledge to all intelligent avatars which are acting in it.

Semantic-based hybrid action planning. Based on the semantic perception of its environment an ISReal avatar agent is capable of planning its action in order to accomplish user-given goals including the answering of complex semantic queries about annotated scene objects. For reasons of efficiency, we adopted the reactive belief-desire-intention (BDI) architecture for agents [17] that is known to be particularly appropriate for implementing a fast perception and deliberative action planning by agents in dynamically changing environments such as virtual 3D worlds. Roughly speaking, a BDI agent is equipped with a plan library of domain-dependent and -independent plan patterns which are used by a BDI planner for reactive action planning from second principles interleaved with the immediate execution of planned actions.

In particular, an intelligent avatar agent selects an appropriate BDI plan pattern of operators, instantiates it with variable bindings taken from its local fact base into actions and then executes these actions until the goal is reached or no further instantiation of the plan pattern is possible. Scene object services as well as agent services are encoded as operators such that the plan execution is performed either by the ISReal client or the agent environment of the ISReal world server. The execution of a semantic scene object service which is grounded in an animation script is outlined in figure

4 above. Agent services which groundings have no effect on the global fact base are not registered at the GSE and directly executed by the agent. In case there is no appropriate BDI plan pattern in the library, the ISReal avatar agent can invoke a semantic service composition planner like OWLS-XPlan [14] to satisfy a given goal by means of state-based planning from first principles over its local fact base.

Development of intelligent avatars. The ISReal platform supports the user in developing intelligent avatars with an integrated configuration tool. In particular, the tool is guiding the user through all stages of the development including the design of the avatar, the customization of default BDI plan patterns and the local ontology of its avatar agent (cf. Figure 5). For this purpose, the tool interfaces with other selected tools that are appropriate for the specific development task at hand. The platform-independent model-driven development of an avatar agent, for example, can be performed with the integrated BDI agent modelling framework DSML4MAS [26], [8].

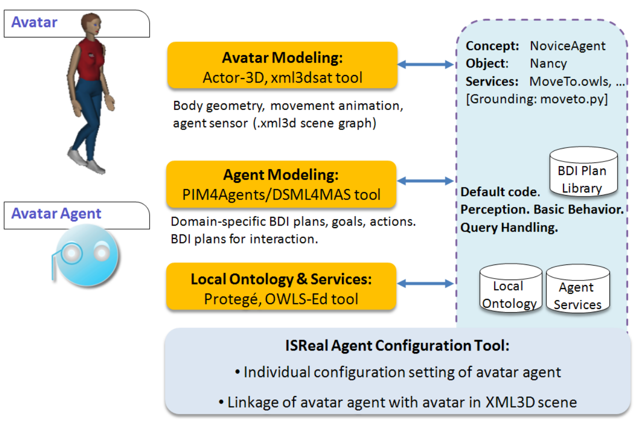

Fig. 5. Development phases of an intelligent ISReal avatar

By default each avatar agent is equipped with domain-independent BDI plans for acting in virtual worlds. These plans concern the basic 3D animation of the avatar, the processing of different types of semantic queries in the local ontology of the avatar agent and the execution handling of its actions (services) via the GSE or by itself. Other domain-dependent plans have to be added or customized for 3D scene simulations by the 3D scene agent developer. In any case, the linkage of an avatar object with its avatar agent can be specified directly in the XML3D scene graph either by the use of special XML3D attributes or preferably in the semantic annotation of the avatar in RDFa.

Simulation of human movement in the 3D scene. For the animation of a more human-like movement of an avatar in a given 3D scene, the corresponding avatar agent passes the respective agent services for its avatar movement animation to a special server-sided component named HumanSim which handles the execution of these services in the browser. The HumanSim component is running in the ISReal world server and is based on the navigation component of the commercial online game AI engine Xaitment. It allows avatars to optimally navigate through complex scene geometries without any collision with scene objects including other moving avatars. In particular, the avatar agent maps semantic information about the target position of its avatar movement such as "Nancy moveNear EngineX"' to a distinct 3D point coordinate in the scene and sends it to the HumanSim component for handling the respective movement animation. In fact, position and orientation updates are continuously pushed by the HumanSim component via WebSockets to the respective Javascript client in the XML3D browser for visualization. In addition, predefined complex animation sequences for the simulation of a large set of walking or idling avatars in a virtual world can also be handled by the HumanSim component by means of its local execution of simple finite state machine based avatar agents.

III. ISReal Application Example

We briefly illustrate the use of the ISReal platform for intelligent 3D simulations by means of a simple example scenario of a virtual training application for production lines. In this scenario the user is represented by an avatar "Nancy" in a small virtual world of a pill production line which includes a pill production station Y in room E and a pill filling station X in room D. Furthermore, the world is populated with one avatar "Bob" which is an expert for both stations and several pharmaceutical retail avatars. Suppose that "Nancy" has to demonstrate its user how to produce one bottle of 20 blue pills with the filling station X.

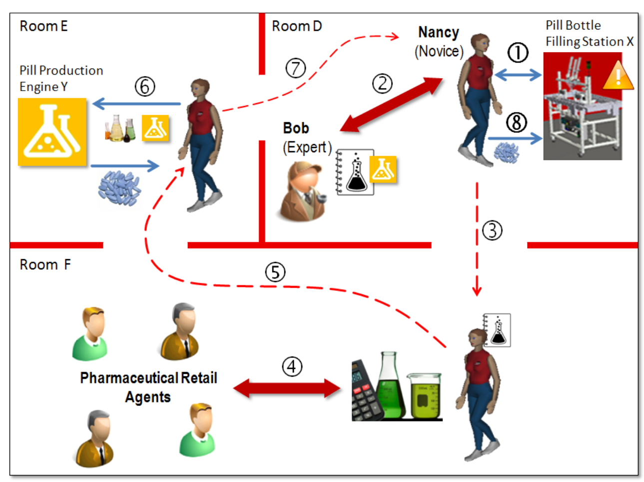

Fig. 6. Application scenario for an intelligent avatar "Nancy" in the small 3D world SmartFactory

For this purpose, the intelligent avatar "Nancy" is moving to the station X, perceives all its components and thereby understands its functionality. Based on this knowledge "Nancy" generates a plan for filling a bottle with 20 blue pills and tries to execute it with station X. In this scenario, however, this plan execution fails since station X does not have enough pills in its depot. Since the alarm signal of the station does not indicate the cause of the plan execution failure, the user avatar is searching for other avatars in the scene for help. After respective interaction with the expert ”Bob” in the same room, the user avatar ”Nancy” knows why its planned filling action with station X failed and got the receipt from ”Bob” for producing the missing number of pills with the other station Y of the production line.

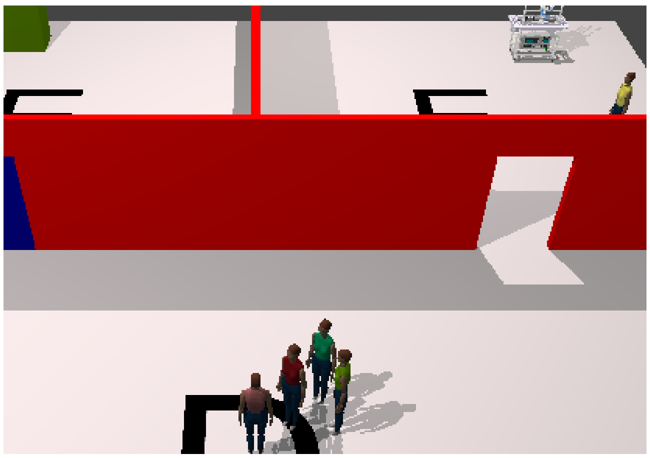

Fig. 7. Snapshot of application scenario running in a XML3D web browser: Negotiation of ”Nancy” with three pharma retail agents.

As a result, ”Nancy” starts to search for the required pill ingredients or other avatars in the scene which can provide them. In this example, ”Nancy” encounters several pharma retail agents in room F and successfully negotiates with them the prize of these ingredients (cf. figure 7). Eventually, ”Nancy” moves to station Y in room E in order to produce the missing pills, then refills the pill depot of station X in room F with the produced pills and finally succeeds with the execution of its original plan to accomplish the given goal. To the best of our knowledge, this kind of intelligent simulation in virtual worlds is not possible to implement with any software other than the ISReal platform yet.

We initially measured the platform performance for the small virtual world example above in terms of (i) the throughput of its global semantic environment, (ii) the average time of interaction between semantics and graphics for service execution and (iii) the average times of semantic perception, BDI planning, and semantic composition planning of an intelligent avatar. Overall the preliminary experiments revealed that the platform performs reasonably fast for small virtual worlds with a medium-sized global ontology, a small number of annotated 3D scene objects and up to a few dozens of agents. For example, it would take only 1.5 seconds of simulation to let 30 agents concurrently open 30 doors in the scene while the only slow down has been caused by the in general PSPACE-hard semantic service composition planning of avatar agents.

IV. Towards Intelligent 3D Collaboration

The ISReal platform allows to perform intelligent 3D simulations of realities in virtual worlds in XML3D. However, its single-user communication architecture does not allow any ad hoc collaboration between multiple users like in multi-player online games. In the following, we sketch our vision of a scalable web-based multi-user 3DI platform for intelligent collaborative 3D engineering and present selected research challenges of realizing this vision.

A. Limitations of 3D Collaboration Today

In the 3DI of today, the support of user collaboration still is very limited. In fact, the joint experience, use and share of arbitrary 3D scenes, media content and services for various applications by users of virtual 3D worlds at any time and anywhere is often claimed to be essential but not supported yet[21], [6], [1]. On the one hand, users can collaborate in publicly accessible but proprietary virtual 3D worlds such as SecondLife or Twinity. Such collaboration concerns, for example, the joint organisation of edutainment live events, the joint exploration of mirror cities and the joint interior design of virtual apartments or houses. In this context, mobile ad-hoc collaboration between users supported by a fast semantic search and composition of relevant 3D content and services is considered essential but not possible.

On the other hand, professional collaborative applications in the 3DI such as for joint virtual engineering of products and production facilities are in use mainly at large companies without web-based 3D simulations. Mobile and ad-hoc collaboration among relevant experts at different locations around the globe for tasks of joint construction, planning, design, formal analysis and training in shared 3D spaces is highly desirable but not yet possible. There are also no systems or tools available that support team members to jointly perform an integrated semantic search, composition and simulation of relevant parts of or alternatives to virtual 3D models of products with test data from different sources.

Currently, the semi-automated semantic annotation and simulation of designed virtual 3D models for product analysis by experts make extensive use of various proprietary and typically not interoperable computer-aided engineering (CAx) and visualisation systems. Only few of those systems are offering limited support of task-specific collaboration to their users in closed modules but without any advanced semantic or reliability checking technologies. None of them support mobile ad-hoc collaborative engineering sessions. In summary, there is no coherent framework of semantic-enabled intelligent support for collaborative 3D engineering in the 3DI available yet.

B. Intelligent Collaborative 3D Engineering

What is our vision of intelligent collaborative 3D engineering? In line with our work on the 3DI research platform ISReal, we assume XML3D to become the standard exchange format for a web-based visualization of and interaction with 3D content on different kinds of computational devices including notebooks, tablet PCs and 3D smartphones. We further envision the underlying system to have a scalable multi-user communication architecture and to provide intelligent support of ad hoc collaborative

3D engineering of products within peer networked user communities. Like in ISReal, intelligent avatars are supposed to represent and assist their users in shared 3D spaces. In particular, the agents are able to jointly perform a reasonably fast high-precision semantic search and composition of relevant annotated 3D models of product units, simulation services or unit testing data in the 3DI. They can even handle vague semantic queries for 3D content and services and adapt to the 3D environments they are involved in at run time such as ad hoc changes of the set of users, their preferences and goals, the set of accessible 3D scenes and simulation services. In the following, we call such a visionary system for intelligent collaboration in the 3D Internet Collaborate3D.

Application scenario. In the following, we illustrate the use of the envisioned Collaborate3D system for a simple example scenario of web-based collaborative 3D product engineering (cf. figure 8).

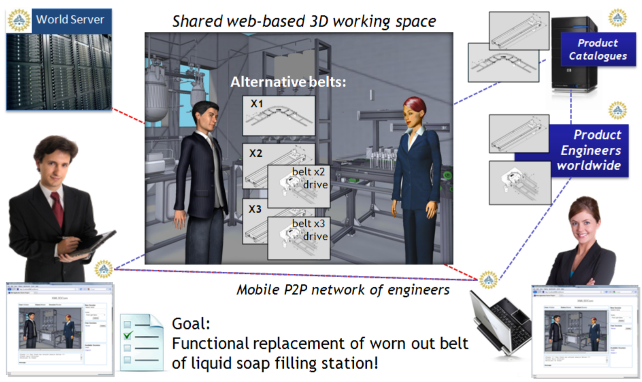

Fig. 8. Example of collaborative 3D engineering with intelligent avatars

Suppose the worn out transportation belt of a liquid soap filling station of some production line has to be replaced with either a new model of a linear belt or a more space-efficient curved belt. For this purpose, the product engineer who is in charge of accomplishing this goal uses his Collaborate3D client to initialise an ad hoc collaborative 3D engineering session with anotehr trusted expert whom he invites to join a shared virtual 3D space from remote in a secured mobile peer-to-peer network. In the future 3D Internet, the set of relevant XML3D models of belts required to accomplish the given design goal will be accessible not only from a few online product catalogues or CAD systems but from many individual experts world wide. Each of them may use a Collaborate3D client to ad hoc join the product design team either for actively participating in the collaborative 3D engineering session or just for granting other team members the access to relevant local items. Like in ISReal virtual worlds, team members can be represented in the shared 3D space of the session by intelligent avatars.

In this example, the intelligent avatar agents of team members are entrusted with jointly searching for and composing of annotated XML3D models of belts, drives, stoppers, mounts and sensors that are semantically relevant for accomplishing the given design goal. Eventually they display the possible alternative belts to all team members in the shared 3D space of their collaborative 3D engineering session. The agents can also search for vaguely described items like belts that are similar to the concept of a linear or curved driven belt for the given soap filling station, or belts that roughly fit into the indicated spatial area of this station. Each of both variants of belts requires the engineering of a different drive and belt system for the transport of soap flasks. The found alternatives of linear and curved belts can be ranked with respect to the degree of their satisfaction of requested technical requirements of functional and safety properties. For this purpose, the agents may apply formal verification means to the respective part of the annotation of the XML3D models of belts.

In addition, the team wants to test whether their virtually designed facility can fill soap flasks of at least two different sizes and shapes. For this purpose, the intelligent avatar agents jointly answer the given semantic query for flasks that are appropriate for liquid soap containment and have the necessary geometrical properties for each of the selected belt alternatives. In particular, the avatars support the discussion of their users on the 3D visualized result of their search in the shared 3D space by indicating the current view position and direction as well as pointing gestures. An additional formal verification of found flasks with different shapes determines whether they are fitting on the conveyor, are able to pass the planned conveyor radiuses and will fit to filling adapters and that the system control software is capable of handling different filling processes. For example, in case of special offers the redesigned filling station must be able to handle also non-standard values of weight, volume, lot size per time unit and duration of the filling process. The agents can refine the constraints for their peer-to-peer search and composition of flask models in XML3D based on the verification result and the feedback that is either explicitly or implicitly provided by their users during the collaborative 3D engineering session.

C. Selected Research Challenges

The realisation of this application scenario of intelligent collaborative 3D engineering requires innovations in the areas of mobile web-based 3D graphics, intelligent agents, verification and semantic technologies. In the following, we discuss only a few selected relevant research challenges in the latter area.

Semantic P2P retrieval of scene objects. How can intelligent avatars perform a highly precise and scalable retrieval of relevant scene object models and services in a mobile P2P network with an unstructured or even hybrid overlay? In particular, how to best exploit the semantic annotations of such objects and services for this purpose?

One common approach to improve the retrieval performance in unstructured P2P networks is to complement the query routing mechanism for each peer with methods for near-optimal dynamic data replication [22]. However, there is no method for dynamic semantic replication available yet.

In this regard, one solution approach is to let each peer continuously determine the demand-supply relations among all peers in the network based on the semantics of both queries and items it knows about. The dynamic adjustment of these relations results in a up-to-date semantic overlay that keeps the individual decisions of peers about item replication adhering to their current interests. It can be proven that any random k-walker search in unstructured P2P networks when combined with this dynamic semantic replication strategy outperforms the combination with a near-optimal non-semantic replication strategy. Further investigation and experimental evaluations are needed to show whether the same holds in comparison with other semantic P2P search algorithms like Bibster or RS2D [13], [25].

Another related challenge is the semantic retrieval of items in hybrid P2P networks. In other words: How to best search for semantically relevant items across connected subnetworks of different type and routing protocol? One idea to address this challenge is to equip those peers that are members of multiple different P2P subnetworks like one structured and one unstructured P2P network with an adaptive intelligent decision procedure for this purpose [19].

Flexible semantic selection of scene objects. How can avatar agents answer user queries about 3D scene objects that include only vague descriptions of their functionality and geometry like "Find something similar to a door that roughly fits into this region"? In this regard, methods for querying XML, RDF and OWL data and services under uncertainty are first-class candidates for research in this direction [7], [23], [20], [3]. Another issue of semantic selection of annotated XML3D scene object models is the optimally weighted aggregation of results which are provided by different kinds of semantic and geometrical 3D matching filters. In this regard, the appropriate adoption of adaptive methods for hybrid semantic service matching like in iSeM [15] could be helpful for solving the problem.

Another strain of related research is centred around the problem of how to efficiently perform the semantic alignment of peer ontologies in a distributed way. Except for a few notable approaches to ontology matching in P2P networks like H-MATCH [4], there are no practical solutions available yet. For a fast, heuristic-based preselection of scene objects, one might extract the semantic annotation records of scene objects from annotated XML3D scene graph files into RDF encoded summaries such that these linked data sets can be efficiently queried with SPARQL [9] and fuzzy extensions of it such as f-SPARQL [5].

One idea to speed up the process of semantic selection of XML3D scene graphs is to use graph kernels for fast computations of structural similarities. In fact, graph kernels are tractable approximated solutions to the NP-hard subgraph isomorphism problem in practice. Since the computation of graph kernels largely bases on operations with label and adjacency matrices, one could make use of GPU programs for extremely fast matrix computations in order to gain a significant speed up of graph kernel-based semantic matching of XML3D scene graphs in practice.

Fast semantic P2P composition planning. In case there are no sufficiently matching XML3D scene objects available, the intelligent avatars of the collaborative 3D engineering session can jointly perform a distributed semantic composition planning in order to accomplish the given goal. For example, the composition of a transportation belt and its belt drive system each of which represented in XML3D with semantic annotation makes up a conveyor or belt. Since the functionality of both components is described by means of semantic services, semantic service composition planning can be exploited to find a combination of XML3D scene objects that would satisfy a given query. However, it remains unclear how to efficiently perform such a semantic composition planning in a distributed way in unstructured P2P networks [18].

One related challenge is to effectively switch between different kind of planners based on the situation at hand in order to improve the overall planning performance under resource constraints. In this regard, hybrid BDI planning methods could serve as a starting point for research into this direction. One idea would be to let intelligent avatars learn when and how to best switch between available planning algorithms for their semantic service composition with a reasonable trade-off between plan quality and planning performance, that is the adaptive planning of good-enough compositions of scene object services for a given query.

V. Related Work

To the best of our knowledge, the ISReal platform is the only open research platform for web-based and semantic-enabled intelligent simulations in the 3D Internet. However, it does not support such simulations between multiple collaborating users in shared 3D spaces online. There are several related approaches that address this challenge including the Semantic MediaWiki platform[5] and the OpenCobalt platform[6]. The first platform extends wiki technology with semantic annotation and querying of content[7] However, the platform does not allow for 3D simulations nor flexible semantic query answering under uncertainty nor does it provide any functionality for a semantic peer-to-peer search and composition of relevant items. The latter also holds for the NSF-funded OpenCobalt platform for a web-based 3D collaborative work space.

The mobile application STEVIE[8] enables users to collaboratively create, share and modify semantic points ofinterests (POIs), i.e. geographic places with explicit semantic properties of a collaboratively created multimedia POI ontology (text, videos, pics) on mobile phones. The semantic POIs and collaborative POI ontology are published as Linked Open Data while data mining techniques are employed to cluster and thus improve the quality of the collaboratively created

[5] http://semantic-mediawiki.org/wiki/Semantic_MediaWiki
[6] http://www.opencobalt.org
[7] European research project ACTIVE: www.active-project.eu
[8] http://isweb.uni-koblenz.de/Research/systeme/csxPOI

POIs. However, the visualization of spatial information is restricted to the use of a 2D map and the semantic search for relevant POIs is performed on a central server rather than by means of a distributed P2P search. The same basically holds for other semantic social Web-based collaboration tools like Kiwi, Taglocity and LiquidJournal. Other social web-based collaboration tools include IBM Lotus Connections, Oracle Beehive, GoogleDocs, Skype, thinkature, bubble.us, and so on. However, these tools do not offer any support of intelligent semantic-enabled collaboration by an integrated utilization of semantic technologies and intelligent agents in mobile shared 3D spaces.

On the other hand, it is common sense that the use of semantics can greatly improve the management and retrieval of 3D content as it has been demonstrated for various applications in different domains such as in arts, bioinformatics, gaming, cultural heritage and virtual museums. However, the functionality of the ISReal platform goes far beyond of these semantic-enabled applications. Related work on virtual agents or intelligent avatars in 3D worlds such as in [2] and STEVE [16] focus on advanced multi-modal user-agent interaction and planning from second principles only. However, unlike ISReal avatars none of these avatars is capable of understanding the semantics of its environment by means of semantic-based perception and reasoning on annotated 3D scenes nor do they perform any semantic planning from first principles. These capabilities are required for acting in virtual worlds that are unknown to the avatar.

VI. Conclusions

We presented the open 3DI research platform ISReal for semantic-enabled intelligent simulations in virtual 3D worlds. Its core innovation is the integration of semantic Web technologies, semantic services, intelligent agents, verification and web-based 3D graphics for this purpose. Since the ISReal platform has a single-user communication architecture, we also described our vision of a web-based multi-user 3DI platform for semantic-enabled intelligent collaboration between multiple users. The realisation of this vision requires to address major research challenges in the areas of mobile web-based 3D graphics, intelligent agents, verification and semantic technologies. However, it remains to be seen whether the underlying idea of creating a semantic 3D Internet and its enabling technologies to be invented will be eventually accepted and widely adopted by industry and the common user in practice.

Acknowledgment

The reported work on the ISReal platform has been funded by the German Ministry for Education and Research (BMBF) under the project grant 01IWO8005.

References

[1] Alpcan, T.; Bauckhage, C.; Kotsovinos, E. (2007): Towards 3D Internet: Why, What, and How? Proc. IEEE Intl. Conference on Cyberworlds.

[2] Anastassakis, G.; et al. (2001): Multi-agent Systems as Intelligent Virtual Environments. LNCS 2174, Springer.

[3] Cassar, G.; Barnaghi, P.; Moessner, K. (2011): A Probabilistic Latent Factor Approach to Service Ranking. Proc. IEEE Intl. Conference on Intelligent Computing and Processing.

[4] Castano, S.; Ferrara, A.; Montanelli, S. (2003): H-MATCH: an Algorithm for Dynamically Matching Ontologies in Peer-based Systems. Proc. 1st Intl. VLDB Workshop.

[5] Cheng, J.; Ma, Z.M.; Yan, L. (2010): f-SPARQL: A Flexible Extension of SPARQL. Proc. Intl. Conference on Database and Expert Systems Applications (DEXA), LNCS 6261, Springer.

[6] Daras, P.; Alvarez, F. (2009): A Future Perspective on the 3D Media Internet. Towards the Future Internet - A European Research Perspective, IOS Press.

[7] Fazzinga, B.; Flesca, S.; Pugliese, A. (2009): Retrieving XML Data from Heterogeneous Sources through Vague Querying. *ACM Transactions on Internet Technology*, 9(2).

[8] Hahn, C.; Madrigal-Mora, C.; Fischer, K. (2009): A platform-independent metamodel for multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 18, Kluwer Academic.

[9] Hartig, O.; Langegger, A. (2010): A Database Perspective on Consuming Linked Data on the Web. *Datenbank-Spektrum*, 10, Springer.

[10] Ibanez-Martinez, J.; Mata D. (2006): Virtual Environments and Semantics. *European Journal for the Informatics Professional*, 7(2).

[11] Kapahnke, P.; Liedtke, P.; Nesbigall,S.; Warwas, S.; Klusch, M. (2010): ISReal: An Open Platform for Semantic-Based 3D Simulations in the 3D Internet. Proc. 9th International Semantic Web Conference (ISWC), LNCS 6414, Springer

[12] Kasneci, G.; Ramanath, M.; Sozio, M.; Suchanek, F.M.; Weikum, G. (2009): STAR: Steiner Tree Approximation in Relationship-Graphs. Proc. 25th IEEE Intl. Conference on Data Engineering (ICDE).

[13] Klusch, M.; Basters, U. (2006): Risk driven semantic p2p service retrieval. Proc. 6th IEEE Intl. Conference on P2P Computing (P2P), IEEE CS Press.

[14] Klusch, M.; Gerber, A. (2006): Fast Composition Planning of OWL-S Services and Application. Proc. 4th IEEE European Conference on Web Services (ECOWS), IEEE CS Press.

[15] Klusch, M.; Kapahnke, P. (2010): iSeM: Approximated Reasoning for Adaptive Hybrid Selection of Semantic Services. Proc. 4th IEEE Intl. Conference on Semantic Computing (ICSC), USA.

[16] Rickel, J.; Johnson, W.L. (2001): Task-oriented collaboration with embodied agents in virtual worlds. Embodied conversational agents, MIT Press.

[17] Rao, A.S.; Georgeff, M.P. (1995): BDI Agents: From Theory to Practice. Proc. Intl. Conf. on Multi-Agent Systems (ICMAS), AAAI Press.

[18] Rao, J.; Kungas, P.; Matskin, M. (2006): Composition of semantic web services using linear logic theorem proving. *Information Systems*, 31(4-5).

[19] Rosenfeld, A.; Goldman, C.V.; Kaminka, G.A.; Kraus, S. (2007): An Architecture for Hybrid P2P Free-Text Search. Proc. 11th Intl. Workshop on Cooperative Information Agents, LNCS 4676, Springer.

[20] Simou, N.; Stoilos, G.; Tzouvaras, V., Stamou, G.; Kollias, S. (2008): Storing and Querying Fuzzy Knowledge in the Semantic Web. Proc. 4th Intl. Workshop on Uncertainty Reasoning for the Semantic Web (URSW), CEUR 423.

[21] Smart, J.; Cascio, J.; Paffendorf, J. (2007): Metaverse Roadmap - Pathways to the 3D Web. www.metaverseroadmap.org

[22] Sozio, M.; Neumann, T.; Weikum, G. (2008): Near-optimal dynamic replication in unstructured peer-to-peer networks. Proc. 27th ACM Symposium on Principles of Database Systems (PODS).

[23] Straccia, U. (2009): A Minimal Deductive System for General Fuzzy RDF. Proc. Intl. Conference on Web Reasoning and Rule Systems (RR), LNCS 5837, Springer.

[24] Sons, K. et al. (2010): XML3D Interactive 3D Graphics for the Web. Proc. Web3D 2010 Conference.

[25] Staab, S.; Stuckenschmidt, H. (2006) eds.: Semantic Web and Peer-to-Peer. Springer.

[26] Warwas, S.; Hahn, C. (2009): The DSML4MAS development environment. Proc. 8th Intl. Conf. on Autonomous Agents and Multiagent Systems (AAMAS).