Towards Integrated Intentional Agent Simulation and Semantic Geodata Management in Complex Urban Systems Modeling*

Klaus Baier¹, Victor Mataré¹, Martin Liebenberg², and Gerhard Lakemeyer²

¹Chair of Engineering Geology and Hydrogeology, RWTH Aachen University, Lochnerstraße 4-20, D-52064 Aachen, Germany ²Knowledge-Based Systems Group, RWTH Aachen University, Ahornstraße 55, D-52056 Aachen, Germany

Abstract Mega-urbanization presents researchers with a network of densely interwoven problems that elude disciplinary boundaries. We report on the development of a spatial knowledge management and agent simulation framework that is designed to integrate closely with the process of trans-disciplinary research into the dynamics of complex human-environment systems. We argue that our choice of knowledge representation languages facilitates cross-domain collaboration. In a runthrough application example, we show how standardized knowledge engineering technologies are used to turn a conventional geodatabase into a self-documenting knowledge base that can flexibly interface with modern open-data infrastructures. The resulting cross-domain world model is then coupled to a graphical actor modeling language that specializes in the formulation of behavioral theories in terms of social roles, intentions, tasks, conditions and interaction. Finally, we describe how system theories expressed in this way are automatically translated into computer simulations.

1 Introduction

When developing an understanding of human-environment interaction, the decision-making of local stakeholders is a crucial aspect, especially if one hopes to predict how altered policies or infrastructure play out in the long run.

The trans-disciplinary research dealing with complex humanenvironment systems faces serious difficulties synthesizing discipline-specific perspectives into coherent theoretical systems. We believe that integrating scientific knowledge management with modeling and simulation tools can provide a huge benefit to this kind of research.

This paper is a progress report on the $SiKAMUS^1$ Project, where we are developing a modeling and simulation framework that combines semantic geodata management with an intentional actor model. The i^* diagram language (Yu, 1995) captures actor intentionality in terms of goals, actions, dependencies and their interaction with the environment.

We use the Protégé OWL editor (an integrated development environment for the Web Ontology Language (cf. Hitzler et al., 2012; Knublauch et al., 2004)) to construct a semantic geodata model that qualitatively describes the environmental system. It makes the expressive power of the OWL family of knowledge representation languages intuitively accessible to users without any technical background knowledge.

Combining OWL and the i^* agent modeling language yields a formal system that minimizes the semantic gap between domain knowledge (as expressed in the descriptive meaning of scientific theories) and the formal-symbolic representation that makes it executable in a computer simulation. The entire framework aims to become a suitable toolbox for integrating knowledge management, modeling and simulation into the process of any trans-disciplinary research that deals with the emergence of complex human-environment interaction. The most crucial aspect however, is the flexibility gained in communication, design and criticizability of the model structure.

In this paper, we will describe the current and planned development of the framework. Section 2 outlines our current application domain. In Section 3, we argue from a knowledge management and engineering perspective that the choice of domainspecific languages matters greatly in a trans-disciplinary modeling framework. Section 4 details our use of OWL to represent environmental knowledge in a way that fits with the i^* actor modeling exemplified in Section 5. In Section 6, we will reflect on some shortcomings of the current approach and sketch some of the work required to remedy them.

2 Application Domain: Mega-Urbanization and Water Resources

Globally, the understanding of highly dynamic urbanization processes increases in importance as already 540 million people are living in cities with more than five million inhabitants – mostly located in developing and newly industrialized countries (Kraas and Nitschke, 2006).

In order to understand the systematics of rapidly urbanizing regions, urban development has to be analyzed by taking into account both structural contexts and action modes. Giddens, 1984 for example, develops a theory of structuration, and Werlen, 2000 outlined the theoretical background for a series of international research in the 1990s.

In 2007, for the first time in human history, more people resided in urban areas than in rural areas. At the same time, megacities' share of the urban population reached more than 40%, highlighting the preponderance of the largest agglomerations over smaller cities in the urbanization process, corresponding to Henderson's urbanization and development theory (Henderson, 2002). In India, for example, cities are expanding rapidly as increasing numbers of migrants stream into urban areas in search of economic safety (The World Bank, 2011), causing the slum population to double during the past two decades

^{*}This paper is a preprint of a journal article published in Computers, Environment and Urban Systems, Volume 51, May 2015 (doi:10.1016/j.compenvurbsys.2015.01.006).

¹ Simulating Knowledge-based Agents in Mega-Urban Systems – a trans-disciplinary research project of the Chair of Engineering Geology and Hydrogeology (LIH) and the Knowledge-based Systems Group (KBSG) at the RWTH Aachen

(Ministry of Housing and Urban Poverty Alleviation, 2009).

Population growth and migration contribute immensely to the development of megacities and mega-urban areas. Megacities are quantitatively defined as cities having a population of more than five (Bronger, 1996), eight (UN, 1987; Fuchs et al., 1994; Chen and Heligman, 1994) or ten million people (Mertins, 1992). However, major challenges for the water management in a megacity include settlement rate, infrastructural requirements and land use. The ever increasing number of factors and complexity, leads to coordination and steering problems (Heinrichs, 2009).

This high-speed urbanization describes a mostly uncontrollable interplay between processes of land use transformation and large-scale migration with far-reaching consequences for the environment and society (Wehrhahn et al., 2008). Massive land use and land cover changes initially take place within already existing build-up areas; then they expand outwards in the adjacent suburban region and the urban fringe (Wiethoff et al., 2011). Typical challenges are

- lack of, or missing land use planning and growth control on the governmental level,
- considerable displacement processes in the real estate and capital market (economic level), and
- lack of access to urban infrastructures on the individual level.

Informal development in all areas of urban activity is fertilized by the general loss of ability to govern and by a law enforcement that struggles with social conflicts and spatial disputes. Due to the multidimensional challenges of poverty, social inequality, marginalization, fragmentation, and environmental degradation, cities are increasingly becoming places of potential risk, linked to various natural and man-made hazards such as heat waves and pollution. The complexity of issues, along with limited steering and control leads to the development of complex formal and informal adaptation and coping strategies (Pelling, 2003).

There has been a comprehensive amount of research on rapidly urbanizing regions since the last two decades, regarding their growth as well as the amplifying problems related to their (natural) resource needs, (cf. e.g. Feldbauer and Parnreiter, 1997; Xu et al., 2002; Kraas and Nitschke, 2008). However, studies that investigate the needs and options of future settlement scenarios in legible levels of detail (such as their real spatial-cultural implications), while being based on objectifiable data (such as related to vulnerabilities or water budgets) – are still rare. Also, heterogeneity of land use and settlement structure is generally dismissed as a threat for spatial planning and resource management. The phenomenon is rarely examined as an enabling condition for adaptive reactions in response to large-scale influences such as population growth or climate change.

Considerations regarding water resources and global city growth are increasingly integrated (Shannon, 2008; Anuradha and Cunha, 2009). Wehrhahn et al., 2008; Baier et al., 2009; Strohschön et al., 2009; Strohschön et al., 2012 and Baier and Strohschön, 2012 analyzed the impact of urbanization on the water resources of Guangzhou, China. Here, foundations were laid for intensifying research on hydrological problems such as water pollution or distribution of water in rapidly urbanizing regions. Driving forces of water shortages in the face of rapid urbanization in Asian developing and newly-industrialized countries include population growth, growth of unplanned (informal) settlements in the (sub)-urban areas, lifestyle changes and resulting increasing per capita water consumption (Uitto and Biswas, 2000; Mohr et al., 2012). Due to the high population density, different and changing forms of land as well as deficient treatment capacities, urban systems do have a faster feedback between potable and wastewater as well as surface and groundwater compared to the rural structures (Azzam et al., 2009; Putra, 2007).

The interaction of urbanization and surface or groundwater quality is considerably controlled by the city's land use structure as different types of land use bear various sources of contaminants and hazards which influence both water quality and quantity (Strohschön et al., 2011; Baier and Strohschön, 2012). Implications for groundwater resources include, for instance, permanent or temporary underground structures, fluctuations in the water level and contamination due to the seeping of different urban pollutants (Morris et al., 2003; Howard, 2004). Functional and spatial relationships between the different types of informality can be observed, and the most important informal processes were classified by Bockhorn et al., 2011 for the Chinese megacity Guangzhou:

- Housing (total housing unit, vertical and horizontal building extension)
- Leisure time and household (well boring, sewage disposal on streets and in rivers, private occupation of public space)
- Work (transport services, workshops, waste collecting, other services)
- Waste disposal (private waste, construction waste, point source waste and debris)

In general it can be stated that the more traditional the area, the higher the degree of informality and that water contamination is higher in units with a high proportion of informality than in units with few informal influences.

The access to drinking water and sanitation is generally better in cities than in rural areas. However, urban areas also exhibit huge disparities regarding the access to qualitatively good water and an adequate disposal of domestic wastewater according to the residential status of the inhabitants (e.g. in fast growing informal settlements or slums or peri-urban areas). Foster et al., 1999 for instance state that using groundwater may reduce pressure upon conventional freshwater supply sources. Therefore, groundwater beneath cities is gaining in importance with complex links to social, legal, economical and political issues. Since lack of good quality water supply often occurs during drought events or in long-term changes over time, the concept of adaptive capacity is closely related. The concept is defined as the difference in the vulnerability under existing conditions and under the less vulnerable condition to which the system could shift (Luers et al., 2003). Different settlement types can have varying adaptive capacities in comparison, both in respect to interrelated water quantity and quality. This fact is yet little investigated by trans-disciplinary research. Even less considered is the more specific question, whether a high heterogeneity of land use types that interact flexibly within one settlement type could raise overall adaptive capacity.

Given the complexity of these highly dynamic processes in mega urban areas, new approaches to knowledge management, modeling and simulation are needed to enable sustainable urban planning and environmental management. Due to this the SiKAMUS project develops new management tools that handle urban structures and the natural environment, along with human activity and their interdependencies in an integrated knowledge representation and simulation framework.

3 Combining Knowledge Management and Simulation – Features and Benefits

One aspect that has so far received little explicit attention in the agent modeling literature is the semantics of modeling languages. Even in systems that strive to improve usability by employing visual modeling languages (cf. North et al., 2007), the particular diagram elements are equivalent to the control flow structures of traditional imperative (i.e. general-purpose) programming languages. This may be a tribute to the sheer breadth of application domains: Macal and North (2010) cite applications of agent-based simulations ranging from biological processes like the immune system, over higher-level systems like the predator-prey relationship, up to complex social dynamics like crowd behavior and economic markets. This extremely wide scope of applications surely makes imperative general-purpose languages a reasonable choice for agent modeling in general, but we believe that this has certain drawbacks in terms of domain adequacy in complex social systems modeling. In particular, general-purpose languages always have to sacrifice conciseness and expressivity for flexibility (cf. Mernik et al., 2005).

The problem domain as laid out in Section 2 poses the most difficult problems to researchers hoping to gain a holistic understanding: An unmanageable complexity of interwoven problems, each of which crosses the boundaries of multiple scientific disciplines.

Both OWL and i^* are languages that can bridge the semantic gap between domain knowledge and formalized computer models built from said knowledge. The communicative artifacts brought about by the Protégé OWL editor and i^* agent diagrams help with the terminology reconciliation required in trans-disciplinary research. OWL concepts can have multiple names, and the same concept can appear in multiple terminological systems. Concept taxonomies never claim to represent a singleton correct ordering. Instead OWL makes it obvious how taxonomies classify concepts according to a certain perspective: Namespaces and ontology imports allow concepts to be used in other taxonomies that follow a completely different paradigm (cf. Bechhofer et al., 2004). Yet OWL reasoners can infer the logical conclusions (i.e. the entailments) of such a multi-paradigm ontology all the same. Integrating data from other sources on the semantic web is just a matter of connecting them to the semantic network, which takes data integration to the next level: Instead of dealing with the technical pitfalls of data structures and formats, we just model the *meaning* of other data in the context of our domain. In Section 4 we will demonstrate how OWL is used to create a semantic geodatabase by formalizing domain knowledge. Ontologies can be published to the semantic web, where they implement the best possible open data infrastructure, according to the five star classification by Berners-Lee, 2009.

 i^* was chosen as a domain-specific language for modeling the intentionality (motivational) and activity structure of urban actor networks. Due to its concise diagram notation that is closely related to the common-sense way of talking about social actors, i^* greatly reduces the effort required for discussing model details with domain experts. At the same time, we gain a lot of flexibility in model development, since all modeling is done at a level of semantic abstraction that can be freely chosen, i.e. domain experts and modelers do not have to deal with technical implementation details. We will elaborate on our use of i^* and its interaction with the OWL ontology in Section 5.

A semantically integrated simulation environment makes the findings generated from model calibration accessible to the process of domain research: Theories developed close to the computational model can be immediately tested on historical data, and their explanatory plausibility can be explored by simulating future scenarios.

4 The OWL World Model

The Web Ontology Language was created as a means to formalizing the meaning of web documents and other data. It was developed to enable the vision of the *Semantic Web* – a vast network of semantically linked information (cf. Berners-Lee et al., 2001).

4.1 Principles of OWL

OWL is designed as a declarative language based on description logic (cf. Nardi and Brachman, 2003), a (mostly) computable subset of predicate logic. The idea is to develop a formalized conceptualization of a certain domain of discourse in the shape of a semantic network. The conceptualization is built around a taxonomy of domain terms, which are further characterized by logical formulas expressing their semantic relations (cf. Borgida and Brachman, 2003). We cannot discuss OWL in all its detail here, but the general idea should be comprehensible by understanding OWL's basic building blocks and the example following in Section 4.2. An OWL ontology is composed of a combination of the following language elements:

- **Individuals** are impartible entities, i.e. things that are indivisible at the chosen level of abstraction. Structurally, they make up the leaves of the taxonomy tree. An example would be a particular, individual person or a specific shop in a certain location.
- **Classes** correspond to the mathematical concept of a set. As such, they can be further characterized by axiomatic formulas defining them as subsets or supersets of each other, or making a general statement about all of their individual members. Structurally, the sub/superclass hierarchy constitutes the domain taxonomy, which also governs inheritance of class properties. An example would be the class *Landuse*, which would be a superclass of the class *Industrial landuse*.
- **Properties** are also called *Roles* in description logic and correspond to the mathematical concept of a binary relation, or a 2-place predicate in predicate logic. A property states a certain relationship between two entities, for example the fact that Bob is married to Alice.
- **Formulas** are logical expressions with description logic semantics. They can combine all the elements explained above into complex axioms using logical operators like quantification, negation, set intersection and so on. Formulas bring a lot of expressive power into the system, but they have to be used with care to keep the computational complexity within manageable bounds.

A knowledge base built from these components can then be used by an OWL reasoner to answer questions regarding everything that logically follows from the knowledge base. When writing down OWL expressions, we use the Manchester syntax, which is an OWL notation that is optimized to be easily readable by non-technical users (cf. Horridge et al., 2006; Horridge and Patel-Schneider, 2009).

In Figure 1, the Protégé ontology editor shows the facts that are known about an individual called railway_work_shop_179. Deduced knowledge is shown with a shaded background and a dashed border. We can see that railway_work_shop_179 is defined to be a Place. The reasoner has determined that it



Figure 1: The Protégé OWL editor showing asserted and inferred facts about the individual railway_work_shop_179.

is also a member of the classes Formal_job and Informal_job. The OWL axioms that lead to classifications like this are discussed in the following section.

4.2 Building a Semantic Geodatabase with OWL

As discussed in Section 2, land use plays a central role in environmental effects and human agency. Thus, land-use units serve as a basic spatial structure in the semantic world model. In this section, we will discuss a small part of a semantic geodata model.

Our example deals with land use and its dispositions for formal and informal economy. Note that this is not meant to be a complete or accurate model of land use or informal processes – for the purposes of this paper, we have to limit ourselves to a very simplified fragment of domain knowledge. We demonstrate how we can construct a minimal semantic geodata model to provide a world view that fits the agent model presented in Section 5.

The spatial basis we use are land-use units identified from a GHMC² development plan, augmented by a manual satellite image analysis. As shown in Figure 2, the land use areas are complemented by the points of interest from Google Maps, which mostly cover small to medium businesses, but also recreational, religious and public institutions. We can map geospatial properties computed by the geodatabase into OWL properties using the -ontopPro- plugin for the Protégé OWL editor (Rodriguez-Muro et al., 2008). This method is called Ontology Based Data Access (OBDA, cf. Rodriguez-Muro et al., 2008). For example, the land use areas as seen in Figure 2 are represented as rows in an SQL table, looking approximately like Table 1 (simplified).

objectid	type	shape
687	informal_housing	[polygon]
467	$super_market$	[polygon]
206	temple	[polygon]
299	tailor	[point]
:	:	÷

Table 1: SQL data excerpt from the *land_use* geodata table. The points were copied from Google Maps.

What we want to show in this example is how to construct a small semantic model that differentiates between formal and informal jobs, which are used by the agent model in Section 5. The only information we have about informality are the informal housing areas shown in Figure 2. From field research and common sense, we hypothesize that informal housing affects the informality conditions of jobs offered in its vicinity. So first, we use the geodatabase to compute a table of all pairs of spatially intersecting shapes. The result will look like Table 2.

Using OBDA mappings with the -ontopPro- Protégé plugin, we automatically convert the data from Tables 1 and 2 into OWL individuals that represent the shapes from the geodatabase. For example, the OWL individual representing the

²The Greater Hyderabad Municipal Corporation is the government body responsible for city planning in Hyderabad, India.



Figure 2: An exemplary map generated from the geodatabase containing land-use units and points of interest (as found in Google Maps, cf. e.g. http://goo.gl/maps/SYwxT, as of 10-2013).

id_1	t_1		id_2	t_2
687	informal_housing		299	tailor
687	$informal_housing$		2329	mosque
660	$formal_housing$		1771	mall
÷	:	÷		:

Table 2: The table lists pairs of intersecting shapes.

point of interest "tailor" with the ID 299 would look like this:

```
Individual: tailor_299
Types:
    Place
Facts:
    has_landuse tailor
    intersects informal_housing_687
```

Here, we have silently introduced the OWL class Place, the trivial definition of which is omitted for conciseness. So far, this is just a translation from the geodatabase to OWL, since no domain knowledge has been added.

What we want to express is the idea that whenever a **Place** has some economic landuse that creates jobs, the informality conditions from the area surrounding it apply also to the jobs it creates. To represent that economic activities "inherit" an

informality inclination from the land use surrounding them, we first need to establish what constitutes our concept of economy. This is done with a simple taxonomy:

```
Class: Landuse
```

Class: Economy SubClassOf: Landuse

```
Individual: tailor
Types: Economy
```

Here we have stated that the land use tailor is a member of the class Economy, which is itself a subclass of Landuse. To allow for a flexible classification of informality, we introduce intermediate land use classes to which we attribute qualitative formality and informality dispositions. Again, trivial classes are defined implicitly:

```
Individual: high_informality
   Types: Informality
Individual: low_informality
   Types: Informality
Class: Informal_landuse
   SubClassOf:
      Landuse
   EquivalentTo:
      has_activity_disposition
```

some Informality

Individual: informal_housing Types: Landuse Facts: has_activity_disposition high_informality

This conceptualization allows us to give arbitrarily finegrained descriptions of the degrees of informality that may be associated with a certain land use. We could even differentiate sectors of informal behavior, as long as we can arrive at some conclusion whether a certain land use is an instance of the Informal_landuse class. Now we can define our concept of an Informal_job:

```
Class: Informal_job
SubClassOf:
Place
EquivalentTo:
has_landuse some Economy
and intersects some (
has_landuse some Informal_landuse
)
```

Now an OWL reasoner will classify tailor_299 as an instance of Informal_job, and it will do the same for all other tailors that intersect with some landuse to which we attribute informality. Varying kinds and degrees of informality could be defined as members of the class Informality, but they will all have the same impact with regard to job conditions within their spatial extension. This should illustrate how we can work with different levels of detail regarding the degree of informality found in a place. However this concept of informal jobs may not be correct for all types of informal activities. For example, we would not expect that a hospital giving a lot of informal treatment influences the informality conditions of a nearby shopping mall. Therefore, we could differentiate several sectors of informality and limit mutual influence to activities that actually imply some sort of economic or social exchange.

5 Intentional Actor Modeling and Simulation

The multitude and complexity of human behavior patterns in densely populated and quickly developing urban areas is a major factor in the excessive challenges posed to urban political and managerial decision making. As pointed out by Zvoleff and An, 2013, agent-based models often suffer from high complexity, which limits their contribution to the domain's scientific discourse. Our intention is to make that complexity manageable by minimizing the semantic gap between domain knowledge and agent models. The i^* language mimics the way humans talk naturally about agents: That is, in terms of intentions, dependencies and actions.

5.1 The i* Dialect used by SiKAMUS

 i^* is a diagram language that originated in process engineering and analysis. Especially in business process re-engineering, there is a need not only to understand *what* is done, but also *why* it is done, *who* is involved as a stakeholder and *how* everybody's goals and intentions interact. i^* was created to be an easily accessible modeling tool for stakeholder relationships (Yu, 1995). For example in early stages of requirements engineering, i^* bridges the gap between natural language descriptions of an organizational environment and implementation oriented modeling formalisms like UML (cf. Yu, 2009). Originally an i^* diagram would be a static description of the intentional i^* graphic notation

explanation



The i^* role is similar to the concept of a social role. In an SR diagram, it can encompass a network of all the elements explained below. An agent playing the role will behave according to the intentionality structure specified by the role.

A **goal** is something the actor wants to achieve. Multiple tasks can be alternative means to that end.

A **task** is some action an actor needs to perform. It can be decomposed into multiple subtasks or subgoals, all of which need to be completed in order to complete the task. In SNet, tasks are given a duration attribute for simulation purposes.

A **softgoal** is a gradual quality that can be associated with a task or a goal. Other goals and tasks can make contributions to it.

The **precondition/effect** element has been newly introduced by the SNet system. It can be connected to tasks and goals. When connected with an effect link, it represents a consequence of completing that task/goal. With a precondition link, the connected task/goal starts only after the precondition has been fulfilled.

A **resource** works similar to a precondition/effect in that it can be provided by completing certain tasks, and be required by other tasks.

Table 3: $SNet/SiKAMUS i^*$ diagram elements

interrelations of a given set of actors. One consequence of this application is that i^* models needed to be easily readable by non-programmers. The graphic representation of an i^* model is designed to convey all the information that is needed to understand the topic, scale, structure and meaning of the model. All language elements are directly related to the natural, commonsense way of talking about people's intentions, goals and interrelations. This way, agent models are not cluttered with technical implementation details that have no meaning in the domain context.

Before we go into the details of agent-based modeling with i^* , some clarification on terminology is in order. The terms 'agent' and 'actor' are used in many disciplines with very specific meanings. Here, we focus on the terminological convention accepted by the i^* community, which takes 'actor' as a superordinate term, while 'agent' and 'role' are regarded as specific types of actors. An agent is a concrete individual which can 'play' certain roles, which means that it behaves according to the behavior specified by these roles. (Yu, 1995, pp. 18–25).

Classic i^* differentiates between Strategic Dependency (SD) and Strategic Rationale (SR) diagrams. Here we deal only with SR diagrams since SD diagrams say nothing about the internal intentionality structure of actors. SD diagrams represent only the macrostructure of the inter-actor network and do not contain enough information to be translated into executable simulation programs (cf. Gans et al., 2006).

The many application fields of i^* brought forth a lot of new diagram elements. To keep things simple, Table 3 explains only

the ones that are common to most dialects and necessary to understand the example in Section 5.

Since its inception, the i^* framework has been adapted to cover many use cases, some generalizing, and some specializing the original idea. The *SiKAMUS* system is based on an i^* dialect that has been described by Gans et al., 2006 to model the evolution of inter-organizational (dis)trust relationships. To accommodate the evolutionary aspect, i^* had to be extended to incorporate parameters that represent an actor's accumulated experiences from previous interactions. An actor network described by an i^* graph would then be evaluated repeatedly, taking into account how previous actions affect trust relationships between actors. This is made possible by translating an i^* actor network into the GOLOG language (see Section 5.3). The result of this translation in SNet is an executable program that ultimately computes which actions an actor will execute given its relations to other actors and its environmental situation.

The modified i^* from SNet and its translation into executable GOLOG programs are extended to support new syntax and semantics, while the trust model implementation is removed since it is considered too specific to the domain of inter-organizational trust relationships. Agents interact with the OWL world model via the precondition/effect elements, which can execute OWL DL queries and updates when triggered.

In our domain, environmental conditions can have a wide influence on actors, especially on softgoals (cf. e.g. the environmental damage effect on the maintain health softgoal in Figure 3). Therefore, we allow precondition/effect elements to make direct contributions to softgoals. Another change from the SNet i^* dialect is how we use precondition/effect elements to define a threshold on a softgoal. If the threshold is crossed, the effect link is triggered, which means the precondition is met. This is needed since we want to be able to express that certain tasks require a softgoal to be fulfilled to a certain level. The SNet framework also removed the resource elements that exist in classical i^* . In the SiKAMUS system, we reintroduce them as simple quantifiable environment conditions that work similarly to precondition/effect elements: A task or a goal can use, consume or provide (a quantity of) a resource. That resource may also be part of the world model, in which case it is spatially shared with other agents and defined in OWL DL queries/updates. Note that all these changes have draft status at best. Some of the syntax and semantics still needs to be formally defined and may even be changed further. i^* constructs may also need to be renamed to better fit with domain semantics.

Figure 3 shows a small exemplary i^* diagram that uses the language elements discussed above. It models (in a very simplified way) the interaction between basic water supply, health and income. In the following explanation, the domain terms from the diagram are written in typewriter font:

There are two roles, the Worker and the Labor Market.³ Agent proactivity starts at the top level goal drinking water. There are two tasks that both work as a means to that end: use tap water and use shallow groundwater. Using shallow groundwater hurts the softgoal maintain health, while the alternative means, use tap water, strengthens it.⁴ Using tap water is a complex task because it requires both access to the water grid and fulfillment of the goal have income. The worker has two alternative means to the end have income: working formally or working informally.⁵ Each of these tasks requires a successful delegation to the Labor Market's respective tasks of providing (in)formal jobs, which in turn depend on the local demand for the respective type of labor. Formal and informal work also differ in their preconditions: Working informally requires a robust health, while the task work formally can be performed regardless of fulfillment state of the maintain health softgoal.⁶ Informal work is further complicated by the environmental damage it causes, since that in turn affects health, which makes it harder to maintain a steady stream of income.

Again, this is not meant to be a complete or correct model of labor markets, job search or environment problems. Instead, we invite the reader to reflect on *the way* the (necessarily incomplete) domain knowledge is expressed in Figure 3 and its description:

- There is no particular reading order, since the diagram signifies a purely declarative structure. Nonetheless it is often advisable to start at the top level goal, but an understanding can be gained just as well by looking from the bottom up.
- The description does two things: It provides an almost homomorphic translation of the diagram into natural language. That is, it is composed almost exclusively of either the domain terms from the diagram, direct translations of *i**'s syntactic elements and some expendable background information in the footnotes.
- There is no explicit representation of decisions. Instead, the choice between alternative means to an end is modeled implicitly by their respective contributions to softgoals.
- All information about an agent's spatial configuration resides in the OWL world model, which is referenced by the precondition/effect elements. The diagram models only roles, which are instantiated as concrete, spatially-bound agents before simulation.

5.2 Agent Instantiation

The role-based actor model shown in Figure 3 is a prototypical abstraction of a behavior pattern. In itself, it does not represent the behavior of any specific individual. A simulation model that aspires to any sort of realism will contain a multitude of such role-based actor models that represent theories of social interaction pertaining to specific aspects of urban life. To run an actual simulation that incorporates the modeled social behavior patterns, we instantiate many concrete, locallybound actors who play certain roles from the diagram. The required information is provided by the OWL world model. In our simple example, we have no data about population density or economic power of certain regions, so the OWL world model will have to make some simplistic assumptions. For example, we could say that one Labor Market role should be instantiated

 $^{^{3}\}mathrm{In}$ a simulation, agents playing these roles would be instantiated in residential and industrious areas, respectively (see Section 5.2).

⁴Here, we consider the situation in Hyderabad, India, where shallow groundwater is easily accessible using (public) hand-pumps, but it is just as easily contaminated due to its high hydro-geologic vulnerability.

 $^{^5\}mathrm{Formal}$ work refers to an actual employment that implies a certain

degree of safety and adherence to regulations, whereas informal work is more of a day-by-day activity, mainly below the radar of regulators, unbound by legal contracts and without security for either the worker or the environment.

⁶This is of course a strong simplification. The idea is that when performing formal work there is usually some kind of legal binding, so that illness does not necessarily cause an immediate loss of income.



Figure 3: A SiKAMUS i* model describing a small aspect of the problem field around income, water supply and health.

for every area that has some economic land use:

```
Class: Job
SubClassOf:
Place,
plays_agent_role Job_Market
EquivalentTo:
Place and
has_landuse some Economy
```

The same thing can be done analogously for the Worker role in all areas that have some kind of residential land use. If we find it too simplistic to instantiate a single Worker agent per residential area, the OWL geodata model could make assumptions on how many agents to instantiate based on the land use type and the area covered by a certain shape. Or, suppose we want to integrate a data source that gives us population densities for the regions we are dealing with, we only have to write another OBDA mapping that relates those data with all the places we already have in the ontology. Then realistically quantifying the number of agents to instantiate takes only a few more OWL rules.

Once we know how many agents of which type we want to instantiate at which locations, we can define their locally-bound parameters. The Labor Market role will need to know whether they have demand for formal labor or demand for informal labor or both. Likewise, each instance of the worker role will have to be instantiated with local parameters determining its access to education, the water grid and hand pumps. In this example, we again have to infer these attributions from the land use type, similar to the way it was done for the Formal_job/Informal_job classification in Section 4.2.

The last thing that is locally bound are softgoal contributions. For example, the finer-grained informality classification shown in Section 4 could be used to define how much the provide informal job task of a labor market instance actually hurts the worker's steady income softgoal. All other contributions to softgoals have to be concretized as inferences from the ontology in the same way.

5.3 Executing the Agent Simulation in Golog

In this section we will shortly describe what happens in the background when an agent simulation is executed in GOLOG. Here we only introduce GOLOG in an informal way. In a real simulation the generated programs contain implementation details that would make them too complex to present here. The reader should note though, that after modeling and instantiating, the concretized model is automatically translated into a GOLOG program.

GOLOG (cf. Levesque et al., 1997) is a programming language for autonomous agents in a dynamic environment. It is based on the Situation Calculus (cf. McCarthy, 1959; McCarthy, 1963; Reiter, 2001), a logical framework used to formalize descriptions of actions and their effects on the surrounding world. The agent's view of the world is described in a so called Basic Action Theory (BAT). A GOLOG program needs such a BAT describing the world and the consequences of actions on it. The generated BAT then contains logical sentences like "there are agents playing the role Worker with the names worker1, worker2, worker3" (as defined by the OWL ontology, cf. Section 5.2). Another example could be "if a Worker wants to use tap water, he or she has to generate income" which is directly translated from the i* model.⁷

⁷Actually, the sentences in the BAT and the GOLOG programs are of course not written in natural English, but in a formal syntax. Explaining that syntax however would not benefit the understanding of

Programs in GOLOG are *complex actions* which are combinations of other actions. Actions which are not complex are called *primitive*. In our example, a primitive action would be "use shallow ground water" because it is not decomposed further. The goal **drinking water** would be translated into a complex action because it needs other actions, for instance "use shallow ground water", to perform the complex action. Note that these examples demonstrate that we translate both goals and tasks into actions in GOLOG.

Let us take a closer look at the goal drinking water in Figure 3. As mentioned in Section 5.1, there are two means to this end: use tap water or use shallow groundwater. Only one of these two means is needed to fulfill the goal drinking water. In this case we do not determine which action the agent should perform. When a program is executed, it will try to find a sequence of actions that allows it to complete the top-level complex action ("drinking water" in this case), given its options expressed in the BAT.

At this point we do not need to go into more detail of the translation. The interested reader is referred to Gans, 2008 for more details on the technical aspects of translating i^* diagrams into GOLOG. The next step after the translation is to start the simulation program. This program coordinates agent interaction and the progress of time. It is also written in GOLOG and is too complex to be discussed here in detail. Therefore, we describe how the simulation works in an informal way.

In the simulation environment, time is divided into discrete time slots of an arbitrary unit. Each primitive action is given a duration of a certain amount of such time slots. The duration of a complex action depends on the combined duration of all primitive actions that have to be performed to complete the complex action.

To fulfill the top-level goal drinking water, a worker agent has to plan a course of action with the best contribution to its softgoals. That is, a worker agent would first try to perform the action "use tap water" because this has a better contribution to the softgoal maintain health than "use shallow groundwater". However using tap water requires the resource water grid to be available and the complex action "have income" to be completed. To that end, the worker agent needs to delegate one of the tasks work informally or work formally to an agent playing the role of the labor market. Therefore, the worker agent requests offers regarding the delegatable tasks from all labor market agents. It will then choose to delegate to the particular labor market agent which promised the best contribution to the steady income softgoal and the shortest duration for the delegated task. When deciding between the "work informally" or "work formally" tasks, the worker agent will also have to consider the fact that informal work causes a certain amount of environmental damage and the immediate effect of that on the maintain health softgoal. Since environmental damage is a spatially shared property stored in the OWL world model, it might only hurt the agent's health after a lot of damage has accumulated, possibly through the actions of other agents. Therefore, the agent is likely to consider the "work informally" task even though it is known to have a possible adverse effect on health that might even prevent it from working informally in the future.

After an agent has completed its top-level action, it will try to do so again in the next time slot. However in the next time slot, the environmental conditions may have changed in the OWL world model, possibly through the agent's own actions, or due to the actions of other agents. Thus, the agent might not perform the action in the same way as before. For example, all shallow groundwater might be used up, so handpumps are not an option any more.

During a simulation, we can observe, protocol, and later analyze the environmental interaction of all agents, for example the environmental damage effect triggered by informal work, or the shifting of formal/informal job allocation. We could also introduce policy agents that change e.g. the availability of education in certain areas and see how that affects informal work dynamics.

6 Conclusion and Future Work

At the time of writing, the SiKAMUS i^* dialect and its translation to GOLOG are neither fully specified nor completely implemented, in contrast to the SNet framework which has been extensively described by Gans et al., 2006, Gans, 2008 and Schmitz, 2010. The domain adequacy of our dialect will have to be tested by incremental prototyping, with results being published as they mature, especially with regard to the changes described in Section 5.

The instantiation described in Section 5.2 is also not completely implemented yet. In particular, we have to work out which instantiation parameters have to be declared statically and which ones have to be read from the world model each time they are used. The performance and scalability impacts of these choices have to be measured, too.

As the *SiKAMUS* framework matures, so will the domain models and data sources. It needs to be seen how well the approach scales with agent instances increasing in number and complexity.

In the OWL world model, reasoning with spatial relations like **intersects** (cf. Section 4) is only practical if individuals only have a few such relations. If e.g. we wanted to reason with distances, it would be impractical to instantiate a numeric literal for each distance that exists between any two places since that would create a number of numeric literals equal to the square of the number of places. For these cases, the geometry of places will have to be stored in the OWL ontology using the Well-Known Text format and the OWL reasoner will have to support spatial queries according to the OpenGIS architecture (both cf. Open Geospatial Consortium Inc., 2011). In this context, the work done by the *STRABON* and *Parliament* projects will be a valuable resource (see Kyzirakos et al., 2012; Battle and Kolas, 2012, respectively).

References

Anuradha, M. and Cunha, D. da (2009). Soak-Mumbai in an Estuary. New Delhi: upa & Co.

- Azzam, R., Baier, K., Neukum, C., and Post, C. (2009). "Sauberes Trinkwasser f
 ür die Zukunft - Grundwasserschutz im Spannungsfeld zwischen Global Change und wachsendem Ressourcenbedarf". In: *RWTH Themenheft* 2009.1, 72–73.
- Baier, K. and Strohschön, R. (2012). "Die Grundwassersituation in der südchinesischen Megastadt Guangzhou". In: DVGW energie / wasser-praxis 4, 18–23.
- Baier, K., Real, A., Strohschön, R., and Azzam, R. (2009). "Ansätze eines alternativen Wassermanagementkonzepts für mega urbane Räume am Beispiel der südchinesischen Megastadt Guangzhou". In: uwf UmweltWirtschaftsForum 17.3, 281–289.
- Battle, R. and Kolas, D. (2012). "Enabling the geospatial semantic web with Parliament and GeoSPARQL". In: Semantic Web 3.4, 355-370. URL: http://www.semantic-webjournal.net/sites/default/files/swj176_2.pdf.

the general idea, so we stick with equivalent expressions in natural English.

- Bechhofer, S., Harmelen, F. van, Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., and Stein, L. A. (2004). OWL Web Ontology Language Reference. Tech. rep. URL: http://www.w3.org/TR/2004/REC-owl-ref-20040210.
- Berners-Lee, T. (2009). Design Issues Linked Data. URL: http://www.w3.org/DesignIssues/LinkedData.html.
- Berners-Lee, T., Hendler, J., Lassila, O., et al. (2001). "The semantic web". In: Scientific american 284.5, 28-37. URL: http://isel2918929391.googlecode.com/svn-history/ r347/trunk/RPC/Slides/p01_theSemanticWeb.pdf.
- Bockhorn, B., Baier, K., Strohschön, R., and Azzam, R. (2011).
 "Informelle Prozesse und deren Auswirkung auf die Wasserressourcen in der südchinesischen Megastadt Guangzhou".
 In: Aachener Studien zur Sozial-Ökonomischen Entwicklungsforschung. Ed. by R. Strohschön, M. Romich, and K. Baier, 25–36.
- Borgida, A. and Brachman, R. J. (2003). "Conceptual Modeling with Description Logics". In: *The Description Logic Handbook – Theory, Implementation and Applications*. Ed. by F. Baader, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider. Cambridge University Press.
- Bronger, D. (1996). "Megastädte". In: Geographische Rundschau 2, 74–81.
- Chen, N. and Heligman, L. (1994). "Growth of the World's Megalopolises". In: *Mega-city Growth and the Future*. Ed. by R. Fuchs, E. Brennan, J. Chamie, F.-C. Lo, and J. Uitto, 242–261.
- Feldbauer, P. and Parnreiter (1997). "Einleitung: Megastädte – Weltstädte – Global Cities". In: Mega-Cities. Die Metropolen des Südens zwischen Globalisierung und Fragmentierung. Ed. by P. Feldbauer, K. Husa, E. Pilz, and I. Stacher, 9–20.
- Foster, S., Morris, B., Lawrence, A., and Chilton, P. (1999). "Groundwater impacts and issues in developing cities: an introductory review". In: *Groundwater in urban environment: selected city profiles.* Ed. by P. Chilton, 3–16.
- Fuchs, R., Brennan, E., Chamie, J., Lo, F.-C., and Uitto, J. (1994). Mega-city Growth and the Future. Tokyo.
- Gans, G., Lakemeyer, G., Jarke, M., and Vits, T. (2006). "SNet: A modeling and simulation environment for agent networks based on i* and ConGolog". In: *Advanced Information Systems Engineering*. Springer, 328–343. URL: http://helios. mm.di.uoa.gr/~rouvas/ssi/caise2002/23480328.pdf.
- Gans, G. (2008). "An agent-based modeling- and simulationmethodology for strategic inter-organizational networks". PhD thesis. Fakultät für Mathematik, Informatik und Naturwissenschaften.
- Giddens, A. (1984). The constitution of society. Outline of the theory of structuration. University of California Press.
- Heinrichs, D. (2009). "Megacity mit Zukunft gesucht!" In: Politische Ökologie 27, 44–46.
- Henderson, V (2002). "Urbanization in Developing Countries". In: The World Bank Research Observer 17.10, 89–112.
- Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P. F., and Rudolph, S. (2012). OWL 2 Web Ontology Language Primer. URL: http://www.w3.org/TR/2012/REC-owl2primer-20121211.
- Horridge, M. and Patel-Schneider, P. F. (2009). OWL 2 web ontology language manchester syntax. URL: http://www.w3. org/TR/2012/NOTE-owl2-manchester-syntax-20121211.
- Horridge, M., Drummond, N., Goodwin, J., Rector, A., and Wang, H. H. (2006). "The manchester OWL syntax". In: Proc. of the 2006 OWL Experiences and Directions Workshop (OWL-ED2006). URL: http://sunsite.informatik. rwth-aachen.de/Publications/CEUR-WS//Vol-216/ submission_9.pdf.

- Howard, K., ed. (2004). Urban groundwater meeting the challenge. Selected papers from 32nd International Geological Congress.
- Knublauch, H., Fergerson, R., Noy, N., and Musen, M. (2004).
 "The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications". In: *The Semantic Web ISWC 2004*. Ed. by S. McIlraith, D. Plexousakis, and F. Harmelen. Vol. 3298. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 229–243. ISBN: 978-3-540-23798-3. DOI: 10.1007/978-3-540-30475-3_17. URL: http://dx.doi.org/10.1007/978-3-540-30475-3_17.
- Kraas, F. and Nitschke, U. (2006). "Megastädte als Motoren des globalen Wandels". In: *Internationale Politik*, 18–28.
- Kraas, F. and Nitschke, U. (2008). "Megaurbanisierung in Asien. Entwicklungsprozesse und Konsequenzen stadträumlicher Reorganisation". In: Informationen zur Raumentwicklung 8, 447–456.
- Kyzirakos, K., Karpathiotakis, M., and Koubarakis, M. (2012). "Strabon: A semantic geospatial DBMS". In: *The Semantic Web-ISWC 2012.* Springer, 295-311. URL: http://cgi.di.uoa.gr/~koubarak/publications/iswc-strabon.pdf.
- Levesque, H. J., Reiter, R., Lespérance, Y., Lin, F., and Scherl, R. B. (1997). "GOLOG: A logic programming language for dynamic domains". In: *The Journal of Logic Programming* 31.1-3, 59–83.
- Luers, A., Lobell, D., Sklar, L., Addams, C., and Matson, P. (2003). "A Method for quantifying vulnerability, applied to the agricultural system of the Yaqui Valley, Mexico". In: *Global Environmental Change* 13, 255–267.
- Macal, C. M. and North, M. J. (2010). "Tutorial on agent-based modelling and simulation". In: *Journal of Simulation* 4.3, 151–162. DOI: 10.1057/jos.2010.3.
- McCarthy, J. (1959). "Programs with common sense". In: Proceedings og the Symposium on the Mechanisation of Thought Processes, 77–84.
- McCarthy, J. (1963). Situations, Actions and Causal Laws.
- Mernik, M., Heering, J., and Sloane, A. M. (2005). "When and how to develop domain-specific languages". In: ACM Comput. Surv. 37.4, 316–344. ISSN: 0360-0300. DOI: 10.1145/ 1118890.1118892. URL: http://doi.acm.org/10.1145/ 1118890.1118892.
- Mertins, G. (1992). "Urbanisierung, Metropolisierung und Megastädte. Ursachen der Stadt"explosion" in der Dritten Welt – sozioökonomische und ökologische Problematik." In: *Dokumentationen, Informationen, Meinungen* 44. Ed. by Deutsche Gesellschaft für die Vereinten Nationen, 7–21.
- Ministry of Housing and Urban Poverty Alleviation (2009). Indian urban poverty report.
- Mohr, J., Baier, K., Jha, R., and Azzam, R. (2012). "Urbanization and the social impact on water resources in India". In: Fourth International Conference on Water Resources and Renewable Energy Development in Asia.
- Morris, B., Lawrence, A., Chilton, P., Adams, B., Calow, R., and Klinck, B. (2003). "Groundwater and its susceptibility to degradation: A global assessment of the problem and options for management". In: UN Environment Programme RS.03.3.
- Nardi, D. and Brachman, R. J. (2003). "An Introduction to Description Logics". In: *The Description Logic Handbook* – *Theory, Implementation and Applications*. Ed. by F. Baader, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider. Cambridge University Press.
- North, M. J., Tatara, E., Collier, N. T., Ozik, J., et al. (2007). Visual agent-based model development with repast simphony. Tech. rep. Argonne National Laboratory. URL: https://cc. ist.psu.edu/BRIMS2013/archives/2008/Tutorials/08-BRIMS-001.pdf.

- Open Geospatial Consortium Inc. (2011). OpenGIS® Implementation Standard for Geographic information – Simple feature access – Part 1: Common Architecture. Ed. by J. R. Herring. Open Geospatial Consortium Inc. URL: http: //www.opengeospatial.org/standards/sfa.
- Pelling, M. (2003). "The Vulnerability of Cities Natural Disasters and Social Resilience". In: *Earthscan*.
- Putra, D. (2007). "The Impact of Urbanization on Groundwater Quality. A Case Study in Yogyakarta City – Indonesia". In: Mitteilungen zur Ingenieurgeologie und Hydrogeologie 96.
- Reiter, R. (2001). Knowledge in Action: Logical Foundations for Describing and Implementing Dynamical Systems. MIT Press.
- Rodriguez-Muro, M., Lubyte, L., and Calvanese, D. (2008). "Realizing Ontology Based Data Access: A plug-in for protégé". In: *ICDEW 2008. IEEE 24th International Conference on Data Engineering Workshop, 2008*, 286–289. DOI: 10.1109/ICDEW.2008.4498333.
- Schmitz, D. (2010). "Managing dynamic requirements knowledge: an agent-based approach". PhD thesis. RWTH Aachen. URL: http://darwin.bth.rwth-aachen.de/ opus3/volltexte/2011/3656.
- Shannon, K., ed. (2008). *Water Urbanisms*. Amsterdam: Sun publishers.
- Strohschön, R., M. Romich, and K. Baier, eds. (2009). Strukturen, Prozesse und Dynamiken der Mega-Urbanisierung in China – Landnutzung und Wasserressourcen. Aachener Studien zur Sozial-Ökonomischen Entwicklungsforschung. Vol. 11. Aachen: Shaker.
- Strohschön, R., Azzam, R., and Baier, K. (2011). "Mega-Urbanization in Guangzhou - Effects on Water Quality and Risks to Human Health". In: *Health in Megacities and Urban* Areas. Ed. by A. Krämer, M. Khan, and F. Kraas, 221–229.
- Strohschön, R., Wiethoff, K., Baier, K., Lu, L., Bercht, A., Wehrhahn, R., and Azzam, R. (2012). "Land use and water quality in Guangzhou, China: A survey of ecological and social vulnerability in four urban units of the rapidly developing megacity". In: Int'l Journal of Environmental Research.
- The World Bank (2011). Urbanization in India Integral part of economic growth. January 31, 2012. URL: http://go. worldbank.org/1UKKDONKTO.
- UN (1987). "Population Growth and Policies in Megacities: Bangkok". In: *Population Policy Paper* 10.
- Uitto, J. and Biswas, A. (2000). Water for Urban Areas: Challenges and Perspectives. Tokyo: United Nations University Press.
- Wehrhahn, R., Bercht, A., Krause, C., Azzam, R., Kluge, C., Strohschön, R., Wiethoff, K., and Baier, K. (2008). "Urban restructuring and social and water-related vulnerability in mega-cities – the example of the urban village of Xincún, Guangzhou (China)". In: *Die Erde* 139.3, 227–249.
- Werlen, B. (2000). Sozialgeographie: Eine Einführung.
- Wiethoff, K., Baier, K., and Strohschön, R. (2011). "Brachflächen: urbane Leerstellen der Megastadt. Auswirkungen der Interimsnutzung von Brachflächen auf die lokale hydrologische Situation in Guangzhou, China". In: Cybergeo 524. URL: http://cybergeo.revues.org/23537.
- Xu, L., Ge, J., and Hua, Y. (2002). "Analysis on the environmental problems accompanying the high-speed urbanization of small towns in China". In: ournal of Environmental Sciences 14.1, 120–126.
- Yu, E. S. (1995). "Modelling Strategic Relationships for Process Reengineering". PhD thesis. University of Toronto. URL: ftp://ftp.db.toronto.edu/pub/eric/DKBS-TR-94-6.pdf.
- Yu, E. S. (2009). "Social Modeling and i*". In: Conceptual Modeling: Foundations and Applications. Ed. by A. T. Borgida,

V. K. Chaudhri, P. Giorgini, and E. S. Yu. Vol. 5600. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 99–121. ISBN: 978-3-642-02462-7. DOI: 10.1007/978-3-642-02463-4_7.

Zvoleff, A. and An, L. (2013). "Analyzing Human–Landscape Interactions: Tools That Integrate". In: *Environmental Management*. ISSN: 1432-1009. DOI: 10.1007/s00267-012-0009-1.