AI Communications 16 (2003) 291–308
IOS Press

Qualitative models of interactions between two populations

Paulo Salles [a], Bert Bredeweg [b], Symone Araujo [a] and Walter Neto [a]

[a] *Universidade de Brasilia, Instituto de Ciências Biológicas Campus Darcy Ribeiro, Brasilia – DF, 70.910-900, Brasil*
E-mail: paulo.bretas@uol.com.br
[b] *University of Amsterdam, Department of Social Science Informatics Roetersstraat 15, 1018 WB Amsterdam, The Netherlands*
E-mail: bert@swi.psy.uva.nl

Abstract. Ecological knowledge is often characterised as being incomplete, sparse and non-formalised. Qualitative reasoning provides means to capture such knowledge that is otherwise difficult to represent in computer programs. An additional feature is that qualitative models can be used to run interactive simulations in learning environments, providing opportunities for learners to acquire causal insights about ecological phenomena. In this paper we present qualitative models of interactions between two populations in biological communities. Our approach further explores a qualitative theory of population dynamics previously implemented. Based on this theory we have developed and implemented qualitative models and simulations that support reasoning about the most common behaviours of two interacting populations. In our models the assumptions are explicitly represented and therefore can be analysed by students and modellers. We also discuss how these models can be organised to create interesting learning routes for teaching learners about population and community behaviour.

Keywords: Population ecology, qualitative modelling and simulation, learning environments

1. Introduction

Qualitative simulations are detailed and articulate *knowledge* models that represent insights humans have developed of systems and their behaviour. Such knowledge models are interesting, both from an ecological and from an educational perspective. This article presents qualitative models and simulations of interactions between two populations in biological communities. In addition, it discusses the organisation of such simulation models into clusters of increasing complexity in order to facilitate their use in interactive learning environments.

Historically, in the field of ecology, models about two populations are based on the logistic or related equations (cf., [10,11]). However, such numerical means may not always be suitable for representing ecological knowledge, because ecological knowledge is often incomplete. Qualitative models provide new opportunities for articulating ecological knowledge. Particularly, to represent those aspects which are normally hard to capture in quantitative models, such as definitions of objects and situations, representation of modelling assumptions, and explanations based on causal relations.

From an educational point of view, qualitative models can be used to generate simulations that form the basis for interactive learning environments. The construction of such *articulate* simulations is of particular interest, because they facilitate "knowledge communication" between the agents involved in the learning process [3,6]. The latter even more now that graphical tools have been developed that support learners in inspecting qualitative simulations [1].

However, qualitative models and simulations of pairs of interacting populations do currently not exist. In this article we address this problem by developing such models. The presented simulations show the typical behaviours as reported in the literature and presented in textbooks. Our work further explores the qualitative theory of population dynamics as presented in [22–24]. Using this library of basic processes it is possible to derive complex community behaviour as illustrated by an implementation of the Cerrado Succession Hypothesis (CSH). The CSH models are based on ecological studies and theories concerning succession in the Brazilian Cerrado vegetation (cf., [4,17]) and capture knowledge about the structure and behaviour of Cerrado communities under the influence of fire and other environmental factors. Simulations based on these models reproduce the "common-sense theories" that domain experts have formulated concerning the vegetation dynamics in the Cerrado.

The models about population interactions presented in this article have been added to the models implementing the CSH, resulting in a large set of models. However, when such a set becomes too complex it cannot be used effectively for teaching purposes. It has to be rearranged into smaller parts and ordered in a sequence for learners to progress through and gradually acquire more advanced insights (cf., [28]). The second part of this article investigates the organisation of qualitative simulation models in the domain of ecology for teaching purposes.

2. Ecology of interactions between two populations

Communities are defined as sets of populations of different species living in the same space during the same period of time. Most ecologists believe these species are not randomly associated, and how communities are structured has been subject of an intense debate. Negative and positive interactions between populations such as competition, predation and symbiosis, working under the influence of physical factors of the environment, have been pointed out as the main organising forces of communities. Particularly, the role of competition and predation in organising communities has been emphasised by many authors (cf., [13,19]).

Relationships between populations of different species can be classified either on the basis of the *mechanism* or on the *effects* of the interaction. Mechanisms of interaction take into account particularities of each species life cycle. When these details are left out, and just the effects considered, interactions can be classified according to combinations of the symbols $\{-, 0, +\}$: "$-$" means that one population is adversely affected by the other; "0" means that one population suffers no effects from the other; and "$+$" means that one population benefits from the other. If the interaction is to be modelled as an equation, the meaning of the symbols is to add a positive or a negative term to the growth equation of both populations. Mechanisms and effects of main interactions between two populations as described by [19] are summarised below. The format used is: *Type(A, B)*: *Description*. "Type" refers to the name of the interaction. "A" refers to the effect the interaction has on population one (written as: population1). "B" refers to the effect on population two (written as: population2).

- Neutralism (0, 0): Neither of the populations affects the other population.
- Amensalism (0, $-$): Population1 inhibits population2, in general by producing some toxic substance (and population1 is not affected).
- Commensalism (0, $+$): Population1 benefits population2, in general providing food or transport (and population1 is not affected).
- Predation ($+$, $-$): Population1, the predator, causes harm to population2, the prey, and benefits from the interaction; often the predator is bigger than the prey and less numerous.
- Parasitism ($+$, $-$): Population1, the parasite, causes harm to population2, the host, and benefits from the interaction; often the parasite is smaller than the host and more numerous.
- Herbivory ($+$, $-$): Population1, the herbivore, causes harm to population2, the plant, and benefits from the interaction; this involves eating fruits, seeds, leaves and other parts of the plant.
- Protocooperation[1] ($+$, $+$): Both populations benefit, but it is a non-obligatory interaction.
- Mutualism ($+$, $+$): Both populations benefit, it is an obligatory interaction (for one or both populations).
- Competition by interference ($-$, $-$): Each population is inhibited directly by the other population.
- Competition by resource exploitation ($-$, $-$): Both species have the same requirement and the availability of this common resource is limited (indirect inhibition).

Models representing interactions between populations are useful for the development of conservation strategies in natural ecosystems, or in programmes of recuperation of degraded land. For example, according to Morosini and Klink [18], "molassa grass" (*Melinis minutiflora*) is one of the most aggressive invading species in the Brazilian Cerrado vegetation. This African species can cause disruptions in the invaded area and benefits from fire. It has been shown that after burning, *Melinis* occupies the space leaving out native species. However, shaded by trees like *Cecropia* the grass can be eliminated. These interactions can be seen as examples of competition between *Melinis* and native species, and between *Cecropia* and *Melinis*.

[1] Protocooperation and mutualism are also called *symbiosis*.

Positive interactions such as commensalism and symbiosis are also reported in the literature about the Cerrado. Mendonça and Piratelli [16] fed animals from eight vertebrate species on fruits and made germination tests with the seeds. They found some good potential dispersors of seeds, like monkeys from genus *Cebus*. Nearly all of the studied species had increased seed germination after passing through the animal digestive system. The relationship can be seen as proto-cooperation, with positive effects for the plants (dispersion and germination) and for the animals that feed on them.

Interactions between two populations may change over evolutionary time, under different conditions or in different stages of the life cycle. This is for instance the situation involving insects of two orders (Hymenoptera and Lepidoptera), described by [26] in the Cerrado. Hymenoptera are *parasitoid* insects, i.e., larvae of Hymenoptera kill caterpillars (larvae of Lepidoptera) but do not affect adult Lepidoptera. This way, interaction between these insects can be described as follows: (a) larvae of Hymenoptera and larvae of Lepidoptera, $(+, -)$; (b) larvae of Hymenoptera with adults of Lepidoptera, (0, 0); adults of Hymenoptera with larvae or adults of Lepidoptera, (0, 0).

In summary, communities can be seen as complex webs of relationships and interactions between pairs of populations. Qualitative models of these interactions may help ecologists to understand how communities are structured and to explain the behaviour of interacting populations in terms of underlying population processes.

3. Qualitative models of interactions between populations

The work presented in this article further explores the qualitative theory of population dynamics presented by [21,23,24]. They describe a fully implemented qualitative model, referred to as the Cerrado Succession Hypothesis (CSH). CSH simulates a common sense theory about the succession of communities as formulated by ecologists for the Cerrado vegetation in central Brazil and that has received support from scientific studies (e.g., [4,17]). The Cerrado vegetation consists of many different physiognomies spanning from open grassland to a more or less closed forest.

According to a widely accepted hypothesis, changes in the fire frequency determine the composition of the Cerrado vegetation. If the fire frequency increases above "natural" levels, woody components of Cerrado communities decrease and graminoid components increase, so that the vegetation becomes less dense. If the fire frequency decreases the vegetation tends to become woody and denser. Experts argue that germination and survival of young plants of tree species are more likely to occur in shaded, cold and moist microenvironments, whereas grass species do better in illuminated, warmer and dryer microenvironments.

An important characteristic of the CSH model concerns the use of "basic processes" (*natality*, *mortality*, *immigration* and *emigration*) that determine the behaviour of each population in a community and from which the overall behaviour of the Cerrado community is derived. For the research presented in this article we also use this qualitative theory based on basic processes to simulate and explain different kinds of interactions between two populations. Below, we first present the basic processes by discussing the behaviour of a single population. Next, we extend this approach by introducing a general schema for modelling interactions between two populations. This is followed by a discussion on the specific interaction types and how they are implemented using this schema.

3.1. Single population behaviour

The models discussed here are implemented in GARP[2] [2], a domain *independent* reasoning engine that implements a compositional modelling approach [6] to qualitative simulation. The engine works on the basis of three constructs: scenarios, model fragments and transition rules. Scenarios specify initial situations for the simulator to start a behaviour prediction. Model fragments capture knowledge about the structure and behaviour of (partial) systems and are used to assemble states of behaviour. Assumptions may be used to further detail the applicability of a model fragment. Transition rules determine valid transitions between states of behaviour. After selecting a scenario the reasoning engine proceeds with the prediction task by recursively consulting the library of model fragments for applicable fragments. This search is exhaustive and each consistent subset of applicable model fragments represents a behaviour interpretation that matches the selected scenario. How many interpretations will be found depends on the kind of scenario,

[2]The software and models can be downloaded from: http://www.swi.psy.uva.nl/projects/GARP/.

particularly on the amount of detail and constraints that have been specified in it.

Following the compositional modelling paradigm an important goal is to construct model fragments that represent elementary behavioural units. These building blocks should reflect the basic concepts and principles for a particular domain from which the behaviour of more complex systems can be explained. To capture the insights that ecologists have concerning the behaviour of populations, the CSH model is based on the "growth" equation that is typically found in textbooks on ecology:

$$\mathit{Nof}(t+1) = \mathit{Nof}(t) + (B + \mathit{Im}) - (D + E)$$

In this equation *Nof* represents the number of individuals of the population at the beginning (t) and at the end of some time interval ($t+1$). B, D, *Im*, and E represent the amount of individuals being born, that die, immigrate and emigrate during that interval, respectively. After introducing this equation, ecological textbooks spend quite a few pages on explaining issues relevant for understanding this equation. For instance, that B, D, and E are functions of *Nof* and that the precise shape of these functions may be different for different species while immigration (I) seldom depends on the number of individuals already present in the population.

Causal dependencies, as introduced by the Qualitative Process Theory (QPT) [7], can be used to capture such knowledge in a qualitative model. Similar to QPT GARP allows the use of positive and negative *direct* influences (I+, I−), and positive and negative *indirect* influences (P+, P) (the latter are also referred to as proportionalities). The flow of individuals being born is typically captured by an I+, meaning that "there is a flow (rate) of B that causes the *Nof* to increase", thus: I+(*Nof*, *B*)[3]. Next, the fact that "changes in the *Nof*, in a particular direction, cause changes in the flow of B, in the same direction" is typically represented by a P+, thus: P+(*B*, *Nof*). Following this approach we can define four basic processes: "natality" (B), "mortality" (D), "immigration" (I) and "emigration" (E), each modelled by a separate model fragment. The quantities B, D, *Im*, and E can be seen as the *rates* of these processes. But notice that there are differences on how they relate to *Nof*. Flows originating from mortality and emigration have a *negative* direct influence on the size of the population; therefore: *I*−(*Nof*, *D*) and I−(*Nof*, *E*). And there is no indirect influence from *Nof* on *Im*, because immigration is seen as being independent of the population size.

An important aspect of a qualitative model concerns the values quantities can have: the quantity spaces. In GARP quantity spaces are uniquely defined *per* quantity and consist of an ordered set of alternating points and intervals. A quantity space may include 0 (zero), which is universal for a model, that is, all 0's are equal. Relationships between other values from different quantity spaces can be defined using (in)-equalities and correspondences. Most of the simulations presented in this article use a three-valued quantity space for *Nof*: QS = {zero, normal, max}, referring to: the population does not exist (*Nof* is zero), the population exists and has a "normal" size, and the population has reached its maximum size. A different perspective may require a different range of values. For instance, to characterise and discriminate between different types of Cerrado communities Salles and Bredeweg [23] use a five-valued quantity space for *Nof*: QS = {zero, low, medium, high, max}. Magnitudes of B, D, *Im*, and E are represented by the values zero and plus (a positive interval), thus QS = {zero, plus}. Derivatives in general can take on the values minus, zero, and plus, represented as QS = {min, zero, plus}. Applied to the derivatives of *Nof*, and B, D, *Im*, and E, these symbols represent that the population and the flows from the basic processes are decreasing, stable or increasing.

A flow of individuals, as e.g., modelled by the "natality" process, does *only* occur when a population exists. Therefore a distinction must be made between situations in which a population exists and in which it does not exist. Two model fragments describe these situations: if $\mathit{Nof} > 0$, there is a population, described in the fragment "existing population", and if $\mathit{Nof} = 0$, there is no population, described in the fragment "non-existing population". The processes "natality", "mortality", "immigration", and "emigration" have the "existing population" model fragment as a condition, resulting in these processes not becoming active for "non-existing populations". However, in some situations, a new population can start to live in a place with the arrival of some individuals. This process is called "colonisation". Colonisation is modelled as a special kind of "immigration", namely one that starts a new population in a space where there is no such population. So the fragment "non-existing population" is conditional for the "colonisation" process to become active.

[3] Following conventions (cf., [7]), the influencing quantity is mentioned as the second argument.

When a quantity is directly or indirectly influenced by more than one process, their effects are combined by *influence resolution* [7]. In our case, B and *Im* are added, whereas D and E are subtracted from the derivative of *Nof*. The final result may be ambiguous depending on the relative amounts of these four rates. Ambiguity is sometimes seen as a problem of qualitative models, because the missing information may lead to enormous state-graphs predicting a large number of possible behaviours. We like to think of ambiguity as a feature, namely one that drives the knowledge acquisition. If ambiguity occurs in a simulation, domain experts can be further questioned about the missing knowledge. If that knowledge is available, it can be made explicit, modelled and the ambiguity can be resolved. If that knowledge is not available the ambiguity reflects the incomplete understanding that experts have of the domain. Notice that, behaviours resulting from ambiguity should not be confused with spurious behaviours (e.g., [14]). Spurious behaviours are undesired. They refer to incorrectly predicted behaviours that do not occur for the real system, but only appear in the model. Ambiguity, on the other hand, refers to alternative behaviours predicted because information is lacking, but given what is know they are *correct* behaviours.

In the case of the processes that govern population behaviour extra information can be represented concerning the relative magnitudes of the flows. This is achieved by aggregating quantities in order to get a different perspective on population growth. As in most mathematical models, we define the "growth" process as an aggregation of the four basic processes, using the intermediate variables *Inflow* and *Outflow* to define growth rate (*Growth*). The qualitative growth equation then becomes:

$$\begin{aligned} \mathit{Inflow} &= B + \mathit{Im} \\ \mathit{Outflow} &= D + E \\ \mathit{Growth} &= \mathit{Inflow} - \mathit{Outflow} \end{aligned}$$

The overall population growth is modelled using a new model fragment, "population-growth", that also introduces the causal dependencies I+(*Nof*, *Growth*) and P+(*Growth*, *Nof*). Different from the four basic processes the quantity *Growth* requires QS $=$ {min, zero, plus} to take care of situations in which *Inflow* is smaller, equal or greater than *Outflow*.

To run simulations that capture different perspectives on population behaviour, assumptions may be used. In our models we use *simplifying* and *operating* assumptions along the lines suggested by [6]. In GARP, assumptions are implemented as labels that affect the applicability of model fragments. If a model fragment becomes active, because an assumption applies, it usually results in additional constraints that have to be taken into account. A typical operating assumption in population ecology is the notion of *open* versus *closed* populations, referring to the situation in which migration occurs or does not occur, respectively. To capture this idea two assumption labels are defined: *open-population* and *closed-population*. Particularly, the "closed-population" fragment always applies when the *closed-population* assumption is active. It excludes migration by specifying that both magnitudes and derivatives are equal to zero ($\mathit{Im} = E =$ zero and $\partial \mathit{Im} = \partial E =$ zero).

Having defined a library of model fragments, scenarios can be specified to run specific simulations. VISIGARP [1] implements a graphical user interface on top of GARP and can be used to control and inspect the simulations. Figure 1 shows some of the results produced by VISIGARP when running a simulation of a single population with *unknown* initial values for all quantities. In the case of unknown values GARP may assume values for certain quantities if there are applicable model fragments that capture knowledge in this respect (for details see [2]). For this scenario, GARP assumes values for the quantity *Nof*. Following this, values are derived for other quantities, as well as possible states and state transitions.

Figure 1 shows different kinds of simulation results, notably the state-graph (right top), the value history (right middle), and part of the causal model (left and bottom). A state-graph starts with a scenario (grey circle, named "input") and shows the qualitative distinct states of behaviour that the system can manifest (numbered black circles). It also shows which states of behaviour succeed each other (arrows between circles). In the example, there are six qualitative states. They implement a single path of behaviour, no branching, starting at state 8 and ending at state 6 $\{8 \rightarrow 5 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 6\}$.

A VISIGARP user can select a set of states from the state-graph and open the value history. The quantities and their values are then shown in the order in which the states have been selected. In each state all quantity values are represented as magnitude/derivative pairs, Value $=$<*mag*, *der*>, representing current value and direction of change, respectively. For instance, in state 8, *Nof* is represented as <*max*, *min*> (that is: it

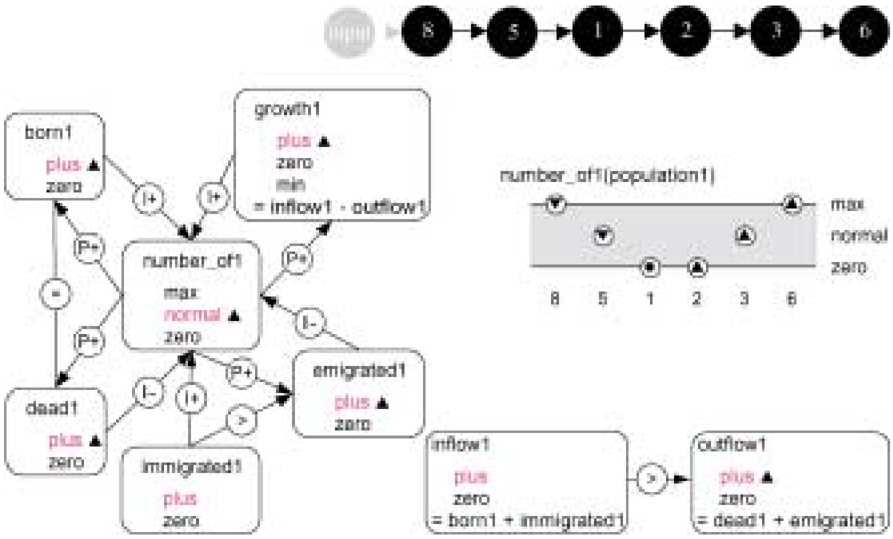

Fig. 1. Basic processes influencing a population. Names used in the drawings relate to the text as follows: number_of (*Nof*), dead (D), born (B), emigrated (E), and immigrated (*Im*). The numbers 1 and 2 are used to distinguish between the two populations. Thus number_of1 refers to the *Nof* individuals of the first population and number_of2 refers to the *Nof* individuals of the second population (if there is a second one). The same procedure is used for the other quantities.

has maximum value and is decreasing). In Fig. 1 the states have been selected following the order of the state-graph (notice that a user may select a different sequence). Thus, the value history diagram actually shows the sequence of values in the successive states. The simulation starts with a maximum sized population (state 8). It decreases and, via state 5, becomes zero in state 1 (population became extinct). State 1 is followed by state 2 in which the population starts to grow again because of colonisation. After colonisation it continues to grow and via state 3 becomes full sized again, as represented by state 6.

The causal model underlying *each* state can also be inspected using one of VISIGARP's interactive screens. Figure 1 shows the causal model details for state 3. This view also shows the quantity space each quantity has as well as the specific values and derivatives in this state of behaviour. For instance, *Nof* has $<$*normal, plus*$>$ (that is: it has a normal value and is increasing, as shown by the scale and by the triangle respectively). Notice that each state has its own causal model, possibly different from the causal model in other states, depending on the model fragments that are applicable in each state of behaviour. In state 3 the four basic processes are all active ("natality", "mortality", "immigration", and "emigration"). Each of these processes influences the size of the population (e.g., the mortality process introduces a negative influence, I−, on the size of the population). Changes in the population affect the basic processes (e.g., when the size of the population increases, D (mortality) will also increase, P+). The figure also shows how different flows are accumulated into the total *Inflow* and *Outflow* and how these flows eventually determine the overall population *Growth*.

The simulation discussed here is only one of many alternative simulations that can be run. Such simulations may typically vary on assumptions, initial values, and initial inequality statements between the quantities involved. In a learning situation, students (working individually or in pairs) may be given assignments that they have to solve by inspecting simulations. E.g., explaining why colonisation does not happen in the case of a *closed-population*.

Using the library of model fragments, which implements a qualitative theory of population dynamics, it is possible to derive complex community behaviour from the underlying basic processes that can be seen as "first principles" in population ecology. The CSH model is an example of that (cf., [21,23,24]). Below we further detail the interactions between two populations using the same library of model fragments to derive population behaviour during the interaction.

3.2. *Base model of interactions between 2 populations*

Suppose there are two populations that do not interact. If there are no constraints, all the possible be-

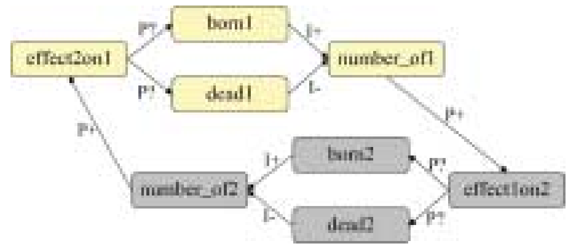

Fig. 2. Base model for representing interactions between two populations.

haviours (that each population alone can exhibit) are expected to appear in a simulation. Therefore, all the combinations of values (magnitudes and derivatives) of all quantities for the two populations will be found. However, when the populations are *not* independent, but interact and affect each other, we expect that some of these behaviours will be restricted. That is, not all behaviours can be expressed by both populations. Modelling these interactions means articulating the *constraints* that limit the set of possible behaviours for the two populations.

As a starting point we define the *basic interaction model* for capturing interactions between two populations. A simplified version of that model is depicted in Fig. 2. It includes the "natality" and "mortality" process for the interacting populations and relates the behaviour of the two populations via a new quantity: "*Effect*" (we use "*Effect*" to refer to both *Effect1on2* and *Effect2on1*). The idea is that the populations are influenced *via* their basic processes. Population1 produces some effect (*Effect1on2*), which affects "natality" (*B2*) and "mortality" (*D2*) of population2. In the same way, population2 produces an effect on population1 (*Effect2on1*) that in turn influences "natality" (*B1*) and "mortality" (*D1*) of population1. These influences are modelled using qualitative proportionalities, represented in Fig. 2 as $\{P+,\ P-,\ P?\}$. Notice the difference between P and I (see [7]). Influences (I) represent flows, which are *direct* influences. Proportionalities (P) propagate changes, which are *indirect* influences. The latter should be used here, because the size of the influencing population is *initially* determined by the basic processes of that population (which represent the direct influences). The changes in size are then propagated to the affected population, hence an indirect influence.

The attributes of interacting populations are thus represented by the quantities {*Nof*, *B*, *D*, *Im*, *E*, *Growth*, and "*Effect*"}. The quantity spaces associated to these quantities are: *Nof* has QS $=$ {zero, normal, maximum}; *B*, *D*, *Im*, and *E* have QS $=$ {zero, plus}; derivatives of all quantities and *Growth* have QS $=$ {min, zero, plus}; *Effect1on2* and *Effect2on1* have the same QS as *Nof*. Finally, the populations affect each other via their basic processes. This knowledge is modelled using indirect influences (qualitative proportionalities).

Having set up the basic architecture, each interaction type can be constructed following a set of modelling steps:

- Defining the quantities that represent the mutual interaction effects. For instance, in the case of predation the *Effect* of the predator on the prey can be called *Consumption* and the *Effect* the prey has on the predator *Supply*.
- Establishing causal links between the quantities *Nof*, *B*, *D*, and *Effect*. Does *Effect* influence both B and D, and what direction does it have for each of them? Notice that it follows from the base-model that the influence from *Nof* on *Effect* is always positive (see also Table 1).
- Defining assumptions that implement correspondences and possibly other constraints between the quantities *Nof*, *B*, *D*, and *Effect*. For instance, a simplifying assumption that we use in all the interaction models is the full correspondence between the *Nof* and the *Effect* it causes. Another simplifying assumption is to state that when the *Effect* influences *both* B and D that the impact will be the same for both processes (see also below).
- Representing conditions for non-existing populations. An interesting issue concerns the representation of things that "disappear" in the real world because of the system behaviour. Should those things also disappear from the model or should the model represent that they have disappeared? In the case of population interactions the non-existence of a population may have an influence on the behaviour of the other population and thus requires reasoning about something that does *not* exist in the real-world system. For example, the idea that the predator population cannot survive when the prey population is extinct.

Before detailing the points mentioned above for specific interaction types, a few issues must be discussed that are relevant to all interaction types. One issue concerns the establishment of the specific causal structure that implements the interaction. Table 1 shows the refinement of the base model for each interaction

Table 1
Main interaction types according to the effects on population growth

Interaction	Influences	Models
Competition (−, −)	Effect1on2 is negative	P−(Born2, Effect1on2)
	Effect2on1 is negative	P+(Dead2, Effect1on2)
		P−(Born1, Effect2on1)
		P+(Dead1, Effect2on1)
Amensalism (0, −)	Effect1on2 is negative	P−(Born2, Effect1on2)
	Effect2on1 none	P+(Dead2, Effect1on2)
Neutralism (0, 0)	Effect1on2 none	no relations between
	Effect2on1 none	the two populations
PredatorPrey (+, −)	Effect1on2 is negative	P+(Dead2, Effect1on2)
	Effect2on1 is positive	P+(Born1, Effect2on1)
		P−(Dead1, Effect2on1)
Commensalism (0, +)	Effect1on2 is positive	P+(Born2, Effect1on2)
	Effect2on1 none	P−(Dead2, Effect1on2)
Symbiosis (+, +)	Effect1on2 is positive	P+(Born2, Effect1on2)
	Effect2on1 is positive	P−(Dead2, Effect1on2)
		P+(Born1, Effect2on1)
		P−(Dead1, Effect2on1)

type as it is implemented in our models. Notice that, a *negative* interaction (e.g., a negative *Effect1on2*) can be represented as decreasing "natality" and/or increasing "mortality" of the affected population. Similarly, a positive influence increases "natality" and/or decreases "mortality" of the affected population. The choices, as shown in Table 1, are not the only possible solution and interactions between certain species may require small adjustments to this set. For instance, not all competition-based interactions may affect both natality and mortality for both populations. Some competitions only affect the mortality process of both populations. However, in our current implementation we have focused on what seems to be the most typical occurrence of the interaction type. Moreover, our models can easily be adapted to include other options in this respect as well.

An important aspect of modelling is the use of assumptions. We have found that some assumptions make sense in most interaction types. One observation to be made is that all interaction types concern non-migrating populations. Migration is apparently not relevant for understanding the interaction types. Hence, all our models include the "closed-population" assumption. This means that for the quantities E and *Im*, which are part of all models because they are part of the basic processes, it is assumed that their influence on the interactions is zero (and that this does not change): $E = $ <zero,zero> & $I = $ <zero,zero>. Consequently, the qualitative growth equation is only affected by "natality" (*Inflow* $= B$), and by "mortality" (*Outflow* $= D$), and the growth equation effectively becomes *Growth* $= B - D$.

A third issue concerns the relationships between *Nof*, B and D in one population. In principle, all kinds of variation is possible, e.g., *Nof* = <plus,plus>, $B = $ <plus,min>, and $D = $ <plus,plus>, or *Nof* = <plus,plus>, $B = $ <plus,plus>, and $D = $ <plus, min>, etc. However, these variations are usually of little importance to the typical behaviour of an interaction type. Thus, to simplify the matter, it is assumed that (1) the values of B and D are fully corresponding, and (2) that D and *Nof* always change in the same direction (in some interaction types B is also included). Another simplifying assumption concerns *Nof* in relation to the "*Effect*" quantity that influences it. In some interaction types they are assumed to be fully corresponding, whereas in other types it is assumed that the impact is less strong.

Notice that these simplifying assumptions have no effect on the typical behaviour resulting from the interactions. Using these assumptions only simplify the state-graph produced by the simulator; making it easier for model users, such as learners, to grasp the essential details captured by the models and the simulations they produce.

3.3. *Predator-prey model* (+, −)

The behaviour we want to represent with this model shows the predator population changing along with

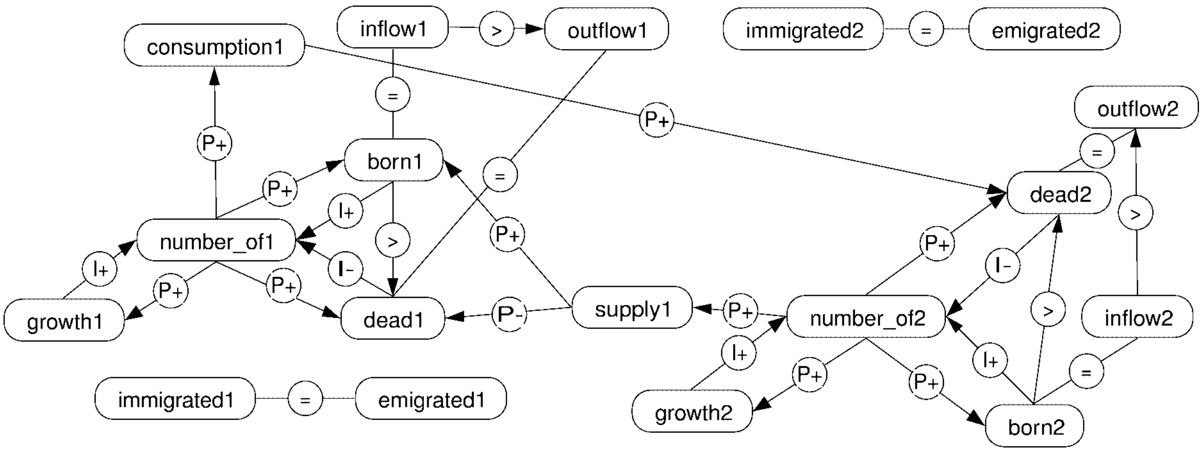

Fig. 3. Causal dependencies in the predator (population1) – prey (population2) model.

the prey population. To achieve that, we have to only slightly adapt the base model shown in Fig. 2: negative influence of the predator (population1) affects *only* the "mortality" of the prey (population2) (and not its "natality"). Figure 3 shows the dependencies as they actually appear in the simulation. As for all simulation results shown in this article, the figure is produced by VISIGARP [1]. Notice that this picture captures the same type of information as shown in the causal model part in Fig. 1 accept that due to space limitations not *all* the available information is actually shown Fig. 2.

The general constraints, as discussed above, imply that the *Consumption* of food (*Effect1on2*) fully corresponds to the predator population size, *Nof1*, and *Supply* of food (*Effect2on1*) fully corresponds to prey population size, *Nof2*. In fact, *Consumption* depends on many factors, such as the ability of predator to catch the prey and the availability of alternative sources of food. *Supply* depends on the ability of prey to avoid the predator and the existence of refuges in the environment. Thus, our assumptions implement an approximation of the full natural phenomena.

Specific predator-prey restrictions are that the predator cannot become bigger than the prey nor survive without it. This is modelled by stating that when *Nof1* is zero, so is *Nof2*. Complementary to that, it is assumed that the *Supply* has to be equal or greater than *Consumption*, and that the latter cannot increase faster than the former.

A simulation with this model is presented in Fig. 4. The state-graph is shown on the left and the value history is shown on the right. Notice that the latter enumerates the states in the order selected by the user. The order does thus not necessarily reflect a behaviour path. The actual behaviour paths are shown in the state-graph.

The state-graph shows the results of simulating a scenario (input) in which both populations start out at "normal" size and an unknown direction of change, thus *Nof* $=$<normal,?>. From this initial situation four interpretations are found, state 1, 2, 3 and 4. Each of these states is the start of a sub-graph representing one of the four typical behaviours of a predation situation:

- **Balanced co-existence**. In state 2 the two populations have a natural balance; they co-exist without further changes.
- **Populations grow to maximum**. State 1 leads to 10, optionally via 11, and shows the case in which both populations grow to their maximum size. Notice that the prey may reach its maximum size before the predator does (state 11), but not the other way around.
- **Populations get extinct**. State 4 leads to 6, optionally via 5, and shows the case in which both populations get extinct. The path via state 5 shows that the predator may become extinct before the prey, but not the other way around.
- **Predator gets extinct**. Finally, state 3 leads to 8, optionally via 7 or 9. It shows that the predator may get extinct without the prey getting extinct. Notice that the opposite is not possible.

3.4. *Competition by interference* (−, −)

The model of this interaction type should express a number of behaviours, including coexistence of the two populations and competitive exclusion of one of the two populations. Compared to the predator-prey

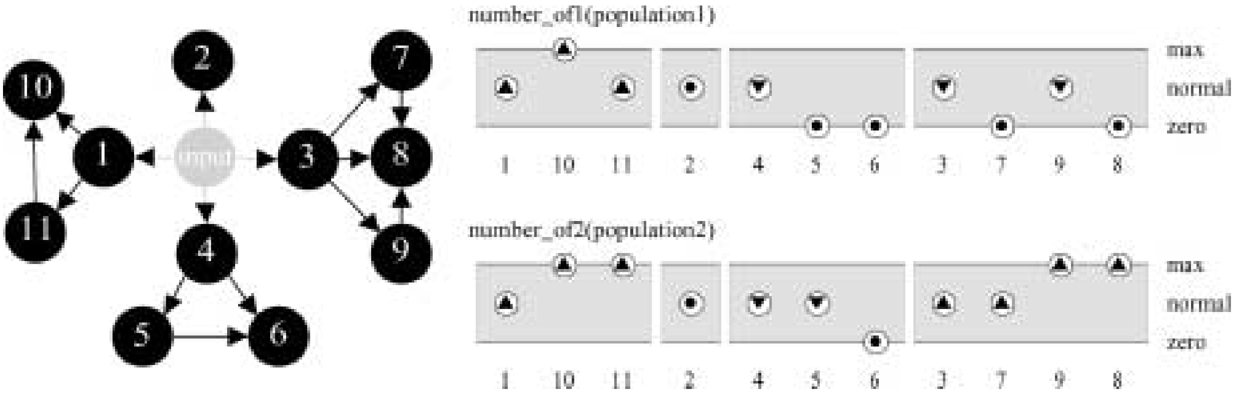

Fig. 4. Behaviour states for a scenario within the predator (population1) – prey (population2) model.

model, there is a small difference: the derivatives of B and D are assumed to be equal within each population. This means that the negative influence caused by one population is the same for both basic processes of the affected population. This assumption simplifies the state-graph, but does not affect the typical behaviour of this type of interaction.

The interaction quantity has not been given a specific name. Hence, *Effect1on2* refers to the harm done by population1 and *Effect2on1* to the harm done by population2. Both *Effects* implement a negative influence on the basic processes of the other population, namely decreasing "natality" and increasing "mortality".

A typical simulation running a competition scenario is shown in Fig. 5. In total, five types of behaviour are possible:

- **Balanced co-existence**. In state 3 the two competing populations have a natural balance. The competitive interaction is symmetric, that is, the mutual effects have the same magnitude and the populations thus co-exist without further changes.
- **Populations grow to maximum**. Despite the competition both populations grow to their maximum size. This interaction is also symmetric. The behaviour starts in state 2 and progresses to state 16, optionally going via state 15 or state 17.
- **Competitive exclusion**. Here the interaction between the competitors is asymmetric. The negative impact one competitor causes on the other population is bigger than the harm if suffers. Thus, one of the competitors becomes extinct while the other grows to its maximum size. As knowledge about which population causes most harm is not specified in the initial scenario, the simulator generates both options. Starting in state 2, and eventually leading to state 13, optionally via state 14 and 12, population2 becomes extinct while population1 grows to its maximum size. The behaviour starting with state 4, leading to state 10, optionally via 9 or 11, implements the opposite behaviour (population1 becoming extinct while population2 flourishes).
- **Populations get extinct**. State 5 leads to 7, optionally via 6 or 8, and shows the case in which both populations become extinct due to their competitive interaction.

An extension of the model above could include the influence of human actions or environmental factors. For example, suppose fire frequency decreases due to management practices. Under this condition, it may happen that *Effect1on2* > *Effect2on1* and that population2 is excluded (as in state 13). Alternatively, if fire frequency increases, *Effect1on2* may become smaller, changing the results of competitive exclusion (as in state 10). This can be used to predict the behaviour of the interaction between *Melinis* and *Cecropia* [18], mentioned above.

3.5. *Commensalism* (0, +)

This model has to represent that population2 increases when the population1 increases, without influencing the latter. In addition, we also want to represent, (a) situations in which the size of population2 (the one that receives the *Benefit*) is limited by the size of the population1 (the one that produces the *Benefit*), and (b) situations that express the effects of *Benefit* with different magnitudes. This is realised as follows:

- In order to limit combinations of possible population sizes, explicit associations involving the

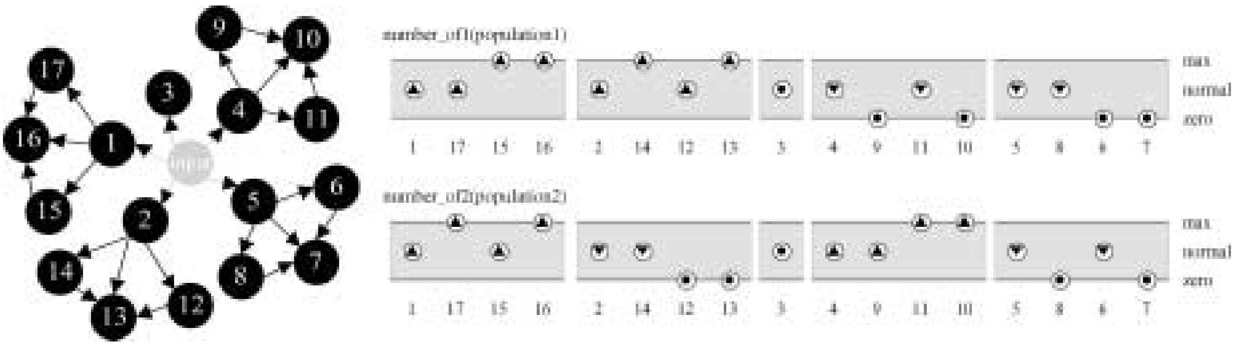

Fig. 5. Behaviour states for a typical competition interaction.

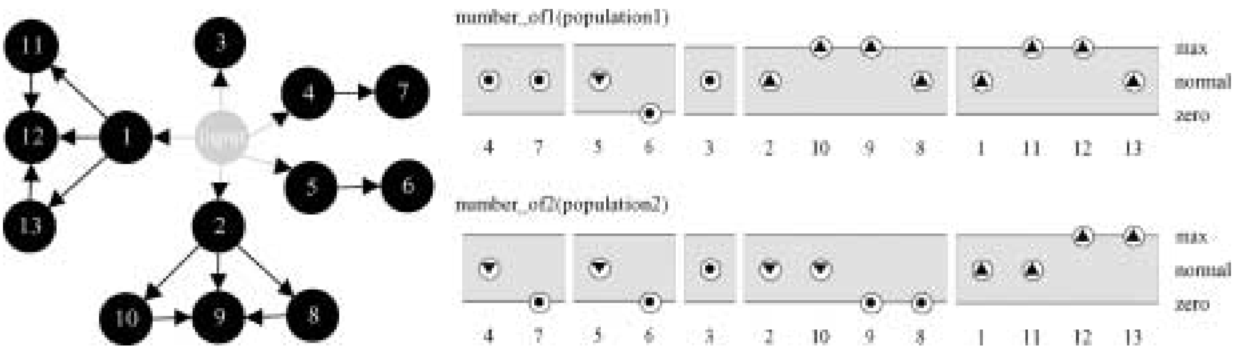

Fig. 6. Behaviour states for a "medium impact" commensalism scenario.

magnitudes of *Benefit* and *Nof2* are introduced in the model. For instance, *Nof2* can only have value "maximum" when *Benefit* is also "maximum". If *Benefit* has value "normal", population2 cannot reach its maximum size.

- To explore the strength of the effect of *Benefit* on *Nof2* we used relationships involving the derivatives. For instance, "medium impact" means that the *Benefit* is partially responsible for changes in population2. This is modelled by assuming that the derivative of *Benefit* is greater than or equal to the derivative of *Nof2*. "High impact" means that changes in population2 are fully determined by the *Benefit*. This is achieved by stating that the derivative of *Benefit* is equal to the derivative of *Nof2*.

The base model (Fig. 2) and the associated set of constraints are enlarged with more details in order to produce behaviour that satisfies the above mentioned requirements. Assumptions about the derivatives of B and D are the same for both populations and similar to those adopted in the competition model. In addition, it is assumed that the *Benefit* produced by the first population is essential for the survival of the second population. This way, if *Nof1* is zero, *Nof2* goes to zero as well.

Comparing simulations, it is possible to see how the assumptions influence the behaviour. A simulation under the "medium impact" assumption starting with the same initial scenario (both *Nof* $=$<normal,?>) produces five initial states, representing combinations of the derivatives of *Nof* in both populations (Fig. 6). However, in none of them the derivative of *Nof2* is greater than the derivative of *Nof1* (remember that the derivative of *Benefit* is determined by *Nof1*). The full simulation produces 13 states showing: both populations stable at normal size (state 3), both populations extinct (state 6, via 5), population1 normal and stable while population2 extinct (state 7, via 4), population1 with maximum size and population2 extinct (state 9, from 2 via 8 or 10), and both populations at maximum size (state 12, from 1 via 11 or 13). Changing the assumption to "high impact" reduces the number of possible states in the full simulation to 10.

3.6. Other models

The neutralism (0, 0) model shows the non-interaction, a situation in which there are no influences between the two populations. Starting with an initial scenario in which both populations have *Nof* $=<$normal,?$>$ the full simulation produces 25 states, showing all the combinations between the values of magnitude and derivative of *Nof*. Given that all the other interactions are modelled by adding constraints on "natality" and "mortality", this model can be seen as the "base simulation model" for interactions between two populations.

The amensalism (0, −) model introduces different impacts of the negative effects depending on the size of population1 (the one that produces the *pollution*). Simulations show that both populations can survive alone, and that population2 cannot become bigger than population1.

The symbiosis (+, +) model represents protocooperation (non-obligatory interaction) and the current implementation assumes that the positive effects on both populations are equal. Simulations show that they increase and decrease together and that both populations can survive alone.

4. Organising the model library for teaching purposes

The models that capture the knowledge about interacting populations are part of a rather large library that also contains the knowledge about the Cerrado Succession Hypothesis (CSH) [21,23,24]. When teaching a substantial complex domain, the subject matter must be divided into units, each unit dealing with a part of the whole, and ordered in sequence that can be traversed by the learner. Below we present arguments in terms of "knowledge characteristics" for effectively dividing a large set of models into "stand-alone units" (that is, full simulations by themselves) and for ordering these qualitative models of populations and communities for tutoring interactions. The approach integrates and expands ideas on model dimensions as discussed in Causal Model Progression (CMP) [28], the Genetic Graph (GG) [9], and the Didactic Goal Generator (DGG) [29].

4.1. Principles for organising the subject matter

CMP focuses on electronic circuits and defines three dimensions for models to vary: *perspective*, *order* and *degree of elaboration*. Perspective concerns the overall view of a system. For instance, functional, behavioural, or physical models. The dimension order further refines the notion of behaviour models. Typically, zero-order models are static, in the sense of not capturing continuously changing behaviour. In zero-order models quantities change values abruptly, such as a light bulb going from on to off. In first-order models behaviour changes gradually, such as a resistor gaining more resistance as power increases. Finally, second-order models include knowledge about relative speed of changes. For instance, one resistor building up resistance faster than another. The third dimension is degree of elaboration. Basically, it refers to the amount of inference detail that is required for deriving a particular behavioural fact. A model is more elaborated if it has more intermediate steps that must be reasoned about.

The GG uses four dimensions to classify elementary sub-skills (i.e., individual rules): *refinement*, *specialisation*, *generalisation* 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 (e.g., colour). A specialisation step refers to further detailing a concept: there are different ways in how it can manifest itself (e.g., there are different colours). A generalisation step is the opposite of a specialisation step (grouping different manifestations under a single concept). Finally, an analogy step refers to identifying other manifestations of the same concept.

DGG adapts the ideas presented in the GG. DGG defines generalisation/specialisation for organising concepts (with less/more attributes) in a hierarchy. *Inversion* refers to concepts being opposites (e.g., in text editors delete versus paste). 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 (e.g., how does a computer application work and how can it be used).

For organising models of ecological systems it turns out that the dimensions defined by GG and DGG provide 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 specialisation of the latter). A "natality process" is analogous to an "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 provide means to handle the “order of behaviour models”. For instance, we can distinguish zero-order (static) models from first-order models, in which things are changing. Summarizing, within the context of the former (GG & DGG) we can talk about the composition and structure of a community. The latter (CMP) can be used explain how active processes cause changes in communities (succession). The next section further discusses how to use the primitives discussed above to effectively divide and sequence qualitative simulation models of ecological systems.

4.2. Decomposing and ordering the CSH model

Libraries used by GARP consist of different types of model fragments (Fig. 7). A single description fragment (S-mf) models features of a single entity (or concept) (e.g., a tree, or a population) and is organised in a subtype (is-a) hierarchy (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. Process fragments can also be organised 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; 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 organised in subtype hierarchies. Next, P-mf and A-mf may apply to an assembly formed by a C-mf (e.g., a process that is only active in a Campo Sujo). P-mf influencing assemblies may again be organised in subtype hierarchies.

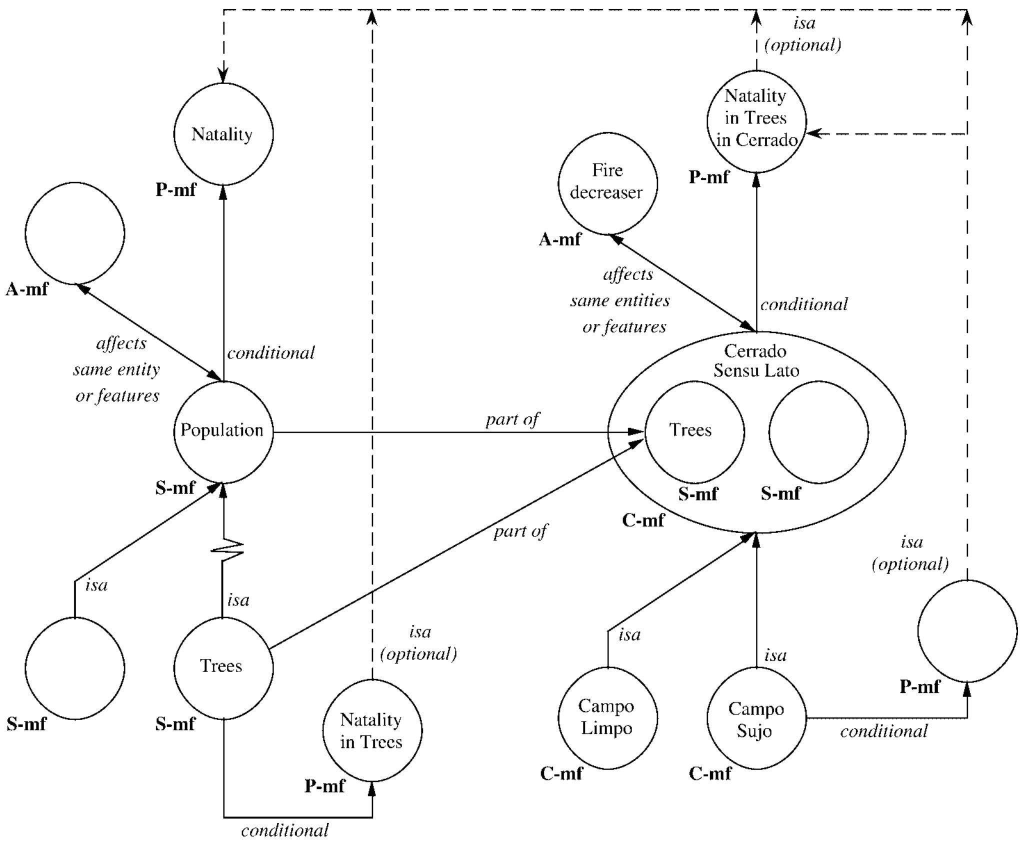

Fig. 7. Technical organisation of model fragments in GARP (illustrating the CSH model).

Combining the organisation of model fragments in GARP with the model dimensions discussed above, gives us the basis for constructing progressive learning routes. The following dimension can now be defined.

Generalisation/specialisation (G/S). The subtype hierarchy is used to organise model fragments on this dimension. A fragment is a specialisation of another fragment if it is a subtype of that fragment. A specialisation specifies at least a new name, but usually also introduces new features. Notice that features may come in many forms, such as quantities, causal dependencies, value ranges, etc. Generalisation is the opposite of specialisation. It refers to "moving-up" the subtype hierarchy. For instance, identifying shrub and tree as being both plant-populations (and e.g., different from animal-populations).

Analogy (A). Two fragments 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 they also differentiate on other features. For example, in the CSH models, the vegetation physiognomies referred to as Campo Limpo and Campo Sujo consist of similar kinds of plants, but each Campo type is characterised by the size of each plant population type.

Inverse (I). Similar to DGG we define inverse as a special kind of analogy. Namely, when two immediate subtypes have opposite features. In terms of ecology this means processes or agents with opposite behaviour (in fact, opposing influences). For example: 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).

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 which is characterised by the specific sizes of these populations (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 explains the behaviour. 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 in 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 when explaining ecological systems there is the need to switch between situations in which "different" entities exist. For example, after explaining the basic behaviour 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 set of the entities present in the scenario that triggers the simulation. Thus a structural change involves adding, or removing, entities. Structural changes have no counterpart in GG, CMP or DGG, but are crucial for explaining particular ecological concepts.

4.3. Ordered scenarios and simulation models

The dimensions listed above provide "natural" constraints to further organise 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* (Fig. 8). **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

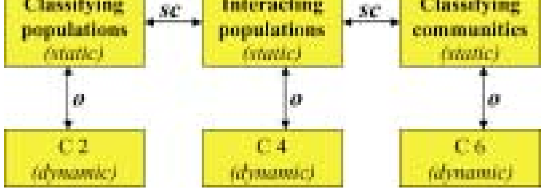

Fig. 8. Cluster organisation of population ecology simulation models.

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 the effects of natality before discussing the structure of the involved population. 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 described above, we have defined six clusters of simulation models (Fig. 8). Below each cluster is briefly described.

C1: Classifying Populations. A typical progression first addresses the zero-order cluster for single populations. Models in this C1 cluster encode knowledge about general features of single populations (no dynamic aspects) (Fig. 9). The main educational goals concern the kinds of populations that exist and what their characteristics are. Within this cluster, questions and assignments follow from the dimensions **G/S** and **A**.

C2: Single population dynamics. An **O** dimension step from C1 leads to the first-order cluster for those populations. Simulations in this C2 cluster show how processes enforce changes in a population. The main educational goals are to discuss the general laws of population growth; to identify the basic processes that cause changes to any population; and to discuss specialised versions of the basic processes (Fig. 10). Within this cluster questions explore the dimensions **G/S**, **A**, and **I**.

C3: Classifying two interacting populations. Next, a **SC** step from C1 leads to the zero-order cluster for two populations. Cluster C3 concerns the structure of pairs of interacting populations. Educational goals are to demonstrate how two populations may affect each other (or some natural resource) and how that happens via the basic population processes. Questions mainly explore the **A** and **I** dimensions, but possibly also the **G/A**.

C4: Dynamics of two populations. An **O** step from C3 leads to the first-order cluster for those populations. Cluster C4 represents the dynamics of the interactions between two populations. Ecological concepts represented in this cluster are the same as in cluster C3. However, the learner can now see the dynamics involved in these relations and notice that they account for community changes. Educational goals are to demonstrate how the values of *Nof* for the two populations change and to express the behaviour of populations involved in different types of interaction. Questions and assignments mainly explore the dimensions **G/S**, **A** and **I**.

C5: Classifying Communities. A **SC** step from C3 leads to the zero-order cluster for the Cerrado community (three populations). This (zero-order) C5 cluster elaborates on the concept of community by using representations of three populations (Fig. 11). The main educational goal is to illustrate different types of Cerrado communities in terms of the values of quantities representing the population sizes. Questions explore the dimensions **G/S** and **A**.

C6: Community dynamics. Finally, an **O** step from C5 leads to the first-order cluster for that community (i.e., the full CSH model). This (first-order) C6 cluster represents the behaviour of Cerrado communities. Environmental factors such as cover, litter, temperature,

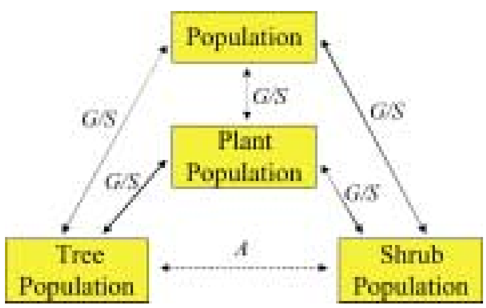

Fig. 9. Dimensions for learning routes in cluster one.

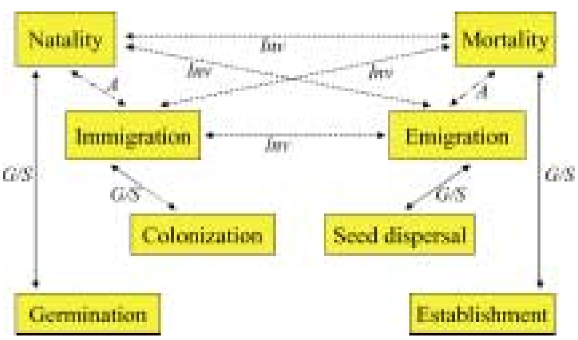

Fig. 10. Dimensions for learning routes in cluster two.

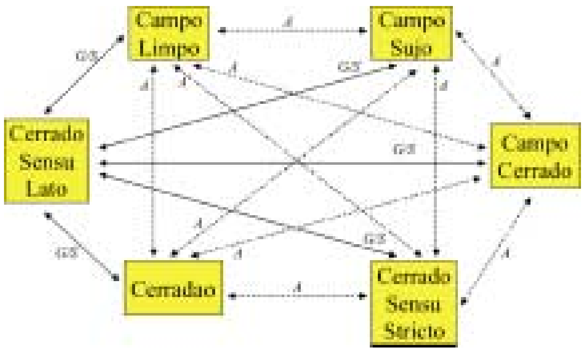

Fig. 11. Dimensions for learning routes in cluster five.

nutrients, water, fire frequency and their influence on different plant species in the Cerrado are included in the models of this cluster. 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. 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 (for details see [21,23,24]).

4.4. *Related research on supporting learners*

Having an organised set of qualitative simulations is an important step towards actual use in teaching practice. It allows learners to gradually progress through the material while acquiring more advanced insights. However, also crucial is the realisation of means that support learners in interacting with the simulations. Graphical user interfaces (e.g., [20]) and diagrammatic visualisations (e.g., [15]) are important in this respect. VISIGARP [1] is a tool that (a) provides a graphical user interface to control the simulation software, and (b) automatically generates diagrammatic representations of simulation results. A study with real learners using the simulations described in this article, albeit in an experimental setting, showed the usefulness of this approach [27]. The results obtained in this experiment support the hypothesis that qualitative simulations enable learners to effectively acquiring domain knowledge. However, to further enhance the *communicative interaction* [5] with interactive learning environments, such as VISIGARP, additional support is needed. Our work in this direction includes domain independent means for automatically generating questions and assignments [8], tracking learner difficulties and misconceptions [12], and generating explanations [22].

5. Discussion and concluding remarks

Community behaviour can be seen as the result of a complex web of relationships and interactions between pairs of populations. Understanding such interactions constitutes an important part of ecological theory and practice. We have presented a set of qualitative simulation models that capture knowledge about the interactions between two populations. With these models it is possible to derive complex community behaviour from what can be seen as the "first principles" in population ecology.

Qualitative models can be used to support simulations in interactive learning environments. They provide modelling primitives for representing aspects that normally are hard to capture in numerical models, such as a vocabulary to describe objects and situations, to represent the assumptions underlying a model, and to represent causal relationships. Models about predation, competition, and other population interactions presented in this article illustrate these points.

Qualitative models force a model builder to explicate the details relevant to a system's behaviour. Initially, when reciprocal influences of two interacting populations are represented (e.g., as proportionalities), the reasoning engine generates all possible behaviours, because the situation is (qualitatively speaking) ambiguous. This means that for each interaction we have to specify exactly how that behaviour is different from "just being ambiguous". The qualitative approach enforces the explication of the assumptions and constraints that must be true for interacting populations to show a certain type of behaviour. Articulating all that knowledge explicitly in simulation models is a major advantage of the work presented in the paper. The result provides an interesting workbench for learners to work with, while constructing their own understanding of interacting populations.

The newly created simulations, about pairs of interacting populations, are integrated with previous work implementing a qualitative theory of population dynamics and models that simulate the Cerrado Succession Hypothesis. The resulting complex library of simulations has been reorganised to facilitate progressive learning routes using ideas on model dimensions from Causal Model Progression (CMP), the Genetic Graph (GG), and the Didactic Goal Generator (DGG). Six clusters have been defined and specific scenarios have been constructed to run simulations within each cluster. The details in the clusters are organized using the dimensions generalisation/specialisation (**G/S**), analogy (**A**), and inverse (**I**). The order (**O**) dimension is used to move from static to dynamic models (and vice versa). The structural change (**SC**) dimension can be used to increase the complexity of the ecological system being modelled and for instance progress from populations, via communities, to ecosystems (and vice versa).

The models presented in this paper have been developed in close collaboration with domain experts. Further validation of the models and their usefulness in classroom settings is still in progress. However, preliminary studies with learners have shown encouraging results, with respect to the latter. An important factor in determining the validity of the models is their ability to scale up and act as building blocks for simulating complex community behaviours. Recent work [25] on successfully modelling the behaviour of a community consisting of *four* interacting species is promising in this respect.

References

[1] A. Bouwer and B. Bredeweg, VISIGARP: graphical representation of qualitative simulation models, in: *Artificial Intelligence in Education – AI-ED in the Wired and Wireless Future*, J.D. Moore, G. Luckhardt Redfield and J.L. Johnson, eds, IOS Press/Ohmsha, Osaka, Japan, 2001, pp. 294–305.

[2] B. Bredeweg, Expertise in qualitative prediction of behaviour, PhD thesis, University of Amsterdam, Amsterdam, The Netherlands, 1992.

[3] B. Bredeweg and R. Winkels, Qualitative models in interactive learning environments: an introduction (introduction to special issue), *Interactive Learning Environments* **5** (1998), 1–18.

[4] L.M. Coutinho, Fire in the ecology of Brazilian Cerrado, in: *Fire in the Tropical Biota*, J.G. Goldammer, ed., Ecological Studies, Vol. 84, Springer-Verlag, Berlin-Heidelberg, 1990, pp. 82–105.

[5] M. Elsom-Cook, *Principles of Interactive Multimedia*, McGraw-Hill, 2001.

[6] B. Falkenhainer and K. Forbus, Compositional modeling: finding the right model for the job, *Artificial Intelligence* **5**(1-3) (1991), 95–143.

[7] K. Forbus, Qualitative process theory, *Artificial Intelligence* **24**(1-3) (1984), 85–168.

[8] F. Goddijn, A. Bouwer and B. Bredeweg, Automatically generating tutoring questions for qualitative simulations, in: *Proceedings of the 17th International Workshop on Qualitative Reasoning*, Brazil, 2003 (forthcoming).

[9] I.P. Goldstein, The genetic graph: a representation for the evolution of procedural knowledge, *International Journal of Man-Machine Studies* **11** (1979), 51–77.

[10] J.W. Haefner, *Modeling Biological Systems: Principles and applications*, Chapman & Hall, New York, 1996.

[11] S.E. Jørgensen and G. Bendoricchio, *Fundamentals of Ecological Modelling*, 3rd edn, Elsevier Science, Oxford, 2001.

[12] K. de Koning, B. Bredeweg, J. Breuker and B. Wielinga, Model-based reasoning about learner behaviour, *Artificial Intelligence* **117** (2000), 173–229.

[13] C.J. Krebs, *Ecology: The Experimental Analysis of Distribution and Abundance*, HarperCollins College Publishers, New York, 1994.

[14] B. Kuipers, Qualitative simulation, *Artificial Intelligence* **29** (1986), 289–338.

[15] Z. Kulpa, Diagrammatic representation and reasoning, *Machine Graphics & Vision* **3**(1-2) (1994), 77–103.

[16] P.R. Mendonça and A.J. Piratelli, Frugivoria e dispersão de sementes por vertebrados no cerrado, in: *Contribuição ao conhecimento ecológico do cerrado*, L.L. Leite and C.H. Saito, eds, Brasília, Universidade de Brasília, Departamento de Ecologia, 1997, pp. 112–116.

[17] A. Moreira, Fire protection and vegetation dynamics in the Brazilian Cerrado, PhD thesis, Harvard University, Cambridge, Massachusetts, 1992.

[18] I.B.A. Morosini and C.A. Klink, Interferência do capim-grodura (Melinis minutiflora Beauv) no desenvolvimento de plântulas de embaúba (Cecropia pachystachya Trécul), in: *Contribuição ao Conhecimento Ecológico do Cerrado*, L.L. Leite and C.H. Saito, eds, Brasília, Universidade de Brasília, Departamento de Ecologia, 1997, pp. 82–86.

[19] E.P. Odum, *Ecologia* (translation of basic ecology, 1983), Discos CBS, Rio de Janeiro, 1985.

[20] J. Preece, Y. Rogers and H. Sharp, *Interaction Design: Beyond Human-computer Interaction*, John Wiley & Sons, 2002.

[21] P. Salles and B. Bredeweg, Building qualitative models in ecology, in: *Proceedings of the 11th International Workshop on Qualitative Reasoning*, L. Ironi, ed., Instituto di Analisi Numerica C.N.R., Pubblicazioni no. 1036, Pavia, Italy, 1997, pp. 155–164.

[22] P. Salles, B. Bredeweg and R. Winkels, Deriving explanations from qualitative models, in: *Artificial Intelligence in Education: Knowledge and Media in Learning Systems*, B. du Bouley and R. Mizoguchi, eds, IOS Press, Ohmsha, 1997, pp. 474–481.

[23] P. Salles and B. Bredeweg, Constructing progressive learning routes through qualitative simulation models in ecology, in: *Proceedings of the 15th International Workshop on Qualitative Reasoning*, G. Biswas, ed., Saint Mary's University, San Antonio, TX, USA, 2001, pp. 82–89.

[24] P. Salles and B. Bredeweg, Modelling population and community dynamics with qualitative reasoning, *Ecological Modelling* (forthcoming).

[25] P. Salles, B. Bredeweg and N. Bensusan, The Ants' Garden: Qualitative models of complex interactions between populations, in: *Proceedings of the 17th International Workshop on Qualitative Reasoning*, Brazil, 2003 (forthcoming).

[26] S. Scherrer, I.R. Diniz and H.C. Morais, Caracterização da fauna de parasitóides (Hymenoptera) de lagartas, no cerrado de Brasília, in: *Contribuição ao Conhecimento Ecológico do Cerrado*, L.L. Leite and C.H. Saito, eds, Brasília, Universidade de Brasília, Departamento de Ecologia, 1997, pp. 131–134.

[27] P. Tjaris, Kwalitatieve simulaties als middel tot kennisoverdracht, Master thesis, University of Amsterdam, Amsterdam, 2002 (in Dutch).

[28] B.Y. White and J.R. Frederiksen, Causal model progressions as a foundation for intelligent learning environments, *Artificial Intelligence* **42** (1990), 99–157.

[29] R. Winkels and J. Breuker, Automatic generation of optimal learning routes, in: *Proceedings of International Conference on Artificial Intelligence and Education*, 1993, pp. 330–337.