Semi-Quantitative Comparative Analysis And Its Application*

Ivayla Vatcheva[1] and **Hidde de Jong**[2]
[1] Department of Computer Science, University of Twente
P.O. Box 217, 7500 AE Enschede, the Netherlands
[2] Institut National de Recherche en Informatique et en Automatique (INRIA) Rhône-Alpes
655, avenue de l'Europe, 38330 Montbonnot Saint Martin, France
Email: ivayla@cs.utwente.nl, Hidde.de-Jong@inrialpes.fr

Abstract

SQCA is an implemented technique for the semi-quantitative comparative analysis of dynamical systems. It is both able to deal with incompletely specified models and make precise predictions by exploiting semi-quantitative information in the form of numerical bounds on the variables and functions occuring in the models. The technique has a solid mathematical foundation which facilitates proofs of correctness and convergence properties. SQCA represents the core of a method for the automated prediction of experimental results.

Introduction

In many situations it is important to compare the behavior of dynamical systems. A population biologist, for instance, may want to predict the consequences of the introduction of a new species into an ecosystem. For an engineer monitoring a chemical process, it may be critical to know whether a particular perturbation could explain observed deviations from the normal behavior.

If quantitative models and precise quantitative information about the initial conditions are available, a *comparative analysis (CA)* of the behaviors of the systems is straightforward. One simply compares the behaviors predicted by means of numerical simulation at the time-points of interest. Often, however, the available information about the systems is incomplete. In such cases we can resort to qualitative models to describe the systems, predict behaviors from an initial qualitative state by means of qualitative simulation (Kuipers 1994), and use qualitative CA techniques to compare the behaviors (Weld 1988; Neitzke & Neumann 1994; de Jong & van Raalte 1997).

A disadvantage of qualitative CA techniques is the imprecision of their conclusions, which hampers their upscalability. When comparing the behaviors of more complex systems, with several structural differences and differences in initial conditions, de Jong and van Raalte's CEC* is likely to generate a large number of possible comparative behaviors. Besides these ambiguities, due to the qualitative nature of the available information, it only characterizes differences as higher or lower, without giving an indication of their magnitude.

In this paper we introduce *SQCA*, a technique which arrives at more precise conclusions than qualitative CA techniques, while retaining their ability to deal with incomplete information. The technique exploits semi-quantitative information about the systems, in the form of numerical bounds on the variables and functions occurring in the models. Although SQCA will be presented as a self-contained technique, it can also be integrated as a filter on comparative behaviors into a qualitative CA algorithm. The implementation of SQCA has been used to answer CA questions involving structural differences in combination with differences in the initial conditions of the systems.

SQCA forms the core of a method for the automated predictions of experimental results that is currently being developed. The approach predicts an estimated value of an unperformed measurement by exploiting available knowledge about experiments already carried out. To this end it uses an integration of techniques from the field of automated modeling and reasoning about physical systems.

The presentation starts with a brief review of semi-quantitative simulation, since semi-quantitative models and behaviors form the input of SQCA. Semi-quantitative CA is basically a constraint propagation process. The next section describes how the requisite constraints are derivable from the models and behaviors of the systems to be compared. The SQCA algorithm is then presented, together with guarantees on its correctness and convergence. In the following sections the results obtained by means of SQCA are given, followed by a short description of the approach to the automated prediction of experimental results. A brief discussion and ideas for further work conclude this article.

Semi-quantitative simulation

We employ the semi-quantitative simulation techniques Q2+Q3, which function as filters on qualitative be-

*A shorter version of this paper appears in the proceedings of the 16th International Joint Conference on Artificial Intelligence, IJCAI-99, Stockholm, 31 July - 6 August 1999.

$$QV(\dot{h}) = QV(v)$$
$$QV(\dot{v}) = QV(a)$$
$$QV(a) = -QV(g)$$
$$QV(\dot{g}) = \langle 0, std\rangle \quad \text{(a)}$$

$$QV(\dot{h}) = QV(v)$$
$$QV(\dot{v}) = QV(a)$$
$$QV(a) = -QV(g)\frac{QV(r)^2}{QV(x)^2} - QV(k)QV(v)QV(|v|)$$
$$QV(x) = QV(r) + QV(h),\ QV(\dot{r}) = \langle 0, std\rangle$$
$$QV(\dot{g}) = \langle 0, std\rangle,\ QV(\dot{k}) = \langle 0, std\rangle \quad \text{(b)}$$

$$QV(\dot{h}) = QV(v)$$
$$QV(\dot{v}) = QV(a)$$
$$QV(a) = -QV(g)f(QV(x)) - j(QV(v)),$$
$$\text{with } f = M^-,\ j = M_0^+$$
$$QV(\dot{g}) = \langle 0, std\rangle \quad \text{(c)}$$

$$\text{range}(g) = [9.83, 9.83],\ \text{range}(k) = [0.0005, 0.001],\ \text{range}(r) = [6.37, 6.37] \times 10^6,\ \text{range}(A) = [1, 1],\ \text{range}(\rho) = [1.29, 1.29],$$
$$\text{range}(m) = [11.72, 11.74] \times 10^3,\ \text{range}(c) = [9.1, 10.9],\ \text{envelope}(f) = [\frac{\underline{r}^2}{x^2}, \frac{\overline{r}^2}{x^2}],\ \text{envelope}(j) = [\underline{c}\frac{\underline{A}}{\overline{m}}\frac{\underline{\rho} v\,|v|}{2}, \overline{c}\frac{\overline{A}}{\underline{m}}\frac{\overline{\rho} v\,|v|}{2}] \quad \text{(d)}$$

Figure 1: QDEs for an object fired upwards in a gravitational field, where the gravitational field is (a) constant, (b) height-varying and completely specified, and (c) height-varying and incompletely specified. In (b) and (c) friction is taken into account, whereas in (a) it is neglected. (d) Ranges and envelopes which turn the QDEs into SQDEs. The variable h stands for height above the Earth surface, v for velocity, a for acceleration, g for gravitational constant, r for Earth radius, x for distance from the center of the Earth, and k for a constant dependent on the air density ρ, projected area A of the object in the direction of motion, object mass m, and drag coefficient c.

haviors obtained by means of QSIM (Kuipers 1994; Berleant & Kuipers 1997). Although other simulation techniques could have been used as well (e.g., Vescovi, Farquhar, & Iwasaki (1995); Kay & Kuipers (1993)), we have chosen Q2+Q3 because they produce a semi-quantitative annotation of the behaviors while preserving their underlying qualitative structure.

The models used for semi-quantitative simulation are *semi-quantitative differential equations (SQDEs)*, that is, qualitative differential equations (QDEs) enhanced with numerical information (figure 1). We use a notation for QDEs which emphasizes their abstraction from ODEs and which simplifies the propositions in later sections. Besides the *basic* qualitative constraints in QSIM, it allows the use of *composite* qualitative constraints (Vatcheva & de Jong 1999). For instance, the constraint

$$QV(a) = -QV(g)\frac{QV(r)^2}{QV(x)^2} - QV(k)QV(v)QV(|v|)$$

in figure 1(b) is composed of $QV(a) = QV(p_1) + QV(p_2)$,

$$QV(p_1) = -QV(g)QV(r)^2/QV(x)^2, \text{ and}$$
$$QV(p_2) = -QV(k)QV(v)QV(|v|), \qquad (1)$$

the latter two being composite constraints themselves. Generally speaking, we deal with constraints $QV(y) = f(QV(x_1), \ldots, QV(x_n))$ (or $QV(y) = f(QV(\boldsymbol{x}))$), where f is a qualitative constraint between the variables $y, x_1, \ldots, x_n$.

The semi-quantitative information completing the QDE takes several forms. In the first place, numerical ranges are added to landmarks. For a landmark l_i, range(l_i) is defined as an interval $[\underline{l}_i, \overline{l}_i]$ with $\underline{l}_i, \overline{l}_i \in \mathbb{R}^*$. Second, envelopes can be added to monotonic function constraints in the QDE. An envelope(f) of a monotonic function is defined as a pair of functions $\underline{f}, \overline{f}$, with $\underline{f}(x) \leq f(x) \leq \overline{f}(x)$, for all x in the domain of f.

Q2 is a technique which uses the ranges and envelopes of the SQDE to refine a qualitative behavior tree produced by QSIM. Given ranges for the variables in the initial qualitative state, it builds a constraint network and propagates the initial ranges through this network. The constraint network relates the variables at each distinguished time-point through constraints on their ranges. Constraint propagation is achieved by recursively evaluating the constraint expressions by means of interval arithmetic and by updating the range of a landmark through intersection of the present range and the newly calculated range. Q2 either rules out qualitative behaviors or produces qualitative behaviors in which the qualitative values are annotated with numerical ranges, so-called *semi-quantitative behaviors (SQBs)*.

Q3 improves upon the results obtained by means of Q2 by following an approach called step-size refinement. First, it locates or creates a gap in a semi-quantitative behavior, that is, it takes a pair of adjacent distinguished time-points t_i and t_{i+1}, such that $\overline{t}_i < \underline{t}_{i+1}$. Then, it interpolates a new state in this gap at an auxiliary time-point t_{aux}, $t_i < t_{aux} < t_{i+1}$, and provides initial ranges for the qualitative value of the variables at t_{aux}. The newly created state adds new landmarks and constraints to the constraint network. A new round of constraint propagation by means of Q2 results in a refined or refuted semi-quantitative behavior.

Figure 2 shows two semi-quantitative behaviors produced by QSIM and Q2+Q3 from the models in figure 1. Both behaviors describe an object falling back to its initial height, in the first case in a constant gravitational field without friction and in the second case in a height-varying gravitational field with friction.

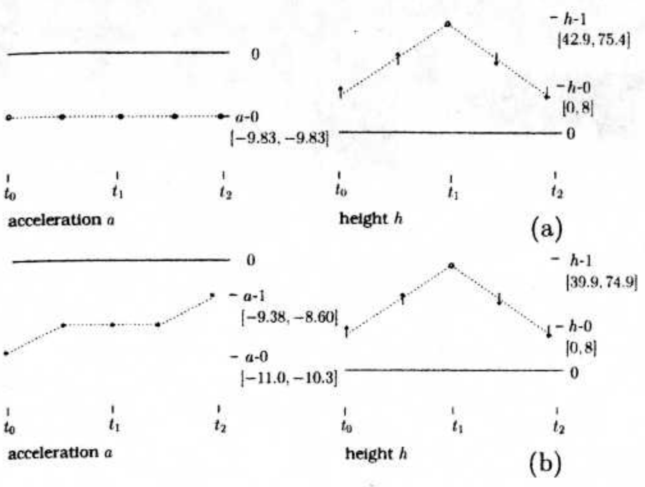

Figure 2: SQBs obtained from the models in figure 1(a) and (b), respectively.

RIVs and RIV constraints

Relative interval values

Consider two SQBs, either topologically equal or topologically different (Weld 1988). Topologically equal behaviors show the same sequence of transitions between qualitative states and the (shared) variables have the same qualitative value in the corresponding states of this sequence.

Let $t_0, \dots, t_n$ and $\hat{t}_0, \dots, \hat{t}_m$ be the sequences of distinguished time-points in the first and the second SQB, respectively. The behaviors are compared at meaningful pairs of comparison (de Jong & van Raalte 1997). A pair of comparison pc is a pair $\langle t, \hat{t} \rangle$ of time-points in the behavior of the first and second system.[1] It is meaningful if one of the following conditions is satisfied: (1) t and $\hat{t}$ are the initial time-points t_0 and $\hat{t}_0$; (2) t and $\hat{t}$ are the end time-points t_n and $\hat{t}_m$; or (3) t and $\hat{t}$ are time-points at which a variable reaches the same landmark 0, $-\infty$ or ∞ in both systems. Pairs of comparison can be ordered by a partial ordering relation: $pc_0 \preceq pc_1$, iff $t_0 \leq t_1$ and $\hat{t}_0 \leq \hat{t}_1$, where $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$. pc_1 is called a successor of pc_0. A comparison of the behaviors in figure 2 yields the meaningful pairs of comparison $pc_0 = \langle t_0, \hat{t}_0 \rangle$, $pc_1 = \langle t_1, \hat{t}_1 \rangle$, and $pc_2 = \langle t_2, \hat{t}_2 \rangle$, with $pc_0 \preceq pc_1 \preceq pc_2$.

A comparison of the shared variables of the systems at a pair of comparison gives rise to *relative interval values (RIVs)*. They provide an estimate of the difference $\Delta x(pc) = \hat{x}(\hat{t}) - x(t)$ of variables x at $pc = \langle t, \hat{t} \rangle$.

[1] As a notational convention, $\hat{}$ denotes variables in the behavior of the second system.

Definition 1 The RIV of a shared variable x at a pair of comparison pc is defined as range($\Delta x(pc)$).

The RIVs at a pair of comparison are related to each other, and to the RIVs at predecessor and successor pairs of comparison. The *RIV constraints* expressing these relations are derived from the SQBs and the SQDEs of the systems which we want to compare. Several types of RIV constraints exist.

Constraints from SQBs

A direct way to obtain a range for the difference of a variable at a pair of comparison is to examine the numerical information in the states of the SQBs.

Proposition 1 At $pc = \langle t, \hat{t} \rangle$ the relative interval value of x is given by

$$\text{range}(\Delta x(pc)) \subseteq \text{range}(\hat{x}(\hat{t})) - \text{range}(x(t)).$$

A special case of this proposition is the difference in the duration of the behavior fragments T and $\hat{T}$, defined by two successive pairs of comparison $pc_0 = \langle t_0, \hat{t}_0 \rangle$ and $pc_1 = \langle t_1, \hat{t}_1 \rangle$:

$$\text{range}(\Delta T(pc_0, pc_1)) \subseteq (\text{range}(\hat{t}_1) - \text{range}(\hat{t}_0)) - (\text{range}(t_1) - \text{range}(t_0)).$$

The behaviors in figure 2 show that the range of the acceleration in the first and second system is range($a(t_0)$) $= [-9.83, -9.83]$ and range($\hat{a}(\hat{t}_0)$) $= [-11.0, -10.3]$, respectively. Applying proposition 1 at pc_0 yields range($\Delta a(pc_0)$) $= [-1.17, -0.47]$.

Constraints from SQDEs at a pair of comparison

Suppose the qualitative value of a shared variable x is constrained in the first and second system as follows:

$$QV(x) = f(QV(\boldsymbol{r})); \quad QV(\hat{x}) = g(QV(\hat{\boldsymbol{s}})), \tag{2}$$

where f and g represent qualitative constraints, and $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$ are vectors of variables. We will allow the models of the two systems to be structurally different, so f and g as well as $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$ may be different.

In order to derive an RIV constraint from (2), f and g need to be made comparable first. This is attained by bringing f and g in the form of a single constraint, the so-called *comparison constraint*. Let $\boldsymbol{q}$ be the vector of variables occurring both in $\boldsymbol{r}$ and $\boldsymbol{s}$, and $\boldsymbol{a}$ a vector of newly introduced auxiliary variables with specified qualitative values, so-called *comparison values*. The constraints f and g are comparable through a comparison constraint h,

$$\begin{aligned} QV(x) &= h(QV(\boldsymbol{q}), QV(\boldsymbol{a})); \\ QV(\hat{x}) &= h(QV(\hat{\boldsymbol{q}}), QV(\hat{\boldsymbol{a}})), \end{aligned} \tag{3}$$

under the following condition: h is satisfied iff f and g are satisfied for every $QV(\boldsymbol{r})$, $QV(\hat{\boldsymbol{s}})$ given the comparison values $QV(\boldsymbol{a})$, $QV(\hat{\boldsymbol{a}})$.

In contrast with (de Jong & van Raalte 1997), the existence of such a comparison constraint can be guaranteed. The set of basic qualitative constraints is restricted and for every pair of f and g a comparison constraint can be easily found due to the simple form of f and g. When f and g are composite, a comparison constraint is obtained by decomposing f and g into basic constraints and composing the corresponding comparison constraints into a composite comparison constraint h (Vatcheva & de Jong 1999).

The acceleration constraint from the model in figure 1(b) can be decomposed as in (1) and the acceleration constraint in figure 1(a) as

$$QV(p_1) = -QV(g) \text{ and } QV(a) = QV(p_1). \tag{4}$$

The comparison constraint of the first qualitative constraints in (1) and (4) is defined as $QV(p_1) = -QV(g)QV(a_1)$ with comparison values $QV(a_1) = \langle 1, std\rangle$ and $QV(\hat{a}_1) = QV(\hat{r}^2/\hat{x}^2)$. The comparison constraint of the second constraints is $QV(a) = QV(p_1) + QV(a_2)$ with comparison values $QV(a_2) = \langle 0, std\rangle$ and $QV(\hat{a}_2) = QV(\hat{p}_2)$. Their composition defines h as

$$QV(a) = -QV(g)QV(a_1) + QV(a_2).$$

Proposition 2 Suppose that $QV(x)$ and $QV(\hat{x})$ are constrained by f and g, as in (2). Let h be the comparison constraint of f and g. The RIV of x at $pc = \langle t, \hat{t}\rangle$ is given by:

$$\begin{aligned}\text{range}(\Delta x(pc)) \subseteq \text{range}(\boldsymbol{d_q})^t \cdot \text{range}(\Delta \boldsymbol{q}(pc)) + \\ \text{range}(\boldsymbol{d_a})^t \cdot \text{range}(\Delta \boldsymbol{a}(pc)),\end{aligned} \tag{5}$$

with $\boldsymbol{d_q}$ and $\boldsymbol{d_a}$ vectors of partial derivatives of the function corresponding to the comparison constraint, i.e., $d_{q,i} = \frac{\partial}{\partial q_i} h(\boldsymbol{mq}, \boldsymbol{ma})$ and $d_{a,j} = \frac{\partial}{\partial a_j} h(\boldsymbol{mq}, \boldsymbol{ma})$.[2] Further, mq_i lies between $q_i(t)$ and $\hat{q}_i(\hat{t})$, and ma_j between $a_j(t)$ and $\hat{a}_j(\hat{t})$.

Proof. The constraints in (3) are abstractions of the mathematical equations

$$x(t) = h(\boldsymbol{q}(t), \boldsymbol{a}(t)); \; \hat{x}(\hat{t}) = h(\hat{\boldsymbol{q}}(\hat{t}), \hat{\boldsymbol{a}}(\hat{t})),$$

where h is a continuously differentiable function. Subtracting $x(t)$ and $\hat{x}(\hat{t})$ and applying the generalized mean value theorem, one finds

$$\begin{aligned}\Delta x(pc) = &\sum_{i=1}^{n} \frac{\partial}{\partial q_i} h(\boldsymbol{mq}, \boldsymbol{ma})(\hat{q}_i(\hat{t}) - q_i(t)) \\ &+ \sum_{j=1}^{m} \frac{\partial}{\partial a_j} h(\boldsymbol{mq}, \boldsymbol{ma})(\hat{a}_j(\hat{t}) - a_j(t)),\end{aligned}$$

[2]Throughout this paper h is used to refer both to constraint and the mathematical function from which the constraint is abstracted. Whenever a confusion is possible, we explicitly speak of the constraint h or the function h.

where mq_i lies between $q_i(t)$ and $\hat{q}_i(\hat{t})$, and ma_j between $a_j(t)$ and $\hat{a}_j(\hat{t})$. □

Ranges for the partial derivatives of h are derived from interval extensions $D_{q,i}$ and $D_{a,j}$ of $d_{q,i}$ and $d_{a,j}$, respectively (Moore 1979). It can be easily shown that such interval extensions always exist and are uniquely specified (Vatcheva & de Jong 1999).

In this way,

$$\text{range}(d_{q,i}) = D_{q,i}(\text{range}(\boldsymbol{mq}), \text{range}(\boldsymbol{ma})) \text{ with}$$

$$\begin{aligned}\text{range}(mq_i) &= \text{span}(q_i(t), \hat{q}_i(\hat{t})) \\ &= [\min(\underline{q}_i(t), \underline{\hat{q}}_i(\hat{t})), \max(\overline{q}_i(t), \overline{\hat{q}}_i(\hat{t}))].\end{aligned}$$

Similar expressions are obtained for range(ma_j) and range$(d_{a,j})$.

In the example above we find at pc_0 the RIV constraint

$$\begin{aligned}\text{range}(\Delta a(pc_0)) \subseteq &-\text{span}(a_1(t_0), \hat{a}_1(\hat{t}_0)) \cdot \text{range}(\Delta g(pc_0)) \\ &- \text{span}(g, \hat{g}) \cdot \text{range}(\Delta a_1(pc_0)) + \text{range}(\Delta a_2(pc_0)),\end{aligned}$$

with $\text{span}(a_1(t_0), \hat{a}_1(\hat{t}_0)) = \text{span}(1, \hat{r}^2/\hat{x}^2(\hat{t}_0))$, $\text{range}(\Delta a_1(pc_0)) = \text{range}(\hat{r}^2/\hat{x}^2(\hat{t}_0)) - [1,1]$, and $\text{range}(\Delta a_2(pc_0)) = \text{range}(\hat{p}_2(\hat{t}_0))$.

Constraints from SQDEs between pairs of comparison

Between pairs of comparison the behavior of a shared state variable x is determined by the derivative constraints in the SQDEs:

$$QV(\dot{x}) = QV(r); \quad QV(\dot{\hat{x}}) = QV(\hat{s}). \tag{6}$$

Derivative constraints give rise to additional RIV constraints. Consider the pairs of comparison $pc_0 = \langle t_0, \hat{t}_0\rangle$ and $pc_1 = \langle t_1, \hat{t}_1\rangle$, which define primitive behavior fragments $[t_0, t_1]$ and $[\hat{t}_0, \hat{t}_1]$, that is, behavior fragments without intermediary distinguished time-points. The intervals will usually contain auxiliary time-points $t_{aux_i} \in]t_0, t_1[$ and $\hat{t}_{aux_j} \in]\hat{t}_0, \hat{t}_1[$, which have been interpolated during simulation.

Since in general $t_0 \neq \hat{t}_0$, we will *synchronize* the behavior fragments first by means of a procedure which shifts the uncertainty in t_0 and $\hat{t}_0$ to subsequent time-points. The ranges of the synchronized time-points t^s in the behavior fragment of the first system are defined as: range$(t_0^s) = [0, 0]$, range$(t_1^s) = [\underline{t}_1 - \overline{t}_0, \overline{t}_1 - \underline{t}_0]$, and range$(t^s_{aux_i}) = [\underline{t}_{aux_i} - \overline{t}_0, \overline{t}_{aux_i} - \underline{t}_0]$. Synchronization of the behavior fragment of the second system is accomplished in the same way. We will henceforth assume that the behavior fragments have been synchronized already.

We now introduce *auxiliary pairs of comparison* by means of the auxiliary time-points. These pairs of comparison allow one to improve the prediction of differences at qualitatively important time-points.

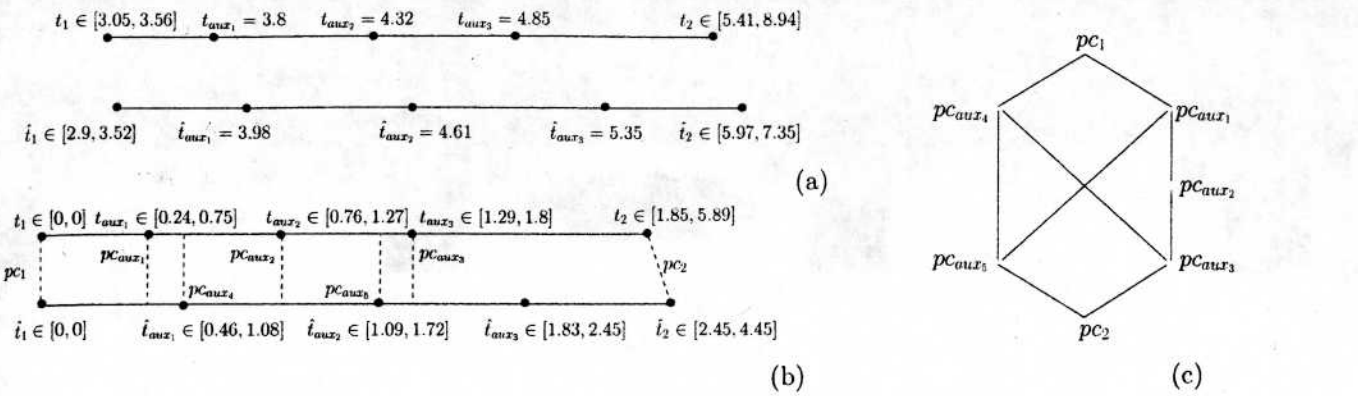

Figure 3: (a) Behavior fragments $[t_1, t_2]$ and $[\hat{t}_1, \hat{t}_2]$ of the behaviors in figure 1. (b) The synchronized behavior fragments with the auxiliary pairs of comparison and (c) the ordered (top to bottom) auxiliary pairs of comparison.

Definition 2 Suppose two systems are compared over primitive behavior fragments defined by pc_0 and pc_1 with n and m auxiliary time-points. Setting $t_{aux_{n+1}} = t_1$ and $\hat{t}_{aux_{m+1}} = \hat{t}_1$, we define auxiliary pairs of comparison

$$pc_{aux_k} = \langle t_{aux_i}, t_{aux_i} \rangle, \ \text{range}(t_{aux_i}) \leq \text{range}(\hat{t}_1), \text{ or}$$
$$pc_{aux_k} = \langle \hat{t}_{aux_j}, \hat{t}_{aux_j} \rangle, \ \text{range}(\hat{t}_{aux_j}) \leq \text{range}(t_1)$$

where $1 \leq i \leq n+1$, $1 \leq j \leq m+1$, and $1 \leq k \leq n+m+1$.

Notice that we introduce auxiliary pairs of comparison only conditionally. The condition $t_{aux_i} \leq \hat{t}_1$ for pc_{aux_k} ensures that t_{aux_i} is a time-point really occurring in the (synchronized) behavior fragment of the second system.

Figure 3(a) shows primitive behavior fragments of an object launched upwards. Q3 has interpolated three auxiliary time-points in each behavior fragment. The synchronized behavior fragments and the auxiliary pairs of comparison are shown in figure 3(b). The pairs of comparison have been ordered with respect to the $\preceq$-relation.

With the help of the auxiliary pairs of comparison, the RIV of the shared state variable x at pc_1 can be expressed in terms of the RIVs of x at auxiliary pairs of comparison between pc_0 and pc_1.

Proposition 3 Given the qualitative constraints (6) and k auxiliary pairs of comparison defined by definition 2. The relative interval value of a shared state variable x at pc_{aux_i} $(1 \leq i \leq k)$ and pc_1 is defined as follows

$$\text{range}(\Delta x(pc_{aux_i})) \subseteq \bigcap_j \{\text{range}(\Delta x(pc_{aux_j})) + (\text{span}(\hat{s}(t_{aux_j}), \hat{s}(t_{aux_i})) - \text{span}(r(t_{aux_j}), r(t_{aux_i}))) \cdot (\text{range}(t_{aux_i}) - \text{range}(t_{aux_j}))\}$$

$$\text{range}(\Delta x(pc_1)) \subseteq \bigcap_j \{\text{range}(\Delta x(pc_{aux_j})) + \text{span}(\hat{s}(t_{aux_j}), \hat{s}(\hat{t}_1)) \cdot (\text{range}(\hat{t}_1) - \text{range}(t_{aux_j})) - \text{span}(r(t_{aux_j}), r(t_1)) \cdot (\text{range}(t_1) - \text{range}(t_{aux_j}))\}$$

where pc_{aux_j} are direct predecessor pairs of comparison of pc_{aux_i} and pc_1, and pc_{aux_0} is set to pc_0.

Proof. Let $pc_{aux_i} = \langle t_{aux_i}, t_{aux_i} \rangle$, $pc_{aux_j} = \langle t_{aux_j}, t_{aux_j} \rangle$, $pc_{aux_j} \preceq pc_{aux_i}$. Applying the mean value theorem for x and $\hat{x}$ in the time interval $[t_{aux_j}, t_{aux_i}]$ and subtracting the resulting expressions we get

$$\Delta x(pc_{aux_i}) = \Delta x(pc_{aux_j}) + (\hat{s}(\hat{mt}_{aux}) - r(mt_{aux}))(t_{aux_i} - t_{aux_j}),$$

where $mt_{aux}, \hat{mt}_{aux} \in]t_{aux_i}, t_{aux_j}[$. Taking into account that r and $\hat{s}$ are qualitatively uniform between the adjacent time-points t_0, t_1 and $\hat{t}_0$, $\hat{t}_1$, respectively, and hence between the auxiliary time-points t_{aux_j}, t_{aux_i}, we conclude that $r(mt_{aux}) \in \text{span}(r(t_{aux_j}), r(t_{aux_i}))$ and $\hat{s}(\hat{mt}_{aux}) \in \text{span}(\hat{s}(t_{aux_j}), \hat{s}(t_{aux_i}))$. The expression above can then be transformed into

$$\text{range}(\Delta x(pc_{aux_i})) \subseteq \text{range}(\Delta x(pc_{aux_j})) + (\text{span}(\hat{s}(t_{aux_j}), \hat{s}(t_{aux_i})) - \text{span}(r(t_{aux_j}), r(t_{aux_i}))) \cdot (\text{range}(t_{aux_i}) - \text{range}(t_{aux_j})).$$

Since pc_{aux_i} may have more than one direct predecessor pc_{aux_j}, there can be a number of estimations of the RIV of x at pc_{aux_i} computed by means of the RIV of each pc_{aux_j}. Hence, the relative interval value of x at pc_{aux_i} is given by the intersection of these estimations.

The proof of the second part of the statement is accomplished in an analogous way. □

In the example of figure 3 the proposition contributes 6 RIV constraints for each of the variables h and v.

As a special case, consider the situation that x is constant in both systems, i.e.

$$QV(\dot{x}) = \langle 0, std \rangle; \quad QV(\dot{\hat{x}}) = \langle 0, std \rangle. \tag{7}$$

Without proof we add the following proposition.

Proposition 4 Suppose that $QV(x)$ and $QV(\hat{x})$ are constants, as in (7), and we compare the systems over behavior fragments determined by pc_0 and pc_1. The RIV of x at pc_1 is now simply range$(\Delta x(pc_1))$ = range$(\Delta x(pc_0))$.

For example, for the gravitational constant g we have range$(\Delta g(pc_0))$ = range$(\Delta g(pc_1)) = [0,0]$.

Redundancy of constraints

Proposition 2 relies on the mean value theorem to obtain more precise estimates of the RIVs of variables at pairs of comparison. One can prove that there are situations in which the RIV constraints thus defined do not improve upon the RIV constraints defined by proposition 1. In particular, this occurs when the semi-quantitative differential equations are *completely specified*. An SQDE is completely specified when it does not contain monotonic function constraints.

Theorem 1 Suppose the SQBs of two completely specified systems are compared. If range$(\Delta x(pc))_1$ is the RIV of a variable x at pc determined by proposition 1, and range$(\Delta x(pc))_{1+2}$ the same RIV determined by proposition 1 and 2, then range$(\Delta x(pc))_1 \subseteq$ range$(\Delta x(pc))_{1+2}$.

Since the models are completely specified, h does not contain monotonic function constraints. In this case the corresponding function and its partial derivatives are real-valued rational functions with corresponding natural interval extensions. The statement is then proved by analogy of the proof of proposition 2 using basic propositions from interval arithmetic.

SQCA algorithm

The algorithm for semi-quantitative comparative analysis takes as input two behaviors SQB,$\hat{SQB}$ and the corresponding models $SQDE$,$\hat{SQDE}$ of the systems, where $SQDE$ and $\hat{SQDE}$ are assumed to consist of basic qualitative constraints only. SQCA generates RIVs for all shared variables at the pairs of comparison from a set of initial relative interval values. The algorithm consists of the following three steps:

1. Establish the meaningful pairs of comparison implied by SQB and $\hat{SQB}$.
2. Generate the RIV constraints from $SQB, \hat{SQB}$ and $SQDE, \hat{SQDE}$, and build a constraint network.
3. Resolve the constraint network for the initial RIVs.

Propositions 1 to 4 define constraint schemata which are instantiated in the second step to yield appropriate RIV constraints from $SQB, \hat{SQB}$ and $SQDE, \hat{SQDE}$. In order to obtain tighter bounds for the RIVs, proposition 2 is not only applied to qualitative constraints of type (2), but also to algebraically equivalent constraints. If $\boldsymbol{q}$ is the n-vector of the variables occuring both in $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$, n additional variants of the qualitative constraints (2) can be formulated

$$QV(q_i) = f_i(\ldots, QV(r_{k-1}), QV(x), QV(r_{k+1}), \ldots),$$
$$QV(\hat{q}_i) = g_i(\ldots, QV(\hat{s}_{l-1}), QV(\hat{x}), QV(\hat{s}_{l+1}), \ldots),$$

where $r_k = q_i$ and $\hat{s}_l = \hat{q}_i$. These variants yield n additional RIV constraints by means of proposition 2. The constraints thus generated form a constraint network linking together the differences Δx of shared variables at the pairs of comparison.

In the third step the constraint network is resolved for the initial RIVs by means of the propagation algorithm included in Q2. The result of the constraint propagation is an RIV for each shared variable x at each pair of comparison pc. If some RIV is $\emptyset$, the initial RIVs are not consistent with the models $SQDE, \hat{SQDE}$ and behaviors $SQB, \hat{SQB}$ from which the RIV constraints have been derived.

SQCA has been shown to be *sound* and *incomplete* (Vatcheva & de Jong 1999). Call range$(\Delta x(pc))_{out}$ the range for a shared variable x at a pair of comparison pc that has been produced by SQCA. We now find:

Theorem 2 SQCA is sound, in that for any pair of solutions of ODEs consistent with the SQCA input it holds that $\Delta x(pc) \in$ range$(\Delta x(pc))_{out}$ for all x and pc.

Theorem 3 SQCA is incomplete, in that for some value riv in range$(\Delta x(pc))_{out}$ there may be no solutions of ODEs consistent with the SQCA input, such that $\Delta x(pc) = riv$.

Soundness is a consequence of the sound derivation of RIV constraints from SQDEs and SQBs and the soundness of the constraint propagation algorithm. Incompleteness is caused by the possibility of excess width in interval arithmetic (Moore 1979) and the use of the weak mean value theorem in propositions 2 and 3.

An important property of SQCA is its *convergence*.

Theorem 4 The relative interval values calculated by SQCA converge to a point value as the ranges in the initial qualitative states converge to a point value and the maximum step-size in the semi-quantitative behaviors converges to 0.

The theorem rests on the convergence of Q3 and holds under the same conditions (Berleant & Kuipers 1997).

Results

The SQCA algorithm has been implemented in Common Lisp. The program interacts with available implementations of QSIM and Q2, and our own implementation of Q3 (Vatcheva 1998): it takes semi-quantitative behaviors produced by QSIM and Q2+Q3 as input and calls Q2 functions for building and resolving constraint networks. In contrast with the implementation of CEC* the process of deriving propagation constraints from the SQBs and SQDEs has been completely automated.

This is possible due to the fact that the models in the SQCA input consist of basic qualitative constraints only.

In the first half of the table below the results of applying SQCA to the behaviors in figure 2 are shown. The trajectory of an object fired upward in a constant gravitational field without friction is compared with that in a height-varying gravitational field with friction (figure 1(a-b)). Although the initial height and velocity are incompletely known in both systems (with ranges [0, 8] and [30, 35], respectively), they are known to be equal, so that the initial relative interval values range($\Delta h(pc_0)$) and range($\Delta v(pc_0)$) are both [0, 0]. The SQCA results show that one cannot predict with certainty whether the maximum height reached by the second object will be higher or lower, i.e. range($\Delta h(pc_1)$) = $[-27.4, 27.8]$. The structural differences work in different directions, the height-varying gravitational field tending to increase and friction tending to decrease range($\Delta h(pc_1)$), while the uncertainty in the initial conditions is too large to distinguish between the two. The prediction of the difference in maximum height is more precise than that obtained in qualitative CA, however.

RIV	h	v	a
pc_0	$[0, 0]$	$[0, 0]$	$[-1.17, -0.47]$
pc_1	$[-27.4, 27.8]$	$[0, 0]$	$[1.4, 2.4] \times 10^{-4}$
pc_2	$[0, 0]$	$[-16.2, 11.7]$	$[0.45, 1.23]$
pc_0	$[-4, -2]$	$[16, 20]$	$[-1.76, -0.7]$
pc_1	$[44.6, 136.8]$	$[0, 0]$	$[1.25, 4.39] \times 10^{-4}$
pc_2	$[-4, -2]$	$[-35.6, 20.7]$	$[-1.55, 3.56]$

In the second half of the table two identical systems are compared, both described by the incompletely specified SQDEs in figure 1(c). In this case the initial relative interval values are range($\Delta h(pc_0)$) = $[-4, -2]$ and range($\Delta v(pc_0)$) = $[16, 20]$, which work in different directions. Can we tell whether the higher initial velocity compensates the lower initial height, even though our knowledge of the systems is incomplete? The results show that the maximum height is greater by $[44.6, 136.8]$ in the second system (range($\Delta h(pc_1)$)), so that the higher velocity compensates the lower height. In this case, CEC* generates 15 comparative behaviors and does not unambiguously answer the question. After combining the comparative behaviors with the SQCA output, only 3 remain.

Omitting the RIV constraints from proposition 2 does not influence the results in the first example. However, for the incompletely specified models in the second example SQCA gets worse results without these constraints: range($\Delta a(pc_2)$) = $[-1.55, 3.87]$ instead of range($\Delta a(pc_2)$) = $[-1.55, 3.56]$. In both examples we obtain worse results when the RIV constraints from proposition 3 are omitted. This shows that semi-quantitative CA cannot be reduced to the trivial approach of subtracting simulation values at pairs of comparison (proposition 1).

SQCA has been tested on a number of examples, including brittle fracture systems in fracture mechanics and prey-predator systems in population ecology. It successfully answers CA questions involving structural differences in combination with differences in the initial conditions of the systems.

Application

SQCA will be embedded in a system for the automated prediction of experimental results in the domain of material science. This system is currently being developed at the University of Twente, as a part of a research project directed at the model-based analysis of scientific measurements (see also de Jong *et al.* (1998)).

Suppose a database or a knowledge base with experimental results is available, a so-called measurement base. More specifically the base consists of property measurements obtained in experiments, supplemented by a description of the experimental circumstances. The description contains general information about the type of experiment, the geometry of the system and environmental conditions. We refer to this description as an *experimental context*. The reported measurements are assumed to bound the true value of the corresponding quantity, that is, the value that would have been obtained in an ideal experiment. Suppose a user is interested in an experimental result which is not in the measurement base. That is, no experiment matching the experimental context specified by the user is found. Is it possible to give an estimation of the desired value by means of the other measurements stored in the measurement base? We suggest an approach that combines techniques for (automated) model construction, semi-quantitative simulation and comparative analysis.

Assume the following question is posed: "What is known about the property measurement m_d, resulting from an experiment carried out in the desired experimental context $\mathcal{E}C_d$?". Suppose this unknown measurement is bounded by the numerical interval $[\underline{m}_d, \overline{m}_d]$. First, the measurement base is consulted for a direct answer, i.e., for a measured value of the same property, performed in an experiment specified by an *actual experimental context* $\mathcal{E}C_a$ that completely matches $\mathcal{E}C_d$. If the measurement base is not able to provide a direct answer, it is scanned for experimental contexts describing conducted experiments of the same type. Suppose a measurement m_a with bounds $[\underline{m}_a, \overline{m}_a]$ of the same property is found with the experimental context $\mathcal{E}C_a$.

In order to make a prediction of m_d, given $\mathcal{E}C_d$, $\mathcal{E}C_a$ and m_a, an approach that consists of the following major steps can be followed (figure 4):

1. Use the actual $\mathcal{E}C_a$ and the desired $\mathcal{E}C_d$ experimental contexts to automatically construct adequate[3] mod-

[3]Informally, a model is adequate if at least one of the simulated behaviors gives correct and precise enough bounds for the measurement of interest.

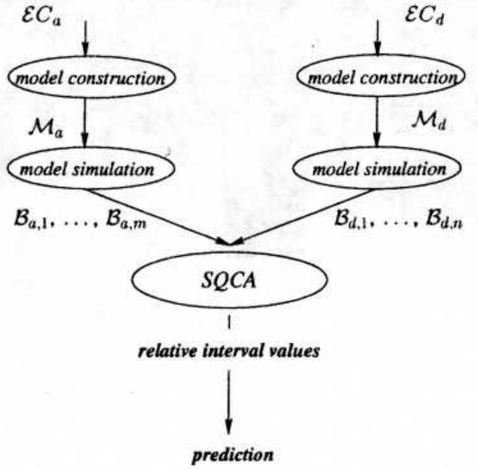

Figure 4: General approach to the automated prediction of experimental results

els $\mathcal{M}_a$ and $\mathcal{M}_d$.

2. Simulate the models $\mathcal{M}_a$ and $\mathcal{M}_d$ to obtain two sets of possible behaviors: $\mathcal{B}_{a,1}, \ldots, \mathcal{B}_{a,m}$, and $\mathcal{B}_{d,1}, \ldots, \mathcal{B}_{d,n}$.
3. Use SQCA to compare each pair of behaviors $\mathcal{B}_{a,i}$ and $\mathcal{B}_{d,j}$ from the sets determined in the previous step.
4. Combine the resulting RIVs to infer interval bounds for m_d.

If $s = n \times m$ pairs of behaviors are compared, SQCA will infer a set of relative interval values $\text{range}(\Delta m)_{out,k} = [L_k, U_k]$, $1 \leq k \leq s$. The interval $[\underline{m}_d - \overline{m}_a, \overline{m}_d - \underline{m}_a]$ representing the actual difference between the two measurements will be contained in the output value $\text{range}(\Delta m)_{out,k}$ (theorem 2). We therefore conclude that

$$m_d = [\underline{m}_d, \overline{m}_d] \subseteq \cup_{k=1}^{s} [L_k + \overline{m}_a, U_k + \underline{m}_a].$$

An estimate of m_d can be obtained by performing semi-quantitative simulation only. The suggested approach improves the bounds resulting from simulation by using knowledge about measurements of the same property obtained in similar experimental contexts. Further, SQCA assures that the estimation of the unknown property value will be contained in the range obtained by simulation (theorem 2 and proposition 1).

The algorithm is straightforward to generalize when more than one experiment in the measurement base matches $\mathcal{E}C_d$. The predicted value of m_d will be estimated by the intersection of the resulting interval bounds. Assuming that the measurement base is free of errors, this intersection will not be empty (for error identification, see de Jong *et al.* (1998)). It is also possible that several adequate models for each experimental context are constructed. In this case all models determined from $\mathcal{E}C_d$ have to be simulated and each of the predicted behaviors has to be compared with the simulated behaviors of the models determined from $\mathcal{E}C_a$.

Semi-quantitative reasoning is essential in the described approach. Incomplete descriptions of experiments occur in a variety of cases when either precise numerical information is not available or one would like to abstract to more general cases. Further, measurements are usually reported in the form of intervals, more specifically confidence intervals obtained from statistically processing a set of individual measurements of a property. SQCA plays a crucial role because it allows one to deal with incompletely specified experiments while exploiting numerical information to obtain predictions of property values as precise as possible. In addition, SQCA provides guarantees on the correctness of these predictions.

Discussion and related work

SQCA borrows ideas from both semi-quantitative simulation and qualitative comparative analysis. As in Q2+Q3, the problem is reduced to a constraint propagation problem. However, SQCA employs constraints dealing with ranges of value differences instead of ranges of values. The constraints are derived from a pair of models instead of a single model, with the additional complication that $SQDE$ and $\hat{SQDE}$ may be structurally different and fragments of SQB and $\hat{SQB}$ unsynchronized.

To our knowledge, only de Mori and Prager (1989) have studied the semi-quantitative comparative analysis of dynamical systems, but their approach is restricted only to linear, time-invariant systems and employs semi-quantitative information on a coarser level of granularity. Moreover, unlike SQCA their technique for qualitative perturbation analysis cannot deal with structural differences between systems and with topologically different behaviors.

The application of SQCA presented in this paper aims at the prediction of unknown property measurements by given a hypothetical description of an experiment. In this respect, the approach can be compared with approaches to answering prediction questions (e.g., Rickel & Porter (1994)). Such methods provide an answer of a query about the behavior or the value of a certain quantity in a dynamical system by a specified scenario description. Answering the query is achieved by combining automated model construction and simulation. When the question asks for a value of a quantity the predictions obtained by semi-quantitative simulation alone are often unnecessarily broad intervals. In our approach this inaccuracy is reduced by comparing the model of the hypothetical experiment with models of performed experiments of similar type.

Given that the SQCA input is valid, the relative interval values $\text{range}(\Delta x(pc))_{out}$ produced by SQCA contain the actual difference Δx at pc (soundness). However, they may overestimate this value due to a loss of information in the process of generating and propagating constraints (incompleteness). By using techniques for the solution of interval CSPs that are more powerful than the constraint propagation algorithm currently employed in SQCA (e.g., Benhamou & Older (1997)), the problem of excess width could be reduced. Also, the RIV constraints defined by proposition 3 can be improved by replacing in some cases the mean value theorem with explicit integration (Vescovi, Farquhar, & Iwasaki 1995).

Even when the predicted RIVs are as tight as possible given the input, they may turn out not to be precise enough. The imprecision of the results is inherent to the incomplete models of the systems we want to compare. For instance, broad intervals for the initial time-points of the two behavior fragments being synchronized introduce wide intervals for the auxiliary time-points and hence for the corresponding relative interval values. The convergence theorem shows that by interpolating additional auxiliary time-points in the SQBs, and thus introducing new auxiliary pairs of comparison and new RIV constraints, we can improve the results of SQCA. This suggests an approach in which the precision of SQCA's predictions is dynamically increased by iterating between semi-quantitative simulation and comparative analysis.

Conclusions and further work

SQCA is a technique for the semi-quantitative analysis of dynamical systems which is both able to deal with incompletely specified models and arrive at precise predictions by exploiting available numerical information. The technique has a solid mathematical foundation which facilitates proofs of correctness and convergence properties. SQCA has been fully implemented, including the derivation of propagation constraints.

Future work will concentrate on the improvement of the precision of the technique, along the lines mentioned in the previous section, and its integration into a system for the automated prediction of experimental results in the domain of material science.

References

Benhamou, F., and Older, W. 1997. Applying interval arithmetic to real, integer, and boolean constraints. *J. Logic Programming* 32:1–24.

Berleant, D., and Kuipers, B. 1997. Qualitative and quantitative simulation: bridging the gap. *Artificial Intelligence* 95:215–256.

de Jong, H., and van Raalte, F. 1997. Comparative analysis of structurally different dynamical systems. In *Proceedings IJCAI-97*, 486–491.

de Jong, H.; de Wit, V.; Mars, N.; and van der Vet, P. 1998. Model-based identification of systematic errors in measurements. In *Working Notes QR-98*, 32–41.

de Mori, R., and Prager, R. 1989. Perturbation analysis with qualitative models. In *Proceedings IJCAI-89*, 1180–1186.

Kay, H., and Kuipers, B. 1993. Numerical behavior envelopes for qualitative models. In *Proceedings AAAI-93*, 606–613.

Kuipers, B. 1994. *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge.* MIT Press.

Moore, R. 1979. *Methods and Applications of Interval Analysis.* SIAM.

Neitzke, M., and Neumann, B. 1994. Comparative simulation. In *Proceedings AAAI-94*, 1205–1210.

Rickel, J., and Porter, B. 1994. Automated modeling for answering prediction questions: Selecting the time scale and system boundary. In *Proceedings AAAI-94*, 1191–1198.

Vatcheva, I., and de Jong, H. 1999. Semi-quantitative comparative analysis. Memorandum UT-KBS-98-12, University of Twente, Enschede, The Netherlands.

Vatcheva, I. 1998. Semi-quantitative simulation with Q3: An implementation. Memorandum UT-KBS-98-12, University of Twente, Enschede, The Netherlands.

Vescovi, M.; Farquhar, A.; and Iwasaki, Y. 1995. Numerical interval simulation. In *Proceedings IJCAI-95*, 1806–1812.

Weld, D. 1988. Comparative analysis. *Artificial Intelligence* 36:333–374.