

Application of Model Driven Techniques for Agent-Based Simulation

Rubén Fuentes-Fernández¹, José M. Galán², Samer Hassan¹,
Adolfo López-Paredes³, Juan Pavón¹

¹ GRASIA, Universidad Complutense de Madrid, Spain
{samer, ruben, jpavon}@fdi.ucm.es

² INSISOC, Universidad de Burgos, Spain
jmgalan@ubu.es

³ INSISOC, Universidad de Valladolid, Spain
adolfo@insisoc.org

Abstract. Agent-based simulation is being recognized as a useful tool for the study of social systems. It is based on the idea that agents can be used as a good abstraction of members of a society, and by simulating their interactions, observe the emergent behavior. Usually, agent based models for social simulation are rather simple, but there are more and more works that try to apply this technique for rather complex systems, both in the typology of the agents and their relationships, and in scalability. For this reason, software engineering techniques are required that would facilitate the development of agent-based simulation systems. In this sense, it can be useful to consider agent-oriented methodologies, as they cope with these requirements. This work explores the application of a model-driven methodology for the development of Multi-Agent Systems, INGENIAS. There are several advantages of this approach. One is the possibility to define specific modeling languages that are conceptually close to the domain-expert, in this case for simulation of social systems. This will facilitate the communication of multidisciplinary teams and, by using generating code from models, the social scientist is alleviated from the concerns of programming for a simulation platform. This is illustrated with a case study of urban dynamics.

1 Introduction

Agent based simulation (ABS) is being recognized as a useful tool for the study of social systems, especially in Social Sciences [1]. One of the main advantages of modeling a social system with agents, and in our opinion the one that distinguishes it from other paradigms, is that it facilitates a direct correspondence between the entities in the target system, and the parts of the computational model that represent them (i.e. the agents) [2].

The process of abstraction to transform the real target system into a model for simulation is complex. This process involves different subtasks and roles, which need

diverse backgrounds and competences in the design, implementation and use of an archetypical agent based simulation. Drogoul et al. [3] identify three different roles in the modeling process: the *thematician*, the *modeler*, and the *computer scientist*. This classification has been expanded by Galán et al. [4] using the framework proposed by Edmonds [5], including an additional role, the *programmer*.

The role of the thematician, which ideally would be performed by an expert in the domain, aims at producing the first conceptualization of the target system. This task entails: defining the objectives and the purpose of the modeling exercise, identifying the relevant components of the system and the relations between them, and describing the most important causal dependencies. The modeler's job is to produce formal requirements for the models starting from the thematician's ideas. These requirements allow the computer scientist to formulate a feasible model that can run in a computer. However, not all the formal specifications can be directly implemented in a computer. The computer scientist role finds a suitable approximation to the modeler's formal model that can be executed in a computational system with the available technology. Finally, the programmer's role is to implement the computer scientist's model to a target simulation platform.

In practical terms, modeling in Social Sciences faces two problems. In individual developments, it is difficult that one person has all the required expertise; in teams, there are communication problems.

The first problem appears when the same person plays all the roles of the process [4, 5]. Minar et al. [6] explain its negative consequences as follows: "*Unfortunately, computer modeling frequently turns good scientists into bad programmers. Most scientists are not trained as software engineers. As a consequence, many home-grown computational experimental tools are (from a software engineering perspective) poorly designed*". Besides, scientists may find difficult understanding the detailed behavior of the underlying software, since doing it would imply a full understanding of its implementation.

On the other hand, many problems require a multidisciplinary perspective, involving members with specialized roles. The second problem arises because effective communication between experts of fundamentally different domains (e.g. sociology and computer science) is not trivial. In most cases, it is difficult to grasp how the social features have been mapped to program constructions. Thus, there are difficulties to assure that the program really implements its conceptual model.

To address these problems, our research promotes the creation of a set of high-level tools, methods and languages, to assist the transfer of models between different roles in the modeling process. These tools should work with modeling languages that include, ideally, concepts close to the thematician's background, but at the same time representing ideas from a software engineering point of view. We must take into account that any mismatch between the specifications and the actual model passed to the next stage, will end up producing an error.

Moreover, a high-level communication tool may also help in the validation. Model validation is the process of determining that the model behavior represents the real system to satisfactory levels of confidence and accuracy, which are determined by the intended model application and its application domain. When dealing with complex systems, as it is frequent on ABS, the traditional methods used in model validation are not widely accepted [7]. In such cases, a good option for the validation of the

conceptual model is to check whether the theoretical foundations and assumptions are reasonable within the context of the objectives of the simulation. This structural validation is sometimes performed on the basis of participatory methods with experts in the modeled domain and stakeholders [8]. Again, these expert panels do not usually have a software engineering background. Intermediate languages between the different roles endowed with a high-level descriptiveness facilitate the communication, modification and criticism of the models in the validation stages.

This paper shows how the application of metamodeling and transformation techniques can facilitate the definition of specific modeling languages for agent-based social simulation, taking as starting point the INGENIAS agent-oriented methodology, which provides tools for agent-based metamodel definition, model transformation, and code generation. The role of metamodels and how to use them for agent-based social simulation modeling is introduced in the next section. Subsequently, the methodological process is illustrated step-by-step by means of a case study on urban dynamics. The paper ends with a discussion and concluding remarks concerning the approach.

2 Metamodels for agent-based social modeling and simulation

An approach for addressing the mentioned problems is the use of domain-specific languages (DSL) [9] to produce intermediate models between the thematician's abstract non-formal models and the final program, but close to the thematician's view. A DSL explicitly defines its concepts, attributes, relationships, and potential constraints applicable to the models specified with it. These elements are "domain-specific" because they are extracted from the target domain. Thus, researchers working with a DSL specify their models using the language of their discipline. Moreover, since the language elements are clearly defined, mapping them to software constructions is significantly easier and more reproducible than in the case of an arbitrary set of elements, which is the situation with non-formal descriptions.

Metamodels [10] allow the specification of DSL. Such specification corresponds to a graph with nodes (i.e. entities or concepts) linked by arcs (i.e. relationships or references), and both of them can have properties (i.e. features or attributes). A metamodel for a given DSL defines the types of nodes and arcs that correspond to the domain concepts. It indicates their names and attributes, and the rules and constraints that they satisfy, for instance the concepts a relationship can link or its cardinality.

Metamodels have two key advantages for defining DSL. First, they can be extended to satisfy specific modeling needs: if the current form of the DSL is not enough to model a given problem, new elements can be introduced as extensions or specializations of the existing ones. These new elements provide an accurate representation of the domain notions according to the definition of the thematician. Second, there is a wide range of support tools to work with them. This facilitates the production of modeling tools for specific domains.

This work starts from an agent-oriented framework, INGENIAS [11], which already provides tools for model driven development. INGENIAS modeling language supports well the specification of organization structure and dynamics, as well as

agent intentional behavior, characteristics that are present in social systems. This is a feature that general purpose modeling languages lack. This language is supported by the INGENIAS Development Kit (IDK) with a graphical editor, which can be extended to introduce new modeling concepts. Second, INGENIAS model-driven engineering approach facilitates the independence of the modeling language with respect to the implementation platform. This is especially important here in order to abstract away programming details and concentrate on modeling and analysis of social patterns. With this purpose the IDK supports the definition of transformations between models and code for a range of implementation platforms.

Table 1 summarizes the main concepts of INGENIAS used in the rest of the paper. Note that this table considers concepts but not their relationships. The relationships used in the case study have self-explanatory names. For instance, the relationship “*WFProduces*” from a “*Task*” to a “*Frame Fact*”. The first two letters indicate the main type of INGENIAS diagram where the type of relationship appears. In this case, “*WF*” stands for “*WorkFlow*” diagram. The rest of the name provides the meaning of the relationship. “*Produces*” shows that the fact is the result of the execution of the task.

Table 1. Main concepts of the INGENIAS modeling language.

Concept	Meaning	Icons
Agent	An active concept with explicit goals that is able to initiate some actions involving other elements of the simulation.	
Role	A role groups related goals and tasks. An agent playing a role acquires the goals and tasks of such role.	
Environment application	An element of the environment. Agents act on the environment using its actions and receive information through its events.	
Goal	An objective of an agent. Agents try to satisfy their goals executing tasks. The satisfaction or failure of a goal depends on the presence or absence of some elements (i.e. frame facts and events) in the society or the environment.	
Task	A capability of an agent. In order to execute a task, certain elements (i.e. frame facts and events) must be available. The execution produces/consumes some elements as result.	
Frame Fact	An element produced by a task, and therefore by the agents.	
Event	An element produced by an environment application.	
Interaction	Any kind of social activity involving several agents.	
Group	A set of agents that share some common goals and the applications they have access to.	
Society	A set of agents, applications and groups, along with some general rules that govern the agent and group behavior.	

It is possible to define an agent based social simulation modeling language by extending the INGENIAS metamodel in a process with the following activities:

1. *Domain analysis.* Thematicians consider the concepts that are required to express their hypotheses and the related information in the group or society.
2. *Determine interactive concepts.* Among the domain concepts, some of them focus on the analysis, and thus they are considered as decision makers that follow certain rationality. Besides, some concepts represent elements that initiate interactions with others, for instance asking for some services. All these concepts are candidate agents in the ABS and therefore they are represented as subclasses of Agent. Any element that engages in interactions with other elements of the system should be modeled also as an agent.
3. *Determine non-interactive concepts.* Passive elements that do not take decisions are regarded as part of the environment. In INGENIAS, these elements are modeled as subclasses of the environment application.
4. *Determine specialization hierarchies between concepts.* The elements introduced for a problem usually share some features. In order to highlight the common aspects of elements and encourage their reusability, these new elements are arranged in inheritance hierarchies. A super-concept contains all the elements/attributes and participates in all the relationships common to its sub-concepts. Sub-concepts only modify their own specific features, constraining or adding some features of the super-concept.
5. *Determine groups and societies.* In case that several agents share common global goals or environment applications, they can be gathered in groups. If they also share common rules of behavior, they constitute a society.
6. *Determine interactions.* To carry out the activities of the system, agents act on environment applications, receive information from them, and communicate with other agents exchanging elements of the system. A group of interconnected activities aimed at satisfying a global goal constitutes an interaction.
7. *Assign roles, objectives and capabilities to agents.* INGENIAS refines agent definitions with the tasks they are able to do, which correspond to their capabilities and goals. A goal is linked to the tasks which are able to satisfy it. These tasks produce some elements (i.e. frame facts). The presence or absence of some frame facts and events (produced by environment applications) are evidence of the satisfaction or failure in the achievement of the goal.
8. *Refine interactions.* A refined interaction indicates the agents and environment applications that participate in it, the tasks that agents execute, the goals they pursue with those executions, and the elements produced and consumed in it.
9. *Validate ABS.* These agent-based models are a refinement of the thematician's non-formal models. They are expected to represent it accurately, while providing additional details that facilitate the transition to the running system.

These activities do not need to be performed sequentially. For instance, if there is a precise idea of the existing interactions, modelers can begin with activity 5 and then use this information to discover the agents and environment applications. Besides, this presentation of the process is necessarily simplified given the space limitations of the paper, but more detailed activities are required to provide a full modeling guideline. Also, note that this process can be used to describe models or metamodels. If the

concepts have a wide application for a domain, they become part of a metamodel; if they are specific for a problem, they remain at the model level.

This process can be regarded as iterative. That is, a given model can be the abstract model for a new and more platform-oriented model. This allows a transition from the abstract non-formal description to the program in several steps, improving the traceability of the process. This approach is in the basis of model driven engineering, and INGENIAS fully implements it. The advantage of this decomposition is the possibility of using specific guidelines and support tools for each step, which crystallize and automate the expertise of thematicians and modelers and helps novel researchers.

3 Case Study: Urban Dynamics in Valladolid

In order to illustrate the usefulness of metamodeling with INGENIAS in social simulation contexts, an urban dynamic model has been selected. Several complementary theories from fields as Sociology, Geography, Political Science or Economics attempt to explain the complex problem of the dynamic spatial occupation [12]. One of the most descriptive models of urban dynamics applied to real systems is the Yaffo-Tel Aviv model developed by Benenson et al. [13]. It considers the stress-resistance hypothesis [14] as the force that explains intra-population movements in cities. This model has been adapted to the Valladolid metropolitan area (Spain) and, together with other socioeconomic models, is used for exploring the dynamics of urban phenomena [15]. In order to adapt it to this new context and develop the different layers of the model, extensive discussions with domain experts (thematicians) were needed. The description presented in this section is a result of the intensive communication between modelers and those thematicians.

The model comprises two different layers. The first layer is retrieved from a vectorial GIS. This GIS explicitly represents every block with households in the studied geographical region and characterizes them by their spatial and socioeconomic characteristics. In the second layer, the computational agents representing the families that live in the area are spatially situated.

The main assumption of the model is that agent's selection of residence is influenced by intrinsic features of the candidate households and by the similarity of the agents' socioeconomic factors with those of their neighborhoods. Thus, some of the rules of behavior of the agents in this model are based on the concept of neighborhood, defined taking into account the centroids of their blocks and the Voronoi tessellation. The variable "residential dissonance" quantifies the dissimilarity between an agent, its neighborhood and its household. The probability that an agent leaves a residence is considered proportional to such residential dissonance. This variable may be influenced by differences in terms of nationality or education level, or by imbalances between an agent's wealth and the value of the house where it lives.

Once each agent has calculated the dissonance, the opportunity to change its current residence is modeled through a stochastic process that transforms dissonance into probability of change. Those selected agents are included in a set M of potential internal migrants. If external immigration is enabled, immigrant agents are also

included into the set M . Subsequently, each agent A in M estimates the attractiveness (one minus the dissonance) of a number of candidate empty households H_A .

The last step of the algorithm entails the assignment of agents in set M to empty households. Each potential migrant chooses the household found with highest attractiveness. If the dwelling is already empty, the agent occupies it with a probability depending on its attractiveness; if it is not, the agent removes the household from its H_A list. The process is iterative until H_A is empty for all the agents in M . In each iteration, the order of the agents is randomly selected to avoid bias in the agent selection. The agents that have not been able to find a suitable house leave the city with probability L_A and remain in its current house with probability $1-L_A$. The immigrant population without household leaves the city as well.

Another submodel dynamically updates the prices of the households [16]. The value of a household depends on the wealth of its family and neighbor families, and the value of the surrounding empty houses. The value of the empty households decreases at a constant rate considered as an exogenous parameter of the model.

4 Following the Process

The development of the model for the case study following the previous process is as follows. Activity 1 corresponds to the descriptive analysis realized by the thematician before the discussions which resulted on the previous subsection, already from the modeler's point of view.

Activities 2 and 3 are made in parallel to identify the entities in the system. In this case, there are at least two active elements, the families and the households. Families initiate interactions to keep low their dissonance level; households perform an interaction with other households to update their value. Thus, both elements are modeled as two different classes of agents.

According to the specification of the problem, when a family wants to migrate, it receives a list of potential candidate households for migration, chosen randomly all over the city. However, communication constraints only allow a family to communicate directly with their neighbors, not all over the city. In order to allow families to get this information about other parts of the city, we introduce a *map* environment application, which is a non-interactive concept as defined in Activity 3. The map is part of the environment and it provides methods to get the list of empty households and retrieve the house data (such as if it is empty).

Activities 4 and 5 are intended to create hierarchies and groups of elements that share features. At this step, there are only two types of agents and one environment application, so the hierarchies are flat. In order to represent the organization of the city, a society is introduced, together with three groups that can be seen in Fig. 1. The *city* society comprehends all the participants in the problem. They share rules of collaborative (e.g. all the agents are willing to provide the requested information) and non-violent behavior (e.g. no family can occupy the household of other family, and no household can eject its family). The three groups in this society are *families*, *households*, and *neighborhoods*. The neighborhoods include the households and the families that are able to communicate among them. This information is used to make

explicit the influence of the neighborhood in the calculation of the dissonance level and the price of households. Note also that this information is useful in the implementation of the model. As Repast (<http://repast.sourceforge.net/>) was chosen as target platform, an agent can only communicate with those placed in adjacent cells. Thus, the neighborhoods tell us which agents must be in which cells. This is a typical example of information with social basis refined to guide the implementation.

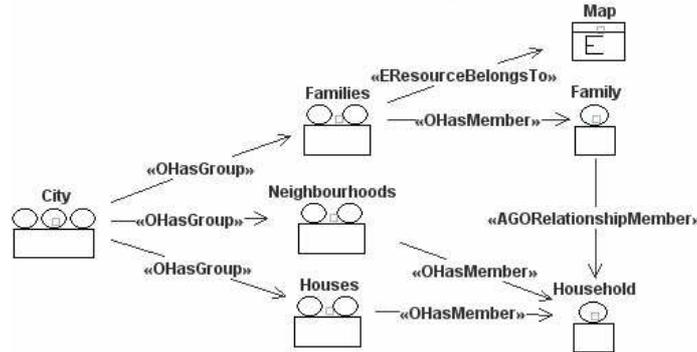


Fig. 1. The city society and its components. For the legend see Table 1.

Activity 6 identifies the relevant interactions for the problem: the calculation of the household prices and the family dissonance levels. Due to space restrictions, the following diagrams focused on the second interaction.

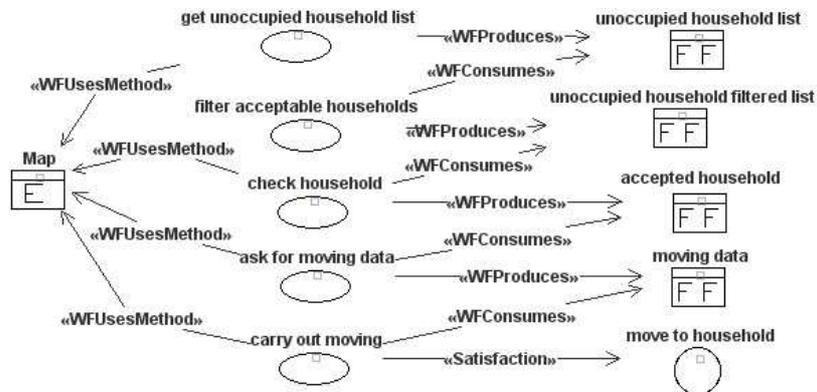


Fig. 2. Migrate in the city workflow for the family agent. For the legend see Table 1.

Activities 7 and 8 identify the tasks and goals of agents, the elements exchanged and their relations to specify the interactions. Fig. 2 and Fig. 3 include part of the results of these activities. They focus on the workflow the family agent performs after it finds out that its dissonance level is too high with respect to its neighborhood.

Fig. 2 shows the part of the workflow where the family agent looks for a suitable and unoccupied household. It begins asking for H unoccupied households using the map. Afterwards, it filters the list to get only the households whose dissonance level is below its threshold. Then, the family uses the map to try to get one of the suitable

households. Note that according to the specification of the non-formal description, all the families that are uncomfortable with their households try to make these tasks at the same time. First, all of them get the list, then they filter it, and afterwards each of them tries to get into a household. Given this order, it is possible that when a family tries to occupy an initially unoccupied household, another discontent family that chose first has already occupied it. For this reason, the task *check household* can fail.

Fig. 3 shows the interaction where the family agent asks the target household agent to accept it. All the involved tasks need to use the map for communication, as it appears in.

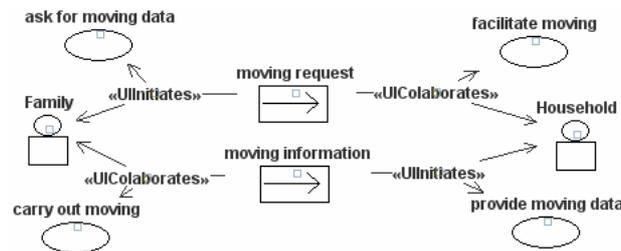


Fig. 3. End of the *migrate in the city* workflow with the interaction between the *family* agent and the target *household* agent. For the legend see Table 1.

5 Discussion and Concluding Remarks

The introduction of a specific modeling language, which is founded on well established agent concepts and close to the domain-expert in the form of understandable diagrams, facilitates communication, specification, implementation and validation of agent-based models for the simulation of social systems. This has been validated by providing guidelines for agent-based modeling using the support of the default modeling tools of a specific agent-oriented modeling language, INGENIAS. And it has been illustrated with a case study on urban dynamics.

This framework will allow the specification of social systems with a graphical modeling language, the simulation of the models of these systems by exploiting the capabilities of existing agent-based simulation tools/platforms, and the identification and analysis of social patterns (at a macroscopic, or aggregate, level) in terms of the atomic elements of the social system specification (at a microscopic, or individual/interaction, level). The advantages go further than usability. As it has been discussed in [17], this solution facilitates the replication of an experiment on different simulation engines, in order to contrast results. The availability of a graphical view of the system facilitates its understanding too and improves the identification of patterns in the system.

It has still to be evaluated the effort of learning a new language, but, in principle, a visual modeling language should be easier to use than a typical programming language. The main issue, however, is the effort that is required to adapt existing agent metamodels for creating domain specific languages. However, in the case of

INGENIAS, this adaptation is feasible as both the language and the tool easily allow extensions to introduce new concepts and relations, together with graphical icons for them. A possible extension could be to differentiate the neighborhood from a standard agents group (such as “Families”), as it should be related with the space in some way.

Despite of these issues, this approach is considered a step forward in the search of more reliable and transparent agent-based models. The increase of formalization associated, together with the facilitation of replication, would restrain the typical criticism of complex models as obscure black-boxes.

References

1. Gilbert, N. and Troitzsch, K.G. *Simulation for the Social Scientist*. Open University Press; 2 edition (2005).
2. Gotts, N.M., Polhill, J.G., Law, A.N.R.: Agent-based simulation in the study of social dilemmas. *Artificial Intelligence Review*. 19 (2003) 3-92.
3. Drogoul, A., Vanbergue, D., Meurisse, T.: Multi-Agent Based Simulation: Where are the Agents? In: Sichman, J.S., Bousquet, F., Davidsson, P. (eds). *Proceedings of MABS 2002 Multi-Agent-Based Simulation*, LNCS Vol. 2581. Springer-Verlag, Bologna (2003) 1-15.
4. Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., et al.: Errors and artefacts in agent-based modelling. *Journal of Artificial Societies and Social Simulation*. 12(1) (2009) 1.
5. Edmonds, B.: The Use of Models - making MABS actually work. In: Moss, S., Davidsson, P. (eds). *Multi-Agent-Based Simulation*, LNAI Vol. 1979. Springer-Verlag (2001) 15-32.
6. Minar, N., Burkhart, R., Langton, C., et al.: *The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations*. Santa Fe Institute Working Paper. 96-06-042 (1996).
7. Brown, T.N., Kulasiri, D.: Validating models of complex, stochastic, biological systems. *Ecological Modelling*. 86 (1996) 129-34.
8. López-Paredes, A., Sauri, D., Galán, J.M.: Urban water management with artificial societies of agents: The FIRMABAR simulator. *Simulation*. 81 (2005) 189-99.
9. Mernik, M., Heering, J., Sloane, A.M.: When and how to develop domain-specific languages. *ACM Computing Surveys*. 37 (2005) 316-44.
10. OMG: *Meta Object Facility (MOF) Core Specification, Version 2.0*. <http://www.omg.org>.
11. Pavón, J., Gómez-Sanz, J., Fuentes, R.: Model driven development of multi-agent systems. In: Rensink, A., Warmer, J. (eds). *Model Driven Architecture – Foundations and Applications*. LNCS Vol. 4066, Berlin Heidelberg (2006) 284-98.
12. Aguilera, A., Ugalde, E.: A Spatially Extended Model for Residential Segregation. *Discrete Dynamics in Nature and Society*. 1 (2007) 48589.
13. Benenson, I., Torrens, P.M. *Geosimulation: automata-based modeling of urban phenomena*. John Wiley and Sons, Chichester, UK (2004).
14. Wolpert, J.: Behavioral Aspects of the Decision to Migrate. *Papers and Proceedings of the Regional Science Association*. 15 (1965) 159-69.
15. Galán, J.M., del Olmo, R., López-Paredes, A.: Diffusion of Domestic Water Conservation Technologies in an ABS-GIS Integrated Model. In: Corchado, E., Abraham, A., Pedrycz, W. (eds). *HAIS '08: Proceedings of the 3rd international workshop on Hybrid Artificial Intelligence Systems*. LNAI Vol. 5271. Springer, Berlin Heidelberg (2008) 567-74.
16. Benenson, I.: Modeling population dynamics in the city: from a regional to a multi-agent approach. *Discrete Dynamics in Nature and Society*. 3 (1999) 149-70.
17. Sansores, C., Pavón, J.: Agent-based simulation replication: A model driven architecture approach. In: Gelbukh, A.F., de Albornoz, A., Terashima-Marin, H. (eds). *MICAI 2005: Advances in Artificial Intelligence*. LNCS Vol. 3789. Springer Verlag (2005) 244-53.