Constructing Progressive Learning Routes through Qualitative Simulation Models in Ecology

Paulo Salles

Universidade de Brasilia
Instituto de Ciências Biológicas
Campus da Asa Norte
Brasilia - DF
70.710-900, Brasil
Phone:+55 61 348-2286
E-mail: paulo.bretas@uol.com.br

Bert Bredeweg

University of Amsterdam
Department of Social Science Informatics
Roetersstraat 15
1018 WB Amsterdam
The Netherlands
Phone: +31-20-525 6788
E-mail: bert@swi.psy.uva.nl

Abstract

Qualitative models support interactive simulations that are well suited to help learners in acquiring causal interpretations of physical systems and their behavior. Such simulation models can be large, particularly if they include many subsystems. When simulations are too big they hardly can be used effectively for teaching purposes. They have to be reorganized into smaller sets of simulation models and ordered in a sequence for the learner to progress through. Model-dimensions and techniques such as Causal Model Progression have been presented as means to address this problem. In this paper we investigate how to decompose a large qualitative simulation into a progressive sequence of smaller simulations, useful for teaching purposes, in the domain of ecology. Based on notions introduced by Causal Model Progression, the Genetic Graph, and the Didactic Goal Generator, we have constructed a set of dimensions that can be used in this respect. Following these dimensions we show how a large qualitative simulation model of the Brazilian Cerrado vegetation dynamics can be rearranged into a sequence of clusters, each representing and simulating distinct features of such ecological systems. These clusters are ordered in evolutionary model progression lines according to movements from static to dynamic models and, by incorporating structural changes, from less complex to more complex models. The approach presented in this paper thus provides means, in terms of knowledge characteristics, to effectively reorganize qualitative simulation models for teaching purposes. In the discussion we briefly argue that this approach may also be applicable to qualitative simulation in other domains.

Introduction

A student usually cannot learn all of the subject matter in one step. Therefore, when teaching a substantial complex domain, the subject matter must be divided into units, each unit dealing with a part of the whole spectrum, and ordered in sequence that can be traversed by the learner. This problem is sometimes referred to as 'curriculum planning' and concerns at least three questions: how to divide the subject matter into appropriate units, with which unit to start the learning, and how to proceed through the available units?

Salles & Bredeweg (1997) describe an implemented qualitative simulation that represents the Brazilian Cerrado vegetation dynamics, following the 'succession hypothesis' described by Pivello & Coutinho (1995). This 'Cerrado Succession Model' includes general concepts relevant for reasoning about population and community dynamics, as well as specific details that are necessary for understanding the ecology of the Brazilian Cerrado. When using this model for teaching students in ecology classes it turns out that the model is too big and too complex to be dealt with in one step. Teachers have a tendency to break the model into parts, e.g. by first discussing populations, then communities and finally the Cerrado specific details (which relates to how this kind of material is organized in textbooks of ecology, e.g. Gotelli, 1998). However, the Cerrado Succession Model as implemented and described by Salles & Bredeweg (1997) does not explicitly support this division and ordering of the subject matter. Moreover, it lacks theoretical underpinning of what progressive sequences can be constructed for this simulation model.

Human cognitive abilities provide constraints on how to organize the subject matter into digestible units for learning. However, there is not a unified theory on cognition and learning that explains how humans learn and how this learning can be aided (by teaching). In fact, there are many competing views and exactly understanding how humans learn, is still subject of current research[1]. In this paper we do not pursue further enhancement of theories on human cognition and learning. Instead, we want to find arguments in terms of 'knowledge characteristics' for effectively dividing a large qualitative simulation model into 'standalone units' (i.e. still simulation models, but

[1] An interesting enumeration of competing theories on cognition and learning can be found in TIP (Kearsley, 1994-2000).

smaller) and for ordering them in a progressive sequence to support learning.

There is not a large amount of research on curricula design within the AI-ED community. Among the older work is the notion of a 'Genetic Graph' (GG) (Goldstein, 1979). It provides a set of primitives, including notions such as 'refinement', 'specialization', 'generalization' and 'analogy', to categorize and order 'knowledge units' within the subject matter. More recently White & Frederiksen (1990) describe the notion of Causal Model Progression (CMP). CMP is designed for sequencing (partly qualitative) simulation models of electronic circuit behavior. As we are dealing with qualitative simulations, this work seems relevant, particularly the concept of 'order' for sequencing behavior models. Winkels & Breuker (1993) describe the notion of Didactic Goal Generator (DGG) in which they, among others, augment the notion of a GG (e.g. including the dimensions 'inversion' and 'abstraction/concretion'). However, they do not relate their work to CMP. Finally, notice that the three approaches mentioned here originate from domains and application areas that are very different from ecology, particularly from qualitative simulation models of ecological systems, which may limit the reusability of the ideas.

This paper presents an argument for constructing progressive learning routes through qualitative simulation models in ecology. First, the Cerrado domain and the succession simulation model are introduced. Second, the reusability of existing approaches is investigated. It will be shown that none of the above mentioned approaches is fully sufficient, which enforces us to reuse and integrate specific parts of the different approaches. Moreover, particular aspects, such as structural changes, cannot be dealt with at all. New primitives are introduced for that purpose. The overall result is a modified set of model dimensions that is better adjusted to qualitative simulation models of ecological systems. Third, this new set of model dimensions is used to re-organize the full Cerrado Succession Model into a sequence of model clusters. Finally, the progressive sequence of model clusters is illustrated by discussing scenarios, resulting simulations and possible learner interactions for each model cluster in the sequence.

Qualitative Simulation of Cerrado Ecology

Cerrado is the vegetation that used to cover almost two million square kilometres in the central region of Brazil until 40 years ago. Today Cerrado is an endangered vegetation, with only 20% of the original area remaining. The Cerrado consists of many different physiognomies, spanning from open grasslands to more or less closed forests. These physiognomies are mainly affected by fire, soil fertility, and the amount of water available during the dry season. It is widely accepted among researchers that if fire frequency increases above 'natural' levels, woody components decrease and graminoid components increase, so that the vegetation becomes less dense. Vice versa, if fire frequency decreases, the vegetation becomes denser and changes toward forest-like communities. This 'Succession Hypothesis' is supported by different studies (see Moreira, 1992; Pivello, 1992; Pivello & Coutinho, 1995; Salles, 1997).

Salles & Bredeweg (1997) present the 'Cerrado Succession Model' (CSM) a fully implemented qualitative simulation model of the succession hypothesis. This model consists of a library of model fragments for reasoning about the behavior of populations and communities and was built to be used in an Interactive Learning Environment, in which learners can run simulations to explore different aspects of the Cerrado vegetation[1]. To represent communities (complex entities consisting of many types of plants and animals) the notion of ‘functional groups’ was applied. This way, communities were modeled as groups of three populations (trees, shrubs and grass). A domain theory of population dynamics was developed to be the ‘first principles’ on the top of which it is possible to make predictions and generate explanations about the behavior of communities. Following Forbus (1984), changes were represented as caused by the direct influences of processes, which then propagate via indirect influences to other quantities; representing the behavior of the system. The library consists of model fragments representing views and processes, which encode knowledge about typical situations, objects, quantities, quantity values, dependencies between quantities, and the conditions for processes to happen, and the mechanisms for changes to propagate.

A kernel of model fragments about populations was built, with representations for different types of organisms (mainly plants), population size (expressed as small, medium and large according to the value assumed by the quantity ‘number of’) and direction of change (increasing, stable and decreasing). Also four basic processes (natality, mortality, immigration and emigration), refinements of these basic processes (e.g. colonization, seed dispersal) and the composite process population-growth, resulting from the aggregation of the basic processes, were identified and implemented. There are five types of communities: Campo Limpo, Campo Sujo, Campo Cerrado, Cerrado Sensu Stricto and Cerradão. Each was modeled as consisting of grass, shrub and tree populations in different proportions. Campo Limpo, for instance, has a population of grass of 'low' size, and has no shrubs and trees ('number of' shrub and tree equal zero). Campo Sujo has a population of grass equal 'high', a shrub population of 'low' size and a population of trees with values ranging from 'zero' to 'medium'. And so on, until the Cerradão community, the most dense community, with number of shrub 'high', number of trees 'maximum' and no grass. The whole model set was further extended with the inclusion of

[1] VisiGarp is model inspection tool that can be used for this purpose (Bouwer et al., forthcomming).

environmental factors such as human actions, fire frequency, cover, litter, and the conditions at the ground level involving water, nutrients, temperature and light.

The implemented Cerrado Succession Model supports many qualitative simulations (using different scenarios). The most important one is the full succession simulation in which the three populations (grass, shrub, and tree) interact with environmental factors such as fire frequency, to change from Campo Limpo to a dense Cerradão. For more details see: Salles (1997) and Salles & Bredeweg (1997).

Principles for Organising the Subject Matter

Causal Model Progression (CMP) (White & Frederiksen, 1990) defines three dimensions for models to vary: perspective, order and degree of elaboration. Perspective concerns the overall view of a system. For explaining electronic circuit's operation three perspectives are identified: functional, behavioral, and physical models. The dimension order further refines the notion of behavior models. Typically, zero-order models are static, in the sense of not capturing continuously changing behavior. In zero-order models quantities change values abruptly, such as a light bulb going from on to off when a switch is turned off and both components are part of the same circuit which also includes a charged battery. In first-order models behavior can change gradually, such as a resistor gaining more resistance as power increases. Finally, second-order models include knowledge about relative changes (in fact, second order derivatives), so that knowledge can be captured such as one resistor building up resistance faster than another resistor. The third dimension defined by CMP is degree of elaboration. Basically, it refers to the amount of inference detail that is required for deriving a particular behavioral fact. A model is more elaborated if it has more intermediate steps in the behavior dependency chain that must be reasoned about (increased number of qualitative rules).

A distinction can be made between a *system* model, a computer model simulating a system existing in the physical world, and a *mental* model, a description of how learners 'understand' a system that exist in the physical world (possibly implemented as a computer model). The distinction is not always clear-cut. In CMP the dimensions focus primarily on mental models. The dimensions define a model space that a learner should master. The hypothesis is that progression along the dimensions will aid this learning process, because 'earlier, less complex models' provide insights needed to understand 'later, more complex models'[1]. For our research problem it is interesting to find out how the CMP characteristics of mental models can be used to classifying system models. For perspective there seems to be no significant difference, both a system and a mental model can be within a certain perspective. Moreover, if a learner has to acquire insights from a particular perspective, probably s/he has to interact with a system model that captures the real system's details from that same perspective. Also for order there seems to be a great overlap. A system model can also be of zero, first or second order, in terms of how behavior dynamics are represented. Again, acquiring a first-order mental model is probably best supported by interacting with a first-order system model. Degree of elaboration, as defined by CMP, does not seem to have this one-to-one mapping between system and mental models. Consider the following examples. A system model of a circuit including regular conductivity can be made very complex (for instance by including many components), but still require only one degree of understanding conductivity and voltage. Whereas at the same time we can make a far more simple model (e.g. having only one or two components) that does require an elaborated understanding of conductivity and voltage (because the circuit has a short). The problem is the mixture of 'structural changes' and 'adding more detail'. For instance, a mental model of a container-piston assembly that includes the concept of friction, between the piston and the container, is more elaborated than a model of the same system without the concept of friction. The same classification can be used for system models of the container-piston assembly. The model with friction is more elaborated (according to CMP), because it uses 'more qualitative rules' to infer how the behaviors of components influence each other. But notice that the structural details represented in both system models are the same (the components of the container-piston assembly). If, on the other hand, we change the *structure* of a system two things are possible. Understanding it may require a higher degree of elaboration in terms of CMP. Or it may be a more complex system but not require a mental model shift. In conclusion: it turns out that 'structural changes' and 'adding more detail' (i.e., elaboration in terms of CMP) are two distinct phenomena that should be dealt with separately when determining model complexity. As it will be demonstrated below, our paper introduces structural changes in the context of model progression for the qualitative simulation of the Cerrado Succession Model.

The Genetic Graph (GG) (Goldstein, 1979) uses four dimensions to classify elementary sub-skills (i.e., individual rules): refinement, specialization, generalization and analogy. If the student masters all the rules s/he will be able to assess the situation at hand adequately and act in the most optimal way. Seen from that perspective, a refinement step refers to identifying a new feature (or a concept), that applies to some entities and not to others (something is dangerous or not). A specialization step refers to further detailing a concept: there are different ways in how it can manifest itself (there are different kinds of dangerous things). A generalization step is the opposite of a specialization step (grouping different manifestations under a single concept). Finally, an analogy step refers to identifying other manifestations of the same concept (sound is dangerous, similar to smell). Winkels & Breuker

[1] An alternative approach could be to more 'randomly' proceed through the model space, as e.g. emphasised by Cognitive Flexibility Theory (e.g. Spiro *et al*., 1988).

(1993) modify and extent the ideas presented in the GG when describing their Didactic Goal Generator (DGG). They define generalization/specialization for organizing concepts (with less/more attributes) in a hierarchy. Inversion refers to concepts being opposites (e.g. delete versus past). DGG also defines analogy (similar to how it is used for the GG). Similarity is defined as a particular kind of analogy, namely as a single concept having two names. Finally, DGG defines abstraction versus concretion, which distinguishes between support and operational knowledge (how does a computer application work and how can it be used).

For organizing models of ecological systems it seems that the dimensions defined by GG and DGG provide us with means to handle 'hierarchies of concepts' (concepts in a broad sense). Consider for instance the following statements. A 'shrub population' is a kind of a 'plant population' (the former has more features and is therefore a specialization of the latter). A 'natality process' is analogous to a 'immigration process' (both increase the number of individuals). A 'mortality process' is the inverse of a 'natality process' (one decreases and the other increases the number of individuals). On the other hand, the dimensions defined by CMP particularly provide us means to handle 'orders of behavior models'. For instance, we can distinguish zero-order (static) models from first-order models, in which things are changing. Within the context of the former we can talk about the composition of a community. The latter we can use to discuss how things are changing because certain processes are active. In the next section we elaborate on how to use the primitives discussed above to effectively divide and sequence qualitative simulation models of ecological systems.

Decomposing and Ordering the Cerrado Succession Model

Libraries for the qualitative reasoning engine GARP (Bredeweg, 1992) consist of different types of model fragments (see also figure 1). A single description fragment (S-mf) models features of a single entity (or concept) (e.g. a tree, or a population) and can be organized in subtype (is-a) hierarchies (e.g. tree-population is-a plant-population, which again is-a population). A process fragment (P-mf) influences features of entities described by a S-mf (e.g. natality in a population), so the latter has to be applicable (it is conditional) before a process can become active. P-mf may (it is optional) also be organized in subtype hierarchies (e.g. natality in trees is-a kind of natality). From a technical point of view, agent fragments (A-mf) are similar to P-mf: they influence, i.e. change, features of entities. But they differ conceptually from P-mf in that they model actions that are exogenous to the system as a whole, i.e. an external agent is enforcing the changes (e.g. a person controlling fire frequency). Compositional fragments (C-mf) specify features of interacting entities (e.g. symbiosis, or populations being part of the Cerrado Sensu Lato). Of course the S-mf describing the entities have to be applicable before the C-mf can become active. C-mf may also be organized in subtype hierarchies. P-mf and A-mf may also apply to the assemblies formed by a C-mf (e.g. a process that is only active in a Campo Sujo). P-mf influencing assemblies may also be organized in subtype hierarchies.

Given a scenario, the simulator searches for applicable fragments from the library[1] and constructs a behavior graph, representing the behaviors that may follow from the initial specification. Notice that a large number of scenarios can be created for the same library. To construct a progressive learning route through this potentially large set, the previously discussed model dimensions can be used. Our proposal is discussed below. In the following section this approach is further illustrated with examples.

Generalization/specialization (G/S) The subtype hierarchy is used to organize model fragments on this dimension. A fragment is a specialization of another fragment if it is a subtype of that fragment. A specialization specifies at least a new name, but usually also introduces new features. Notice that features may come in many forms[2], such as quantities, causal dependencies, value ranges, etc. Generalization is the opposite of specialization. It refers to 'moving-up' the subtype hierarchy, i.e., grouping different fragments into a single immediate super-type. E.g. identifying shrub and tree as being both plant-populations (and e.g. different from animal-populations).

Analogy (A) Two fragments they are in principle analogous when they are both immediate subtypes of the same super-type. They share at least the knowledge specified in the super-type, but often they also differentiate on particular features. For example, in the succession model, Campo Limpo and Campo Sujo consist of similar kinds of populations but the populations differ in size for each Campo type.

Inverse (I) Similar to DGG we define inverse as a special kind of analogy. Namely, when two immediate subtypes have opposite features. Inverse is in principle reserved for 'activities'. In terms of ecology this means agents, or processes, with opposite behavior. In terms of model fragments it means that inverse agents or processes have opposing influences. For example, with regard to how a population changes: natality is the inverse of mortality, and immigration is the inverse of emigration (whereas natality is analogous to immigration and mortality is analogous to emigration).

---

[1] Notice that fragments may apply multiple times. For instance, the fragment 'population' will apply to all the populations mentioned in a scenario.

[2] For a detailed description of model fragments in GARP see Bredeweg (1992).

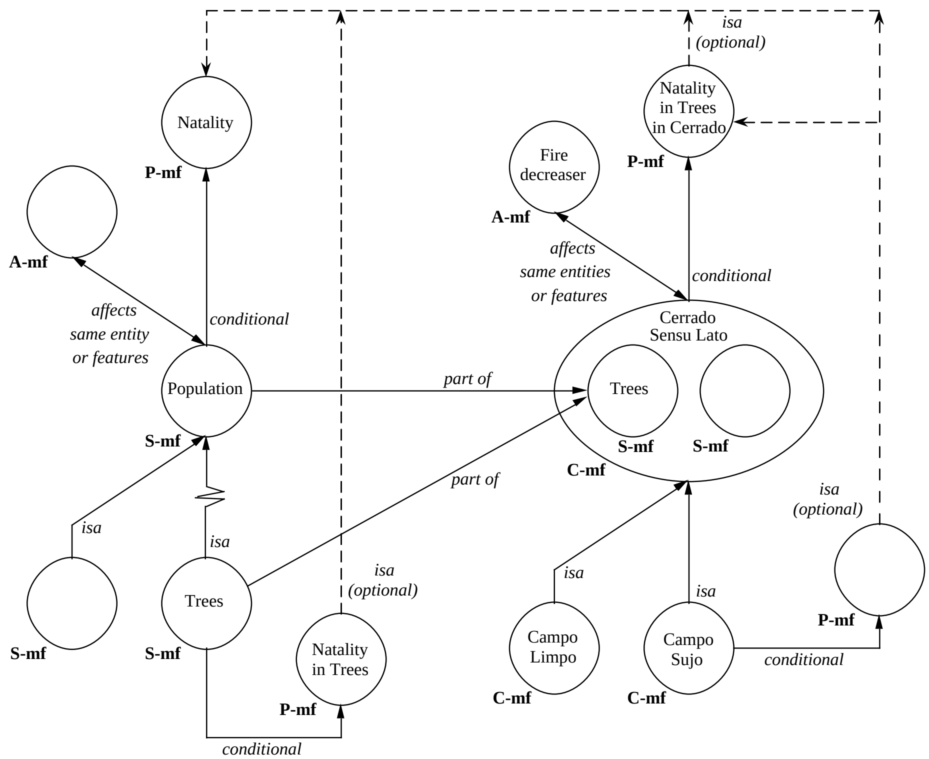

Figure 1: Technical organization of fragments in Cerrado Model

Order (O) The order of a model is defined as zero, first or second, mainly following CMP. However, a strict zero-order model, in which 'values are on or off' does not make much sense when discussing ecological models. The notion is therefore widened, in the sense that a quantity can have different values (e.g. low, medium or high). This allows for discussing different kinds of ecological situations. For example, a Campo Limpo in which certain populations are active (the value is on, using CMP terminology), forming a community because the populations differ in sizes (one large, the other small, etc.) and how that differs from another community in which the same populations exist, but with different sizes. First-order models include changes, which means processes or agents will be active. Moving to a first-order model is an important step, because it introduces parts of the causality that explain the systems behavior. If multiple processes (and/or agents) are active it may be known that certain processes are stronger than others, and thus that the system evolves is a particular direction. For example, when both mortality and natality processes are active, but the latter is bigger, the population increases. Second-order models represent relative changes, e.g. both immigration and emigration decrease, but the former decreases faster.

Structural change (SC) Often in explaining ecological systems there is the need to switch between situations in which 'different' entities exist. For example, after explaining the basic behavior of a single population, one may want to move to discuss 'competition' which requires the existence of at least two populations. Switching between such situations is a structural change. In terms of the simulation model a structural change always requires a modification of the entities present in the scenario that triggers the simulation. Thus structural change involves adding, or removing, entities into the systems. Structural changes have no counterpart in GG, CMP or DGG, but are crucial for explaining particular ecological concepts.

The dimensions listed above provide 'natural' constraints to further organize the set of possible simulations. Notice that moving along the dimensions **G/S**, **A** and **I** always involves *one* super-type and its immediate subtypes, whereas moving along the dimensions **O** and **SC** always introduces a new primitive (e.g. a process or a new population). To exploit this distinction we use the notion of *clusters*. **G/S**, **A** and **I** dimensions exist within a cluster, **O** and **SC** dimensions exist between clusters. Second, by definition it now follows that clusters are always of a certain order (zero, first, or second). Going to a higher order cluster requires an **O** change and moving to a more complex cluster, of the same order, requires a **SC** change. Third, a partial ordering among the clusters follows automatically. A zero-order cluster always precedes the adjacent first-order cluster. For instance, there is no point in discussing a natality process until the involved population has been discussed. Similarly, a more complex zero-order model (e.g. Cerrado Sensu Lato) can only be discussed after the three populations involved have been introduced (i.e. tree, shrub and grass populations).

However, the sequence is not fully determined, that is, not all aspects within one cluster have to be dealt with before someone can move on to another cluster.

Following the principles discussed above, the Cerrado Succession Model consists of six clusters[1] (figure 2). A typical progression would first address the zero-order cluster for single populations (C1). An **O** step would then lead to the first-order cluster for those populations (C2). Next, a **SC** step from C1 would lead to the zero-order cluster for two populations (C3), and after that an **O** step would lead to the first-order cluster for those populations (C4). A **SC** step from C3 then leads to the zero-order cluster for the Cerrado Sensu Lato community (three populations) (C5), and finally, an **O** step leads to the first-order cluster for that community (C6) (i.e., the full succession simulation model).

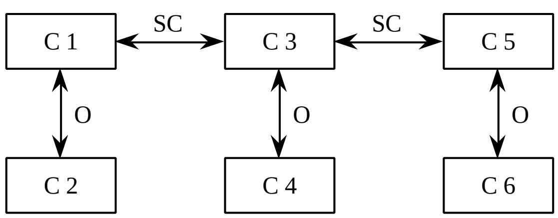

Figure 2: Clusters and organization in Cerrado Succession Model

Ordered Scenarios and Simulation Models

In this section we present implemented examples of model progression along the six clusters discussed in the previous section. Each of the six clusters that can be traversed along the Structural Change and Order dimensions is commented (see also figure 2).

C1: Classifying Populations The starting point for the model progression is this cluster of zero-order models. It encodes knowledge about general features of single populations (no dynamic aspects). Specializations of the population concept may represent plant populations and different types of plant populations (tree, shrub, or grass). This is a typical descriptive set of models. Ecological concepts represented in this cluster are: plant, tree, shrub, grass, population, plant population, (tree, shrub and grass) populations. General knowledge may also be represented (e.g. the importance of grass biomass as fuel for burning). The main educational goal to be achieved in this cluster is to define what kinds of populations exist and what their characteristics are. Within the cluster, questions follow from the dimension **G/S**. There are no **A** movements because in this cluster we are considering just one population. Also, no **I** steps exist because inverse is reserved for opposite influences. Questions concern knowing: What is X? Why Y is a specialization of X? For instance: What is a tree population? (A plant population); What are the characteristics of grass? (Its biomass is highly flammable in dry season).

C2: Single population dynamics From C1 we can move to C2 using the **O** dimension. Using this group of first order models, it is shown what active processes enforce changes in a population. This knowledge about processes is applicable to general populations, plant populations, tree populations and so on, following the definitions of C1. Ecological concepts represented in this cluster are: (a) populations and directions of change; (b) basic population processes; natality, mortality, immigration, emigration and population growth; closed and open populations (c) more specific processes of plant populations such as germination; establishment; colonization; seed dispersal. The main educational goals to be achieved in this cluster are to discuss the general population growth laws; to identify the basic natural processes that cause changes to any population; and to discuss some specializations of the basic population processes. Within this cluster questions explore the dimensions **G/S**, **A** and **I**. Also dynamic changes allow for queries about quantity values and relations between quantities changing over time. Questions within this cluster typically take the following forms. What are the differences /similarities between processes X and Y? How can the effects of X and Y be compared? What is the size of population X in state Y? Is the value of X increasing? Is X bigger than Y? What are the primary causes of changes in X? How can changes in X propagate to Y? For instance: What is colonization? (It is a specialization of the immigration process); What happens if natality is bigger than mortality in a closed population? (Its size increases); What is the value of 'number of' shrub in state 4? Why does the value of 'number of' tree decrease from state 2 to 3?

C3: Classifying two interacting populations Moving from C1 to C3 models progress on the **SC** dimension. This cluster of zero-order models concerns the structure of models involving two populations that interact. Ecological concepts represented in this cluster are: communities, symbiosis, comensalism, parasitism, predation, herbivory, neutralism, amensalism, competiton. Educational goals to be achieved include: to demonstrate how two populations may affect each other features or some natural resource; to demonstrate that influences coming from other populations are related to the basic population processes. Questions about the models in C3 explore the **G/S** and **A** dimensions about the knowledge represented in clusters C1 and C2. Comparisons between two populations are possible within this cluster, using the **A** dimension. For example: What are the differences or similarities between X and Y? New concepts support queries such as: How do populations X and Y use natural resource Z? Examples of questions are: What do tree and shrub populations have in common? (They are analogous/they are plant populations); What are

[1] In fact nine clusters, if the second-order cluster is fully separated from the first-order.

the differences between tree and grass populations? (Size, potential combustibility).

C4: Dynamics of two populations C4 is an **O** change from C3. Simulations in this cluster show how two populations interact and how their values change simultaneously. Ecological concepts represented in this cluster are the same as in cluster C3. However, the learner can see the dynamics involved in these relations and notice that they account for community changes. The main educational goals to be achieved in this cluster are: to establish different types of how two populations interaction; to demonstrate how the values of quantities representing population size change simultaneously. Simulations show changes on the sizes of both species. So a simulation involving a symbiosis relation (e.g. plant and bacteria that produces Nitrogen-based organic matter) should show that when plant increases, its positive influence causes the bacteria population to increase. Questions may explore the dimensions **G/S**, **A** and **I**, and knowledge represented in clusters C1, C2 and C3. New concepts may support queries such as: What will happen to population X, if population Y increases? What happens to mortality in population X, if emigration in Y increases? Some examples are: What does happen to a population of preys if its predator species increases? (Prey population decreases); What happens to natality of parasites if immigration of hosts increases its population? (It may increase)[1].

C5: Classifying Communities Progression from C3 through the **SC** dimension leads to the structurally more complex C5. This cluster of zero-order models elaborates on the concept of community by using representations of three populations. Cerrrado communities are defined as consisting of tree, shrub and grass populations, with different sizes. Ecological concepts represented in this cluster are: Cerrado Sensu Lato, Cerradão, Cerrado Sensu Stricto, Campo Cerrado, Campo Sujo and Campo Limpo. The main educational goals to be achieved in this cluster are: to illustrate different types of Cerrado communities; to compare different community types defined in terms of the values of quantities representing population size. Questions may explore the dimensions **G/S** and **A**, and knowledge represented in previously clusters. New concepts may support queries such as: What kind of community is X if it has a tree population size T, a shrub population size S, and a grass population size G? How do communities X and Y compare? Some examples are: What populations form the typical Campo Cerrado community? (Grass, shrub and tree populations); What are their qualitative values? (grass = shrub = tree = 'medium').

C6: Community dynamics Moving from C5 along the dimension **O**, we reach C6. This cluster of first-order models represents the behavior of Cerrado communities. Environmental factors such as cover, litter, temperature, nutrients, water, fire frequency and their influence on different plant species in the Cerrado are included in the models of this cluster. Note that the notion of ecosystem refers to the interaction between community and environmental factors. So, they could have been included in two population models (clusters C3 and C4) as well (a suggestion for future research). Exploring C6, it is possible for the learner to see the effects of things such as fire influencing other environmental factors and eventually affecting the basic processes involved in population growth. Ecological concepts represented in this cluster are: communities, succession, human actions (e.g. conservation), and causal relations involving the influence of environmental factors such as fire frequency. The main educational goals to be achieved in this cluster are related to the process of succession: to observe community changes due to the effects of human actions and natural processes; to understand causal relations between the environment and the basic population processes. Questions follow exploring the dimensions **G/S**, **A** and **I**. They take the form: What will happen to community X if environmental factor Y increases / decreases? Of course, the main question to be asked in this cluster is: What will happen to the Campo Limpo vegetation if human actions cause fire frequency to decrease? The answer is the full simulation, showing succession in the Cerrado vegetation, a process in which grass populations disappear and tree and shrub populations are introduced and become dominant, so that the Campo Limpo, evolving through all the other community types, becomes a Cerradão.

[1] Model fragments needed for this cluster are less well developed in the succession model presented by Salles & Bredeweg (1997).

Conclusion and Discussion

In this paper we have discussed progressive learning routes through large qualitative simulation models of ecological systems using ideas on model dimensions from Causal Model Progression (CMP), the Genetic Graph (GG), and the Didactic Goal Generator (DGG). The conceptual integration of the three approaches turned out to be essential for that purpose. CMP allows for distinguishing between simulation models of zero-order, first-order, and second-order (**O**). Within each set of models of a particular order, GG and DGG provide the vocabulary to talk about concepts in a classification hierarchy from a tutoring point of view (i.e., generalization/specialization (**G/S**), analogy (**A**), and inverse (**I**)). To complete the set of dimensions, we introduced the notion of structural change (**SC**), which we regard as a dimension being of particular importance for reasoning about ecological systems. With these five dimensions large qualitative simulations can be reorganized into clusters of smaller simulation models. Each cluster is then always of a particular order (in our examples mainly zero-order and first-order). Zero-order models capture the static features of an ecological system, whereas first-order models represent the dynamics of such a system. The details within each cluster are organized using the dimensions **G/S**, **A**, and **I**. Using the **O** dimension we can move from static to dynamic models (and vice versa). With the **SC** dimension we can increase the complexity of the ecological system being modeled and for instance progress from populations, via communities, to ecosystems (and vice versa).

The presented approach is implemented as a series of scenarios that can be run using an already existing qualitative simulation model, the Cerrado Succession Model, that expresses a widely accepted hypothesis about vegetation dynamics. Following the dimensions discussed in this paper the Cerrado Succession Model is reorganized into six clusters. Specific scenarios have been constructed to run simulations within each cluster.

Qualitative reasoning engines typically use a library of model fragments for generating simulations. The definition of the model dimensions presented in this paper is related to the way model fragments are organized in such a library. An interesting next research step would be to automatically generate learning routes of clusters from such a library. A related aspect would be to use the dimensions as a basis for explanation generation. Further work could also investigate whether the dimensions can successfully be applied to new domains. Although further refinement may be needed, we strongly believe that this should be possible for qualitative simulations in which processes play an important role.

References

Bouwer, A., Bessa Machado, V., & Bredeweg, B. *Interactive Model Building Environments.* Pages 1-23. (forthcoming).

Bredeweg, B. 1992. Expertise in qualitative prediction of behaviour. Ph.D. thesis. University of Amsterdam, Amsterdam, The Netherlands.

Forbus, K.D. 1984. Qualitative process theory. *Artificial Intelligence*, 24:85–168.

Goldstein, I.P. 1979. The Genetic Graph: a representation for the evolution of procedural knowledge. *International Journal of Man-Machine Studies*, 11:51-77.

Gotelli, N.J. 1998. *A primer of Ecology* (2nd edition). Sinauer Associates pub. Sunderland, MA.

Kearsley, G. 1994-2000. Explorations in Learning & Instruction: The Theory Into Practice Database. TIP: http://www.gwu.edu/~tip/

Moreira, A. 1992. *Fire Protection and Vegetation Dynamics in the Brazilian Cerrado*. Ph.D. thesis, Harvard University, Cambridge, Mass.

Pivello, V.R. 1992. *Expert system for the use of prescribed fires in the management of Brazilian savannas.* Ph.D. thesis. Imperial College of Science, London.

Pivello V.R. & Coutinho L.M. 1995. A successional model to assist on the management of Brazilian cerrados. University of Sao Paulo, Department of Ecology, Pages 2-22 (unpublished manuscript).

Salles, P.S.B.A. 1997. *Qualitative models in ecology and their use in learning environments.* Ph.D. thesis, University of Edinburgh, Edinburgh, Scotland, UK.

Salles, P., & Bredeweg, B. 1997. Building qualitative models in ecology. *Proceedings of the International workshop on Qualitative Reasoning, QR'97.* Istituto di Analisi Numerica C.N.R. Pavia, Italy, L. Ironi (ed.). pages 155-164.

Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D.K. 1988. Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. *Proceedings of the tenth annual meeting of the cognitive science society*, pages 375-383, Hillsdale, NJ, Lawrence Erlbaum Associates.

White, B.Y., & Frederiksen, J.R. 1990. Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, 42:99–157.

Winkels, R., & Breuker, J. 1993. Automatic Generation of Optimal Learning Routes. *Proceedings of International Conference on Artificial Intelligence and Education*, 330-337.