Geosimulation in Education: A System for Teaching Police Resource Allocation

Vasco Furtado, Eurico Vasconcelos, University of Fortaleza, Master’s Program in Applied Informatics, Av. Washington Soares 1521, Edson Queiroz, Fortaleza, CE, Brazil
vasco@unifor.br, jose_eurico@yahoo.com

Abstract. This article describes the ExpertCop tutorial system. ExpertCop is a geosimulator of criminal dynamics in urban environments that aims to train police officers in the activity of preventive policing allocation. ExpertCop intends to i) help students comprehend the task of resources allocation and identify the variables involved; ii) induce students to reflect about their actions related to resources allocation; and iii) help students understand the relation between preventive policing and crime. In ExpertCop, the students (police officers) configure and allocate an available police force according to a selected geographic region, and then follow the simulation process. The students interpret the results with the aid of an intelligent tutor, the Pedagogical Agent, by observing how crime behaves in the presence of the allocated preventive policing. The interaction between domain agents representing social entities such as criminals and police teams allocated by the students drives the simulation. Assisting the user, the Pedagogical Agent aims to define interaction strategies between the student and the geosimulator in order to make simulated phenomena better understood. The innovative character of the work is in the strategy adopted by the Pedagogical Agent in order to assist the user in understanding the simulated model. It seeks to explain the model at the macro-level (global or emergent behavior) and micro-level (behavior of the agents individually). The evaluation of the system in training held with students of the domain has shown that the system attains its intended pedagogical goals.

Keywords. Intelligent tutoring systems, geosimulation, multiagent simulation, explanation, law enforcement

INTRODUCTION

Simulations make it possible to study activities and phenomena without the costs, risks or damages that such activities may bring about in a real environment, in addition to allowing the simulation of long periods on a reduced scale of time. Natural and urban disasters, migration of people, terrorist attacks, urban growth, criminal activity, and epidemics are examples of such phenomena and activities.

Geosimulation (Benenson & Torrens, 2004) proposes the aggregation of Multi-Agent Systems (MAS) and Geographic Information Systems (GIS) for modeling and simulating urban phenomena. Although it permits the modeling, representation and simulation of these activities, the process of geosimulation does not guarantee that these simulated models and the results obtained therefrom will be understood by the user. The complexity of the relations between the entities that compose these activities and the global behavior that emerges from these relations makes it difficult to understand the simulation process and consequently, its results. This problem is particularly relevant in addressing the use of geosimulation for educational purposes. It is therefore necessary to construct a pedagogical model that, among other things, helps the user in understanding the simulation and events that it represents.
In analyzing related works in the area of geosimulators and training based on multi-agent simulation, we observed that most such works are limited in their effective strategies of pedagogical support to students in their understanding of the simulation (Mann & Batten, 2002; Gibbons et al., 2001; Benenson & Torrens, 2004). Such works normally consider the simulation per se as being the pedagogical strategy to be adopted. Strategies for motivating the student or mechanisms that lead to reflection on the reasons why the simulated events occurred, which are commonly addressed in tutorial systems, are not systematic in geosimulators. Therefore, this analysis made it possible to identify the premises that guided our proposal, which are the following:

- Geosimulation becomes a good strategy for applying a “learn by doing” approach, in the case of urban activities that generally involve risks, costs or long periods of study. It is up to the student to interact with the system, describing his/her beliefs, and validating such beliefs with the simulations.
- The student needs to understand the simulation model which, in turn, must represent and manipulate characteristics of the physical environment of the domain studied.

In this article we describe an intelligent tutorial system that runs geosimulations of urban phenomena. In particular, we developed—in the domain of the Public Safety—the ExpertCop system to assist police officers in training for the urban activity of resources allocation. The allocation of resources is a key activity for the public safety sector, whereby police managers outline preventive policing strategies for a given region. The activity addresses the allocation of the available police teams for specific time periods and points of the city, with the intention of minimizing crime. ExpertCop requests a plan of resources allocation from the user and simulates how crime behaves vis-à-vis the police resources allocated. The objective of the tool is to allow users—generally police officers still in the academy—to understand the factors involved in the activity of resources allocation, to learn the procedures through simulated practice, and to reflect upon the causes-effects of the chosen allocations and the consequence of such choices on crime rates.

ExpertCop’s main contribution lies in the way we designed the Pedagogical Agent, which seeks to aid the student to understand the simulation through explanations. These explanations are supplied basically on two levels that we call macro- and micro-levels. At the macro-level, we use concept formation algorithms to identify emergent patterns in the simulation and which are provided in the form of “hints” to the students. At the micro-level, the Pedagogical Agent can provide explanations of the individual events reflecting the agent’s cognitive decision-making. In addition to these explanations, ExpertCop also features a gaming structure as a strategy for motivating and challenging the student, and possesses a set of interactive tools that allow for a statistical analysis of the results obtained from the simulation. The evaluation of the system was made by applying it in a course for police officers, and the results obtained show the effectiveness of the proposal.

This article is structured as follows: initially, we present the state of the art, in which we describe the adopted technologies and their correlations. In the next topic, we present the ExpertCop system, architecture, pedagogical purpose, tests, and results. In the final portion of the article, we present correlated works which lead to comparisons and conclusions.
BACKGROUND KNOWLEDGE

Urban Simulation and Intelligent Tutoring Systems

Simulations based on Multi-Agent Systems (MAS) are live simulations that differ from other types of computational simulations because simulated entities are individually modeled through the use of agents. According to Gilbert (Gilbert & Conte, 1995), the multi-agent approach (bottom-up) is appropriate for the study of social and urban systems. Social or urban environments are dynamic, non-linear, and composed of a great number of variables. MAS are also appropriate when the environments are made up of a great number of entities whose individual behaviors are relevant in the general context of the simulation.

A particular kind of simulation, called geosimulation, addresses an urban phenomena simulation model with a multi-agent approach to simulate discrete, dynamic, and event-oriented systems (Benenson & Torrens, 2004). In geosimulated models, simulated urban phenomena are considered a result of the collective dynamic interaction among animate and inanimate entities that compose the environment. The Geographic Information System (GIS) is responsible for providing the “data ware” in geosimulations.

Simulation is widely used as an educational tool because the computerized simulation of the activity studied allows the user to learn by doing (Piaget, 1976) and to understand the cause-and-effect relationship of his/her actions. According to Kolb (Kolb, 1984), learning is favored when the learning process occurs within the following successive steps:

- **Concrete Experience**: Obtained through the activity itself or its simulation in a virtual environment.
- **Reflexive Observation**: The experience is followed by the reflection phase. It is recreated internally in the user’s mind under different perspectives.
- **Abstract Conceptualization**: In this stage, the experience is compared and its patterns, processes, and meanings are analyzed. Within this context, abstract concepts and new knowledge are created. The knowledge is generated in two moments of the cycle, in this step and in that of concrete experience. The knowledge generated in the concrete experience phase comes only from the simple observation of the external event, while the knowledge generated in the abstract conceptualization phase emerges as a consequence of an internal cognitive process of the student.
- **Active Experimentation**: In this stage the student will conduct a new experiment with the newly acquired or modified concepts.

The simulation per se is not a sufficient tool for education. It lacks the conceptual ability on the part of the student to understand the simulation model. Therefore, some works (Taylor & Siemer, 1996; Angelides & Siemer, 1995) have tried to integrate the notions of Intelligent Tutoring System (ITS) and simulation in order to better guide learning and to improve understanding of the simulation process. The idea of an ITS is the integration of artificial intelligence in computer learning systems. It aims at emulating the work of a human teacher who has knowledge of the content to be taught, as well as how and to whom it should be taught. To achieve this, we need to represent i) the domain of study, ii) the pedagogical strategies, and iii) the student to whom the teaching is provided. A fourth component may also be considered (Kaplan & Rock, 1995; Woolf & Hall, 1995)—the interface with the user. The user interface determines how the interaction with the system is. Through the interaction
of these components, the ITS adapts pedagogical strategies on a domain at the level of the student for his/her individual needs.

**Knowledge Discovery in Databases**

Fayyad (Fayyad et al., 1996) defines Knowledge Discovery in Databases (KDD) as a non-trivial process for identifying valuable and useful patterns on data. The KDD process involves several phases, namely data selection, pre-processing, mining, and interpretation. These phases are iterative and some of them are often realized with human-interaction. Typically, the mining phase is an automatic one, where machine learning algorithms are used to discover clustering, classifications, and associations on data.

Concept formation algorithms such as COBWEB (Fisher, 1987) generate a probabilistic concept tree as that shown in Figure 1, which shows the specialization of a more generic concept in two more specialized concepts. Probabilistic concepts have attributes and values with an associated conditional probability of an entity having an attribute \( a \) with a value \( v \), given the fact that this entity is covered by the concept \( C \), \( P(a=v|C) \). The figure below depicts an example of a probabilistic tree formed from examples describing objects with three attributes (form, size and color). The concepts \( C_1 \) and \( C_2 \) are specializations of \( C_0 \), as \( C_21 \) and \( C_22 \) are of \( C_2 \). We interpret the concepts as the following: concept \( C_2 \), for instance, represents the objects with black color that have equal probability of being small or medium and equal probability of having a triangle or circle shape. Concept formation algorithms such as COBWEB are a kind of conceptual clustering whereby the data partitioning is driven by the meaning of a particular concept. Such an approach is particularly useful in data mining for constructing automatic and informative clusters, whereas the interpretation of the user for naming the clusters cannot be easily obtained. Furthermore, the conditional probabilities of both the properties and the concepts enrich the representation and allow the description of fuzzy concepts. As in this work we intend to discover crime patterns and show them to the user in terms of hints without any human intervention, we use COBWEB as the engine for generating the patterns that will originate these hints about crime.

**THE EXPERTCOP SYSTEM**

**Motivation**

Police resource allocation in urban settings in order to perform preventive policing is one of the most important tactical management activities that is usually decentralized by sub-sectors in police departments of the area. What is required from those tactical managers is to analyze the disposition of crime in their region and to perform the allocation of the police force based on such analysis. We agree with the principle that by knowing where crime is happening and the reasons associated with this crime, it is possible to make an optimized allocation and, consequently, to decrease the crime rate.

The volume of information that police departments have to analyze is one of the main factors for providing society with efficient answers. Tactical managers who perform police allocations, for instance, often have a lack of ability related to information analysis and decision-making based on such analysis. In reality, understanding criminal mapping activities, even using GIS, is a non-trivial task. In addition to that, experiments in this domain cannot be performed without high risks because
they result in loss of human lives. In this context, simulation systems for teaching and decision support are a primordial tool.

**Fig.1. Example of a probabilistic concept tree.**

**Objectives**

Our claim is that that educational systems aimed at the study of urban phenomena and activities must be modeled:

- in a multi-agent simulation process that offers “learning by doing”;
- in the entertaining, challenging, and motivating context of games;
- with the support of a tutorial agent which offers learning strategies to help users in understanding both the level of the agent and its actions as well as at the system’s emergent level;
- with appropriate representation and manipulation of the model as a whole, especially of the physical environment where the studied phenomenon occurs.

The ExpertCop system was developed based upon these premises, aiming to support education through the induction of reflection on simulated phenomena of crime rates in an urban area. The system receives as input a police resource allocation plan and thereby creates simulations of how the crime rate would behave in a certain period of time. The goal is to lead the student to understand the consequences of his/her allocation as well as to understand the cause-and-effect relations.

In the ExpertCop system, the simulations occur in a learning environment along with graphical visualizations that help the student’s learning. The system allows the student to manipulate parameters dynamically and to analyze the results.

**EXPERTCOP ARCHITECTURE**

The ExpertCop Architecture is composed of a user interface, a MAS platform, a GIS and a database. Each part of the ExpertCop architecture is described below.
The User Interface

The interface in ExpertCop allows the user to move among the functionalities and processes of the system in a logical, ergonomic and organized way. The interfaces that make up the system follow the logical flow of the use of the application. The system offers interfaces for configuration, allocation, simulation, analysis and evaluation. Each interface is endowed with a menu that allows the user to navigate through the functionalities of the system according to his or her needs.

The Geographical Information System—GIS

The GIS is responsible for generating, manipulating and updating a map of the area on a small scale. The map contains geographical and statistical layers, representing the characteristics of the area such as streets, demographic density, traffic signs, slums, commercial areas, etc. The GIS agent makes it possible for the other system agents to interact with the map by creating patrol areas, identifying structures and distances, identifying the domain agent positions, plotting them and allowing them to move about the map.

The System Database

The system database contains i) the information about each student and about his/her simulations, ii) the configuration data, iii) the real data and statistics on crime, and iv) the domain ontology. Such ontology is a definition of the basic concepts used in the domain and was produced by an expert on public safety. Abstractions such as the term “near” for defining a place that is between 0 and 200 meters are also represented in the ontology.

The MAS platform

The MAS platform makes it possible for the entities that make up the domain to be represented appropriately, both individually in their characteristics and behaviors as well as in their interrelationships. The structure, communication, administration, and distribution of the agents is provided by the Java Agent Develop Framework—JADE (TILab, 2003). The multi-agent platform in ExpertCop is made up of three groups of agents: Control Agents, Domain Agents and the Pedagogical Agent.

The Control Agents are responsible for the control, communication and flow in the system. The Control Agents are the GIS, the Manager, the Log and the Graphical Agent. In the following sections, we will describe in detail the domain and the pedagogical agent.

THE DOMAIN AGENTS

The domain agents are the actors of the studied domain. They act in the simulation process by representing the behavior of the entities intended to be “simulated”. These agents are defined by the specialist in the domain, contemplating the dynamic part of the model and the means of interaction between student and model. They can participate in the process of simulation with no direct intervention by the student or can be configured by him/her. It is the latter case wherein the student interacts with the system’s pedagogical model, demonstrating his or her beliefs about the domain
while attributing values for the characteristics of these agents and/or defining behaviors that will guide the agents during the simulation process.

Expertise on the domain, modeled in terms of the characteristics and behaviors of these agents, stems from interviews with specialists in public safety and from classic theories of criminology (Siegel, 2003) and sociology (Cohen & Felson, 1979).

**Police Teams**

The mission of the police teams is to patrol the areas selected by the student during the work period and work shifts scheduled for the team. An agent represents each team and has a group of characteristics defined by the student, such as means of locomotion, type of service and work shift that will influence his patrol. The team works based on its work period and work shift. The work period determines the beginning and end of work, and the work shift determines the work and rest periods. The patrol areas are composed of one or more connected points and are given to the police team as a mission. These areas are associated to intervals of time so as to fill out the work period of the team. A team with a work shift of 8 x 16 should patrol an area (or areas) for repetitive periods of eight hours every twenty-four hours.

One or more points that determine the area make up each patrol area. Initially, the police teams begin their activities at a common initial geographical point (the neighborhood police station). From that point onwards, the working teams (during their work period) verify the schedule during which their areas must be patrolled. After identifying the area which must be patrolled at that time, the police team lists, in order, the points that form that area and utilizes the first point as the objective point (OP). With the objective defined, the team should move towards it. To obtain its next position (NP) at each moment of the simulation, the team asks the GIS Agent for the next point between the current position (CP) and the objective point (OP) according to the speed (S) of the means of locomotion used. The Best First algorithm is used to define among the possible points between CP and NP.

The calculation of the agent’s walking time takes into account the time elapsed between the time at the last point request (TPrevious) and the current time (TCurrent). The equation \( D = (S \times (T_{current} - T_{previous})) \) defines the distance to be covered according to the time that has elapsed between the last request and the current one and the team’s speed. When arriving at its objective point, the agent considers the new point as its objective. Following this flow, the agent moves along the points that make up the patrol area. This process of going to different patrol areas and different patrol points is repeated until the end of the team’s work shift.

**Criminals**

**Criminal Agent’s Internal Architecture**

The criminals are the most important agents of the model. Their internal architecture is depicted in Figure 2 in three modules: perception, cognition and performance.

The criminal agent initiates its activities starting from the reception of information. This receiving process occurs in the agent’s Reception module (1), where messages sent by other agents, through the communication channels offered by JADE in Agent Communication Language (ACL) are received (2.1) and their content is interpreted (2.2) in accordance with ExpertCop’s messages protocol (2.3). The protocol specifies the format of the domain’s messages, allowing—during the interpretation—the
values and information proceeding from the message to be easily retrieved and mapped in values of local variables that contemplate the agent’s internal state (2.1). As for the message content, it addresses the criminal’s objective, carrying the mission’s data. The data contained in the message are interpreted and mapped in values for variables such as target_type, target_position, hour_opportunity, opportunity_date, opportunity_type.

With its modified internal state containing the data referring to its objective, the cognition process is initiated (2), whereby the agent executes its reasoning process (2.2). The Cognition module is responsible for the agent’s decision-making. This module uses the agent’s internal state with static and dynamic information on the agent and on the environment, and the ontological base which contains the necessary concepts and discerners for message interpretation. The process of inference of the cognition module is developed based on the Abstract-and-Match PSM. Problem-Solving Methods (PSM) provide architecture and reusable components for the implementation of the reasoning part of KBSs (Fensel et al., 2003). A PSM provides a structure in which a given type of problem can be classified facilitating the inference process. It is basically divided into the following phases:

1. **Abstract**: Initially the data are simplified, or “abstracted” in accordance with a set of rules preset by the specialist. Continuous data are discretized; for example, the value of the variable police_distance, collected from the GIS, is discretized from 152.25 to near. This discretization is executed from the rule: IF distance < 200.00 THEN police_distance = near. These discretizations modify the values of the variables that make up the internal state of the agent.

2. **Specify**: In this task, the specific evaluation criteria are selected in the internal state, such as risk, opportunity and reward. An agent can make several types of evaluation in its cognitive process, having a set of specific criteria for each type. In the case of ExpertCop, only one type of evaluation is made, to commit a crime or not, therefore the selected criteria will always be the same.

3. **Select**: The criteria specified in the specify stage are selected one by one and forwarded to the evaluation.

Fig.2. Agent’s architecture.
4. Evaluate: Each criterion has its value defined based upon a set of production rules; in this task, the criterion selected in the previous task is evaluated by means of scanning the production rules that define its value. A example of a rule that defines the risk criterion:

\[
\text{IF distance\_patrol = near AND type\_crime = robbery AND type\_victim = bank} \\
\text{THEN risk = high;}
\]

5. Comparison, match: The comparison task evaluates (executes based on an inference engine) a specific set of production rules in order to obtain the results of the evaluation. These rules are composed of the evaluation criteria. Therefore, at each criteria evaluation, it is verified—based on the criterion (or criteria) evaluated up to then—whether it is possible to arrive at an evaluation result or match. If the result is not reached based on criterion (or criteria) selected and evaluated up to then, the process returns to the Select task (3), where a new specified criterion (2) will go through the same process. An example of the comparison rule is presented below:

\[
\text{IF benefit = high AND risk = low THEN decision = commit\_crime;}
\]

The values of the variables regarding crime (type of crime, type of victim, geographical location of crime, date, and time) are sent to the criminal by the Criminal Manager. But to obtain the data on the environment (geographical factors), the criminal exchanges messages with the GIS Agent, which furnishes the geographical location, date, and time of the crime. Steps 1, 4 and 5 make use of a base of rules containing the structure of the decision support process and an inference engine, in our case JEOPS (Figueira & Ramalho, 2000), that sweeps these rules and associates them with the data collected on the crime.

Finally, the performance module is responsible for the actions of the agent. It contains the methods responsible for the agent’s internal and external actions and executes message sending. A decision is made as a result of the reasoning process (evaluation), whether to commit a crime or not given the presented opportunity. This decision is implemented by the performance module (3) where the action-taking process occurs (3.2) and the rendering of the action by means of sending messages (3.1) where victim or environment undergo the agent’s decision by means of the message.

**Criminal’s Knowledge Base**

The Criminal Manager creates each criminal agent in the simulation, with the mission of committing a specific crime. After the selection of the area and simulation period by the student, the Criminal Manager loads from the system database all the crimes pertaining to the area and period selected, and places the crimes in chronological order. When beginning the simulation, by observing the chronological order of the events, it creates a criminal agent for each crime. The criminal’s task is to evaluate the viability of committing the crime. The evaluation is based on risk, benefit and personality factors, defined on the basis of a set of interviews with specialists in crime from the Public Safety Secretariat and research in sociology and criminology.

- **The risk** is defined based on the variables:
  - **Type of crime:** Each type of crime is associated to a risk level, which is based on the type of punishment for the crime, on the level of experience and on the apparatus of the criminal. ExpertCop works only with robberies, thefts, and burglaries, which are types of crime influenced directly by preventive policing.
  - **Type of target:** The type of target indicates the capacity of resistance against a crime. The targets
considered in the system are persons, vehicles, residences, gas stations, drugstores, lottery houses, banks and commercial establishments. These targets are associated with the types of crime mentioned previously. Table 1 associates the risk value for each type of crime to each type of target. The possible values are: very low, low, average, high, very high and N/A when there is no association between type of crime and target.

**Police Presence:** Police presence (distance in relation to the scene of the crime) is the main factor that influences risk. The greater the distance between the closest team and the scene of the crime, the lower the risk is. We considered three categories for the evaluation of the criminal as the distance from policing. Any distance between 0 and 200 meters is considered near, between 200 and 500 meters is considered as average distance, and above 500 meters is considered far. We understand that the police presence (conspicuousness) is perceived radially; the criminals do not observe only the street where they will commit the crime, but evaluate the area all around the place in order to identify the presence of police officers and possible escape routes.

**Public illumination:** When the crime occurs at night, public illumination in the area is a factor of evaluation. Areas with deficient illumination facilitate criminal action, thus directly influencing the risk. The areas can be classified as poorly illuminated or well illuminated.

**Existence of escape routes:** The existence of places such as slums, woods, or deserted areas close to the place of the crime facilitates escape, thereby diminishing the risk involved in committing the crime. The classification as to the proximity of escape routes follows the same parameters as the distances of police teams. These areas may be near (0 to 200m), at average distance (200 to 500m), or far (above 500m) from the place of the crime.

- **Benefit** is defined based on the type and amount of goods the target can offer. In Table 2 we describe the possible values of goods that each type of target can offer.

- **Personality** defines the criminal’s “courage” level vis-à-vis the crime. When being created, a type of personality is randomly associated to the criminal (apprehensive, careful, fearless, and bold). A “bold” criminal evaluates risk with fewer criteria, giving more weight to the benefit. But an “apprehensive” criminal does the opposite, giving much more weight to risk.

<table>
<thead>
<tr>
<th>TARGET /TYPE OF CRIME</th>
<th>ROBBERY</th>
<th>THEFT</th>
<th>BURGLARY</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON</td>
<td>Low</td>
<td>Very Low</td>
<td>N/A</td>
</tr>
<tr>
<td>VEHICLE</td>
<td>Average</td>
<td>Very Low</td>
<td>N/A</td>
</tr>
<tr>
<td>DRUGSTORE</td>
<td>Average</td>
<td>Very Low</td>
<td>N/A</td>
</tr>
<tr>
<td>LOTTERY HOUSE</td>
<td>High</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>GAS STATION</td>
<td>Average</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>COMMERCIAL ESTABLISHMENT</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>BANK</td>
<td>Very High</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>RESIDENCE</td>
<td>Average</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>
Having collected all the necessary information for the decision support process of the crime to be executed, the agent uses the evaluation PSM (abstract-and-match) that will evaluate the viability of committing the crime. This process results in the decision of whether or not to commit the crime.

THE PEDAGOGICAL PROPOSAL OF THE SYSTEM

ExpertCop considers the simulation as a part of a pedagogical tool. The student can learn by doing. He/she initially interacts with the system allocating the police resources. It is a moment to expose his/her beliefs on the allocation of such resources. A simulation of the agents’ interaction is then done and the student’s beliefs can be validated by means of a phase of results analyses. This cycle can be repeated as many times as the student finds necessary. Moreover, the Pedagogical Agent (PA), as a strategy of support to the student, uses an explanation of the model simulated both at the micro-level—the behavior of the individuals (agents), and macro-level—the system’s emergent or global behavior.

Individual Explanation of the Events of the System (Micro-Level)

The micro-level explains the simulation events (crimes). ExpertCop uses a tree of proofs describing the steps of reasoning of the criminal agent responsible for the event. This tree is generated from the process of the agent’s decision-making stored in the database. The student can obtain the information on the crime and the process that led the agent to commit it or not, by just clicking with the mouse on the point that represents the crime on the map. Each crime located on the map is represented by a point in the color green (crime prevented) or red (crime occurring); these points function as hyperlinks to the explanation tree of the agent’s decision-making process. When the student clicks on a crime, the Pedagogical Agent will present an explanation containing the crime data and a hyperlink to the explanation of the reasons this crime was committed or not. When this explanation’s hyperlink is clicked, it leads to another interface containing the explanations based upon domain-specific concepts, where each concept is also a hyperlink, allowing the student to explore a more specific definition of the concept in the event that the concept *per se* is not sufficiently clear or unknown to the student. An example of this can be viewed in Figure 3. At the top frame, an explanation of the PSM is done. The
student can access the results obtained in the different steps of the PSM by means of links offered by the questions Which one and Why. The example shows an explanation of an abstraction and of a decision rule which were followed by the criminal at a certain moment of the simulation.

This type of explanation is obtained in ExpertCop because the agent’s architecture has a cognitive module modeled explicitly in terms of ontologies, production rules, and the abstract-and-match PSM as described in the section on the criminal agent's internal architecture. This tree of proofs is automatically generated by the PSM’s components, since such components were modeled adding the concept of pattern of explanation (Pinheiro & Furtado, 2004). The explanation pattern is coupled to the PSM and allows, in a transparent fashion, the PSM designer to generate trees of proofs not only of the PSM steps but also of the chaining executed by the inference engine based on the rules. The tree of proofs is represented in node sets of the Proof Markup Language (PML) (Pinheiro da Silva, McGuinness, & Fikes, 2004) which is one of the components of the Inference Web infrastructure. This infrastructure enables applications to generate portable and distributed explanations for any of their answers (McGuinness & Pinheiro da Silva, 2003). Thus, proof fragments in PML can be shared with other applications, besides using the IW’s infrastructure to abstract proofs into explanations and to present proofs and explanations to users.

**Explanation of the Emerging Behavior (Macro-Level)**

In ExpertCop, we understand as emerging behavior, the effects of individual events in crime, its increase or reduction, criminal tendencies, and seasonableness. For the explanation of the system’s emerging behavior, the Pedagogical Agent tries to identify patterns of behavior from the database generated in the simulation. The Pedagogical Agent (PA) does the KDD process automatically as illustrated in Figure 4.

First, the agent takes (requesting the LOG agent) the simulation data (events generated for the interaction of the agents such as crimes (date, hour, motive, type) and patrols (start time, final time, stretch)), and pre-processes them, adding geographic information such as escape routes, notable place coordinates, distance between events, agents, and notable places. After pre-processing, in the mining phase, the PA identifies patterns by means of COBWEB, a probabilistic concept formation algorithm. This algorithm generates a hierarchy of probabilistic concepts, which are characterized according to their attribute/value conditional probabilities. That is to say, a conceptual description is made of attribute/values with high probability. Having the probabilistic concept formation hierarchy constructed, the agent identifies and filters the concepts adequate for being transformed into questions to the student.

The heuristics used to filter which concepts will generate questions to the student, and the features that will compose these questions follow the steps below:

- The root of the hierarchy is ignored (not appraised), because it aggregates all the concepts that are too generic.
- The hierarchy is read in a bottom-up fashion from the most specific to the most generic concepts.
- The criteria used in the analysis of the concepts for selection are:
  - A concept must cover at least 10% of the total number of examples. We believe that less than 10% of the examples would make the concept poorly representative.
  - An attribute value is only exhibited in the question when it is present in at least 70% of the
The total number of the observations covered by an example.

- A question must contain at least three attributes.

- When going through a branch of the tree considering the previous items, in case a concept is evaluated and selected, the nodes superior to this concept (parent, grandparent, etc.) will no longer be appraised in order to avoid redundant information. This does not exclude the nodes in the same level of the hierarchy from those nodes that may be appraised in the future.

Fig. 3. Individual Events Explanation at two levels of abstraction. The abstract level explains the PSM strategy (first paragraph of the top frame). The concrete level explains the reasoning in the domain context.

An example of a COBWEB result is the concept exposed in Figure 5. That concept is displayed to the student as the following question (Hint): “Did you realize that crime: theft, victim: vehicle, day of week: Saturday, period: night, place: residential street, neighborhood: aldeota frequently occur together?” Having this kind of information, the student can reflect on changes in the allocation, aiming to avoid this situation.
SYSTEM FUNCTIONING

Initially, the student must register with the system and configure the simulation parameters using a specific interface. After that, he/she determines the number of police teams to be allocated and the characteristics of these teams. Based on the geographic and statistical data available on the map of the area and his/her knowledge about police patrol, the student determines the areas to be patrolled and allocates the police teams available on the geoprocessed map. To perform the allocation process, the student selects the patrol areas on the map for each team. After that, he/she defines the period of time that the police team will be in each patrol area. The sum of each period of time must be equal to the team’s workload. Figure 6 shows the interface for the allocation process and describes its main functionalities.
Agents representing the police teams monitor the patrol areas defined by the student following the programmed schedule. The patrol function is to inhibit possible crimes that could happen in the neighborhood. We presume that the police presence is able to inhibit crimes in the scope of a certain area. The goal of the student is to provide a good allocation, which prevents the highest number of crimes.

After the configuration and allocation process, the student can follow the simulation process on the simulation interface. At the end of the simulation process, the student accesses the pedagogical tools of the system. Figure 7 shows the functionalities for visualization.

In addition to the visualization functionalities, the student can access the explanation capabilities (see the previous section). A micro-level explanation can be obtained by clicking the mouse on any red or green point, which indicates occurred or avoided crimes, respectively. Figure 5 is a screen-shot of the screen at the moment of a micro-level explanation. The student can request a macro-level explanation pushing the Hint button represented on the screen. Figure 7 also shows how the concepts discovered by the probabilistic concept formation algorithm are presented on the screen. A set of questions (Hints) is shown to the student in order to make him/her reflect about possible patterns of crimes.

At each new allocation performed, the system will comparatively evaluate the simulated moments, showing the student whether or not the modification brought about a better effect on the crime rate. The PA also makes comparisons among results obtained in each simulation tour for evaluating learning improvements achieved by the student. The student can also evaluate the results among a series of simulations on the evaluation screen. On this screen the results of all simulations made by the student are shown in a bar graphic.

**EVALUATION**

ExpertCop was used to support a course at the Ministry of Justice and the National Secretariat of Public Safety—SENASP. The objective of this course was to emphasize the importance of information technologies in public safety. ExpertCop was intended to help police officers reflect on the forms of treatment and analysis of information and how this influences the understanding of crime. The audience was made up of three groups of thirty professionals in the area of public safety: civil police officers, chiefs of police, and military police (which are the majority). These groups of professionals compose the target public towards which this tool is geared.

**Methodology**

ExpertCop’s workflow tries to improve the learning process proposed by Kolb (Kolb, 1984) in the following four successive steps: i) the process of simulation as a concrete experience; ii) reflection and observation of the results with the aid of the support tools offered by the system; iii) abstraction and conceptualization, supported by the hints offered by the pedagogical agent on the patterns revealed in the simulation; and iv) new experimentation with the concepts acquired in a new process of simulation. The use of ExpertCop occurred in two distinct stages, one explanatory and the other evaluative.
In the first stage, the participants were instructed on the process of allocation of police resources—what it is all about, how it occurs in practice, and the factors involved in this process. After this contextualization, ExpertCop was presented, with its objectives and functionalities. Concluding this stage, the participants made use of the tool in an illustrative simulation to familiarize themselves with the functionalities.

In the second stage, training was carried out by a set of at least three simulations in city areas. In the first simulation, the participants had to create and configure a certain number of teams (according to the size of the area), allocate them on the map, and activate the simulation. At the end of the first simulation we asked the participants to identify, according to their beliefs, five factors (concepts) that influenced the occurrence of the crimes. They did so by observing the map of the crimes that occurred and those that were avoided. We requested that the participants not mention complex factors of political or socioeconomic order, such as unemployment or taxes, because we focused on geographical and/or visual factors that directly affect the crime rates. After collecting the participants’ concepts, we allowed them to use the pedagogical support of the system (clues, explanations and evaluations). After the use of the pedagogical support tools, the collection of factors influencing the crimes was carried out again. In the subsequent simulation, we repeated the same area to serve as a comparison with the initial simulation already completed, and allowed them to make their allocations and use the pedagogical support of the tool according to their needs. Afterwards, we performed two other simulations with different areas than the first one. Using different areas for each simulation allows us to evaluate if the student was able to abstract the modified or acquired concepts by applying them in different contexts (characteristics), since the new concepts are applicable to all of the characteristics that the environment presents. During the simulations, the time needed to accomplish the allocation
process in the training simulations was measured for a sample of the participants (two from each
different group, civil police officers, chiefs of police, and military police).

Fig. 7. Visualization Functionalities.

**Hypotheses**

**1st Hypothesis**

Bruner (1966) describes learning as an activity, in which learners build new ideas or concepts, based
on their past and actual experiences. Based upon this notion, we formulated the first hypothesis: the
system will make the students improve their understanding of the results (causes and effects),
acquiring new concepts or modifying old ones regarding crime and the allocation process.

**2nd Hypothesis**

Based on the belief that the percentile of crimes avoided in relation to the total number of crimes
attempted represents the participant’s performance in a simulation, we formulated the second
hypothesis: the acquisition of new concepts and beliefs would make the students improve their
allocation and consequently obtain better results in the simulations.
Analyses and Results

Analysis of results in the first hypothesis

As each student was supposed to write down, in each collection phase, at most five relevant factors, the total number of collected concepts was 450. In fact, in the second round the total number of concepts was reduced (by 2%), since some concepts were excluded because the students no longer considered them relevant. From the analysis of the collected concepts, compiled in Figure 8, it was possible also to observe that:

Fig.8. Evaluation of the second group of collected beliefs.

- 23% of students included new concepts about motives and causes of crime which were not mentioned in the first round. These concepts represent totally new factors that were not mentioned before. New concepts about specific factors such as public illumination, patrol route distances, the existence of slums and work shifts were included by the students in the second collection. Concepts describing details of the geography of the area came to be described. The identification of escape routes by means of the plot of parks is an example of that.
- Concepts that were more specific and practical replaced those initially observed, which were more generic. 64% of the initial concepts were changed, in the sense that they are shown to be more informative mainly due to specialization. For instance, in the first collection, the students mentioned the existence of attractive targets as a factor relevant to the occurrence of crime. In the second collection, they specified that the existence of residential targets is relevant for a specific type of crime (burglary) at night. Time factors, such as the relationship between the day of the week and the periods of the day to the number of crimes occurred, began to be specified.
- 11% of the concepts were maintained (they were written down again in the same way as the first collection). Among them, 85% came from veteran military police officers with some experience in the allocation process.

According to the data, it is possible to conclude that the first hypothesis is valid, confirming that the system offered support so that the student modified his/her beliefs on the domain or acquired new concepts (learning).

Analysis of results in the second hypothesis

Based on an analysis of the relation between the number of times the system’s pedagogical support was accessed and the results of the simulation, it was possible to observe that:
• There is a positive linear relation between the number of times pedagogical support was accessed and the results of the simulation, as presented in the dispersion graph shown in Figure 9;
• Students with a higher number of accesses obtained better results.

Also regarding the second hypothesis, we observed that the average of the results of the simulations after the use of the pedagogical support (second and third simulations) was 9.09% greater than the average of the results in the simulation before the use of the support offered by the system. The average for the first simulation was 28.65% while in the second one, the average results were 31.25%. Analyzing the significance of this increase (9.09%) by means of the t-Test for similar samples, it was possible to argue that the result is relevant with 99% of certainty (significance level 1%). Even though the subsequent simulations were with different regions, they did not present greater differences in terms of complexity, and maintained the levels of improvement in relation to the two initial simulations.

**Discussion**

We analyzed how sensitive ExpertCop was to the gaming effect. Some students, after having obtained high scores in the first simulation, decided to play with the system and defined routes that were totally unfeasible. None of them were able to improve scores previously obtained. In virtue of the complexity of the factors involved, it is not easy to find a gaming strategy that leads to good results without a complete understanding of the problem. Excluding those cases where gaming was identified, there was an even greater improvement in the percentage of learning of the students, on the order of 12%.

We also observed that although the students demonstrate more ability in handling the tool as time goes by, they spend more time during the allocation process. We think that this indicates a more reflexive allocation process. Another observation is that an atmosphere of cooperation (or competition) was created among the students, and they often compared results and patrol routes seeking to identify similar strategies among themselves. Effectiveness of learning depends on the student profile. A novice in computer science tends to concentrate on the tool instead of on the allocation process. Experienced students in the allocation process improve their performance to a lesser degree. Based on the results, we may also conclude that the learning level is higher in participants with little or no experience in the domain or in the treatment of information. Also
evaluating these results, we conclude that the pedagogical support offered by the system helps the participant understand and better identify the factors that affect crime, allowing thus for better performance in their allocations and consequently a reduction in crime levels. The students were capable of noticing the importance of the analysis of the data in the allocation process. The tool was revealed as easy to use and attractive to the students. They continued using the system even after the end of the course.

We observed as a negative fact that students who obtained very low results were not motivated to follow the subsequent simulations. This fact occurred with three out of the total of ninety students who took part in the course. Another important aspect that merits discussion is how learning is influenced by the quality of the model. In our discussions with educators, we were advised to be careful that students do not come away thinking that everything they did in ExpertCop will be reproduced equally in their area of work. For this reason, we prefer not to present students with the areas where they actually work. The objective of the tool is to make them reflect on causes and effects. We based ourselves on the ideas of Dewey (1910) that educational environments must support the practical methodology of problem solving. Moreover, Schön (1987) perceived that during the exercise of their activities, professionals often resort to non-planned activities in order to solve problems. These strategies are conscious, since they are actually making use of a series of rules to achieve success, in a true epistemology of practice. Thus, the idea is reinforced that the student is capable, through practical situations, of reflecting upon alternatives of allocation of police officers.

The matter of modeling the criminal also deserves mention. In this work we have chosen to model the criminal as a cognitive agent. That choice was based on the idea that the teaching process would obtain leverage from the explanations that such a modeling can provide. The cognitive modeling of the criminal is a difficult task, and it is practically impossible to acquire it and to represent it completely. This certainly impacts the accuracy of the simulation model as a whole. However, as a complete and totally accurate crime simulation model has not yet been proposed (even though we are pursuing this goal as we will state in the Conclusion section), our intention here is to model the basic knowledge that is found in the literature as well as heuristics supplied by police officers. In this way, the explanation of why crimes are committed leads the student to know some of the factors that were considered during the elaboration of the crime. This knowledge is shown opportunistically within a precise context. The design choices made in ExpertCop were based on police expertise and on criminology theories. One might argue, for instance, that police patrolling one area may in reality result in moving crime to another area, instead of lessening it. Empirical evidence and sociological studies (Tonry & Farrington, 1995) have shown differently. Crime spatial distribution follows a Zipf law in which most of the crimes are concentrated in a few places whereas most other places have few crimes, because crime recurrence is preferred by criminals. The crime migration phenomenon actually happens but not without implications on the amount of criminal activity during a period of time. The mere presence of police at a vulnerable point is not enough to completely change the motivation of criminals. However, forcing the migration is a strategy to reduce crime because criminals feel uncomfortable when they are obliged to act in unfamiliar places. Doing so, the police reduce the criminals’ productivity or diminish their impetus to practice crime by leading them to commit errors and being captured. Briefly, the assumption that by policing a hot spot leads to crime reduction is not incorrect for learning purposes even though this reduction was obtained by means of crime migration.

Finally, we point out that whether, for any reason (perhaps due to his/her field experience), the student is led to disagree with the performance rules of the criminal agent or to elaborate other
characteristics that were not observed, ExpertCop has nonetheless fulfilled its role of making the student reflect upon these factors. Moreover, maintaining these rules that determine the cognitive model of the criminals, since they are explicit in an ontology, can be easily modified and adapted to different realities with no need to change the source code.

RELATED WORK

Some approaches have influenced our work more strongly, such as multi-agent simulation, intelligent tutorial systems, educational games, and the common ground among these areas.

Based on these approaches, there are many projects and systems that describe solutions that are similar with parts of our system design. Virtual environments for training, such as Securevi proposed by Querc et al. (2003), which is a system based on the Mascaret model, use multi-agent systems to simulate realistic, collaborative and adaptive environments for training simulation. Intelligent GIS, such as the system proposed by Djordjevic-Kajan (1995), intends to provide computer support in fire rescue. The system has a “Fire Trainer”, an agent that covers the activities connected to education. The Phoenix system (Cohen et al., 1989) is a discrete event simulator based on an agent architecture. The system is a real-time, adaptive planner that simulates the problem of forest fires. The DEFACTO system (Schurr et al., 2005) is also an example of geosimulation for the public safety domain for training the coordination of various disaster response entities. Another similar application is the CACTUS multi-agent simulation system (Hartley, 2001) which is a training system to exercise police officers in managing public order events. The simulation and debriefing is intended to develop strategic and contingency planning skills. The trainees work as a team, communicating as they would in a real situation, mediated by a facilitator (trainer). Instead of providing tools for helping the understanding of the decision-making process, emphasis is given to the human facilitating process.

The main difference between these works and ours is the fact that we propose a pedagogical agent to help the understanding of the simulation model while these works consider the simulation as the only pedagogical support to the student.

Intelligent Tutoring Systems like the one built by Wisher (2001) describe intelligent tutoring for field artillery training, or the Sherlock system by Lesgold (1992) that provides advice for impasses while using a simulated system. The architecture proposed by Atolagbe (1996) and Draman (1991) for educational simulation also has points similar to this work, although they do not emphasize the power of simulation in GIS or the use of KDD to improve student learning. In relation to these works, our approach presents an innovation in the way we provide assistance to the student by means of explanations of the events. Individual explanation was possible due to the particular manner in which we modeled the criminal agent, by associating an explanation component with the agent’s reasoning process. Moreover, we use concept formation algorithms to identify patterns on simulation data to explain macro behavior.

Several works in games and entertainment (Galvão et al., 2000; Leemkuil et al., 2003) use simulation with an educational purpose. Even though they present some similarities to our approach, game simulators have a different pedagogical strategy. They focus on the results of the simulation while we believe that the most important aspect is the process itself.
CONCLUSION AND FUTURE WORK

This work describes the ExpertCop system, a pedagogical geosimulator of crime in urban areas. The system architecture is based on the integration of MAS, GIS, and a Pedagogical Agent producing an Intelligent Tutoring Geosimulator. ExpertCop aims to train students (generally police officers still in the academy) in the urban activity of resources allocation, a key activity in the effective combatting of crime. In order to provide support to the student in understanding the results of the geosimulation and the model proposed, the Pedagogical Agent uses, as primary strategies, the explanation of the model both at the micro-level, or level of the agent’s individual activity, as well as at the macro-level, or level of the system’s global, emergent behavior. At the micro-level, the agent uses the steps given in the decision-making process of each agent to explain individually each event that occurred in the simulation. At the macro-level, the agent uses the data mining process to identify the system’s patterns of behavior. This approach of pedagogical support and its implementation provide the innovative character of this work.

Initial training sessions with police officers interacting with the system were performed, aiming to evaluate learning by using this tool. As a complement to the use of the system, a course was held where ExpertCop was used as a tool for analysis and reflection of practical situations. The methodology adopted to analyze the learning of students in ExpertCop has shown a significant improvement in the students’ data analysis abilities, in the process of resource allocation with ExpertCop, and in the identification of factors that influence crime.

We intend to continue this research on the ExpertCop system, enhancing its functionalities and increasing the training support, with the aim of making it not only an educational tool but a decision-making support tool as well. The next steps are to render ExpertCop multi-user and to make it available on the Web, to study and develop a student model suited to the system, and to analyze the viability of implementing a framework for educational urban simulation based on the principles contemplated in ExpertCop.

Another ongoing work aims at transforming ExpertCop into a decision-making support tool. In doing so, we are aware that the accuracy of the simulation model is essential. Therefore we have adopted a different approach for the crime simulation model. Instead of cognitive agents, criminals are reactive agents with behavior driven by real crime data and the statistical distribution of crime in urban areas. We model the criminal as distributed entities with the ability to demonstrate self-organization from their individual (local) activities as well as taking into consideration the influence of other criminals in the community where they live (Furtado et al., 2006). We are also designing an evolutionary approach that integrates with the simulation tool and is devised to assist police officers in the design of effective police patrol routes. Our approach is inspired by the increasing trend of hybridizing multi-agent systems with evolutionary algorithms. Our idea is to uncover strategies for police patrolling that cope with the dynamics of the crime represented by criminals that learn ‘on the fly’. To uncover good police patrol routes in this context, we are integrating into the simulation model a genetic algorithm. Preliminary experiences have shown that such an approach is very promising (Reis et al., 2006). Our long-term challenge is then to design and integrate educational strategies into a much more accurate crime simulation model.
ACKNOWLEDGEMENTS

This work was partially supported by the Brazilian Research Council (CNPQ) grant number 552009/2002-4, FUNCAP, and Motorola Brasil. The authors would like to thank Bill Clancey and the anonymous reviewers for their insightful comments and suggestions which helped greatly to improve the quality of this paper.

REFERENCES


