submitted to: the third international workshop
on Qualitative Physics

also

submitted to **IJCAI 89**
area: B6. Patterns of Commonsense Reasoning

# Stratification – A New Method of Analyzing Discontinuous Change


Toyoaki Nishida and Shuji Doshita
Department of Information Science
Kyoto University
Sakyo-ku, Kyoto 606, Japan
phone: 81-75-751-2111 ext. 5396
email: nishida%doshita.kuis.kyoto-u.junet%japan@relay.cs.net


December 7, 1988


**Abstract**

In this paper, we describe *stratification*, a new method of reasoning about discontinuous change. In stratification, we make use of a family of more detailed models with a control parameter and predict the outcome of discontinuous change by investigating and summarizing the behavior of a detailed version of a given model in an extreme. Stratification consists of four techniques of: (a) recognizing qualitatively different region in phase portrait when the value of the control parameter becomes sufficiently small or large, (b) making transition analysis between recognized regions, (c) extending a history of a trajectory to the past and the future, and (d) abstracting the result by elminating those states which persist only for infinitesimal period of time. Stratification serves as an accurate method of analyzing discontinuous behaviors without sacrificing the efficiency.

**Stratification – A New Method of Analyzing Discontinuous Change**

Toyoaki Nishida and Shuji Doshita
Kyoto University

# 1 introduction

Notion of discontinuous change is important in characterizing behaviors of dynamical systems from an abstract point of view. In a previous paper [6], we studied a couple of methods of analyzing discontinuous changes caused by piecewise linear differential equations. One, called the direct method, is make use of heuristic rules for predicting occurence and outcome of discontinuous change. The other is called the approximation method which handles discontinuous changes as if they were very rapid continuous changes. The former was preferred because the expected amount of computation was less than the latter and no detailed model is required for prediction. The former method was implemented and worked well for many cases.

However, there does exist an example for which the direct method fails. This is due to the lack of information for predicting the outcome of discontinuous change. Compelete analysis is not possible for such cases unless more information is provided from more detailed model.

In this paper, we explore *stratification*, a technique of predicting the outcome of discontinuous change by investigating and summarizing the behavior of a detailed version of a given model in an extreme. Like the approximation method, the stratification technique presumes the use of detailed models. This makes the internal structure of discontinuous change visible, resulting in more accurate prediction.

# 2 Stratification and Phase Space Analysis

Like recent approaches to qualitative reasoning, stratification is based on phase space analysis. Mathematical theory of phase space analysis is dynamical systems theory, which is usually referred to as a "qualitative theory of ordinary differential equations" [3]. Dynamical system theories derive qualitative properties of dynamical systems by investigating geometric features of phase portrait. In qualitative reasoning, such approach enables to take into account the consistency of overall phase portrait, contributing to suppressing spurious behaviors which are hard to eliminate in conventional frameworks.

## 2.1 Phase Space Analysis

Dynamical system theories use the notion of *phase space*, the Cartesian product of *state variables*, and identify a state of a system of differential equations with a point in the phase space.

A system of differential equaions defines a *vector field* on the phase space, specifying how the state should evolve at each point. A *trajectory* is a trace of a state change over time, and it can be viewed as a geometric representation of a solution to the differential equations defining the vector field. In fact, there are obvious correspondence between geometric properties of trajectories and dynamic properties of solution. For example, a closed trajectory corresponds to a repeated behavior of the system, a fixed point an equilibrium, an attractor a stable oscillation exhibited after sufficiently long elapse of time, *etc.* Thus, we can study geometirc properties of *phase portrait*, a collection of trajectories, to capture of dynamic properties of differential equations. This is the approach taken by dynamical system theories.

A couple of important properties of phase portrait are the non-intersection of different trajectory and no branching of any single trajectory. Both of them are derived from the uniqueness of solution of differential equations. Struss [10] and Lee and Kuipers [5] made use of these properties to tame spurious behaviors predicted by QSIM [4].

## 2.2 Modeling Discontinuous Change in Phase Portraits

The most straightforward way to introduce discontinuous change in phase space based approach is to allow discontinuous change to occur in phase space. However, this implies that we have to abandon the advantages of strong constraints of phase portraits and hence it is not attractive.

Alternatively, we take an approach of making use of a family of more detailed models with a control parameter.[1] The family of more detailed models are chosen so that discontinuous changes of the original model may be approximated in each model as continuous change, and the degree of accuracy increases as the value of the control parameter is decreased.[2] *Stratification* is a technique of reasoning about discontinuous change by predicting and summarizing what will happen to the phase portrait when the control parameter becomes extremely small.

Before going into the stratification technique, we point out that our approach is at least feasible in the sense that we can actually obtain a family of more detailed models in most circumstances, by taking into account time delay that has been neglected in less detailed models [7].

However, this might introduce new kind of indeterminacy, since it means the introduction of one or more new state variable, which in turn means the increase in dimension of phase space.[3]

This criticism does not apply to stratification, since stratification is in fact a technique for reducing the dimension *before* behavior is predicted.

# 3 The Idea behind Stratification

Stratification makes use of an interesting property that can be oberved when one makes the value of a parameter smaller and smaller.

[1]For simplicity of discussion, we assume the control parameter to be zero or positive in this paper. Of course, this is not an essential limitation. In actual implementation, we do not pose such constraint.

[2]Again, this is only for simplicity of discussion; actual implementation does not impose such constraint.

[3]This is similar to the problem pointed out by de Kleer as a side effect of the "push-a-level" approach [1].

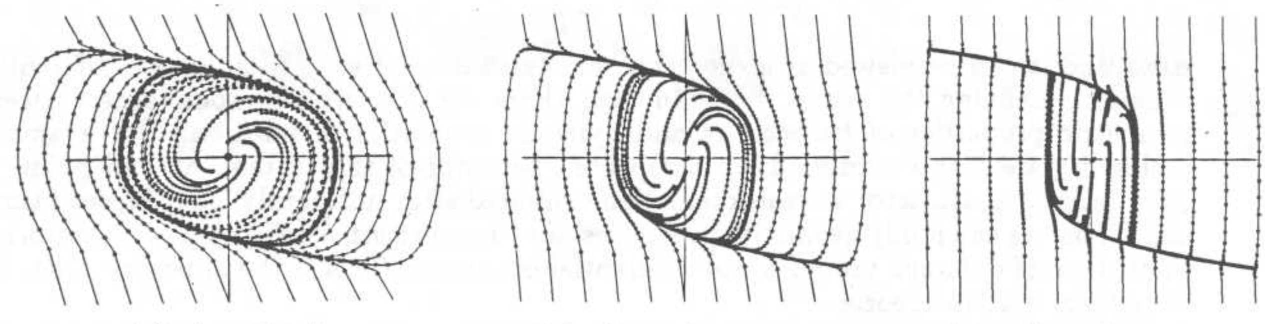


Figure 1: Phase Portraits of the Equations (1) with Different Value of the Control Parameter $\epsilon$

Consider the following family of equations with a positive parameter $\epsilon$:

$$\begin{cases} \dot{y} = (-y^3 + y - x)/\epsilon \\ \dot{x} = y \end{cases} \qquad \text{where, } \dot{\alpha} \equiv d\alpha/dt \qquad (1)$$

With each $\epsilon$, we have specific differential equations. What will happen if we make $\epsilon$ sufficently small? Figure 1 demonstrates several results of numerical simulation. As $\epsilon$ becomes smaller, the whole phase space is divided into several qualitatively different regions. In some region such as *A1*, the flow becomes so fast that the transition of the area becomes almost instantaneous, while in some other region such as *A2* the flow is not heavily affected by the change of the control parameter and is moderate in speed.

The region of slow flow becomes smaller as the value of the control parameter is made smaller. In two dimensional case, the slow region converges to a thin line, and in three dimensional case, the slow region comes to form a thin plane, "a stratus".[4] Although all of these are visible and understandable to humans, it is not so trivial for programs. Stratification is a technique which enables programs to understand what is going on in the phase space if the control parameter is exaggerated. As shown below, stratification is a symbolic technique rather than numerical. Thus, the result of stratification is not affected by numerical errors. This is mainly because reasoning about limit is rather a matter of logical inference.

# 4 Stratification Techniques

Stratification consists of four techniques:

1. a technique of recognizing qualitatively different regions in a phase portrait when the value of the control parameter becomes sufficiently small
2. transition analysis between recognized regions
3. trajectory tracking which extends a history of a trajectory to the past and the future

[4]This is the origin where the terminology stratification comes from.

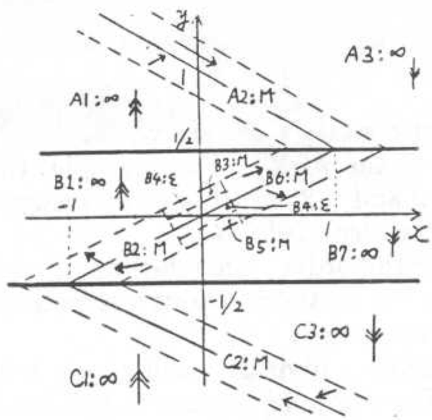


Figure 2: Division of the Phase Space

4. abstraction technique which will elminate those states which persist only for infinitesimal period of time

In what follows, we concentrate on a limited version of stratification which applies to piecewise linear differential equations. By this convention, we can temporalily step aside from the problem of devising symbolic representations for general geometric opbjects, which is really a hard issue and seems almost impossible [2].

We use the following piecewise linear approximation of Van der Pol's equation as example:

$$\left[ \begin{array}{l} 1/2 < y \left\{ \begin{array}{l} \dot{y} = (-2y + 2 - x)/\epsilon \\ \dot{x} = y \end{array} \right. \\ -1/2 \leq y \leq 1/2 \left\{ \begin{array}{l} \dot{y} = (2y - x)/\epsilon \\ \dot{x} = y \end{array} \right. \\ y < -1/2 \left\{ \begin{array}{l} \dot{y} = (-2y - 2 - x)/\epsilon \\ \dot{x} = y \end{array} \right. \end{array} \right. \tag{2}$$

The phase space for this equation is two dimensional. We discuss stratification resulting from making the free parameter $\epsilon$ approach to *zero*.

## 4.1 Recognizing Qualitatively Different Regions

The first stage is to divide the phase space by order of magnitude of vectors. The regions in which order of magnitude of vectors is very small, medium, and very large are called *slow*, *moderate*, and *fast* regions, respectively. Figure 2 shows how the phase space for (2) should be divided. Signs $\epsilon$, $M$, and $\infty$ associated with each region represent slow, moderate, and fast regions, respectively. Note that unlike conventional boundary-based approaches, the boundaries between those regions are vague, for the distinction in order of magnitude is not a matter of quantitative difference. In figure 2, we used thick and broken lines to draw exact and broken boundaries, respectively.

Some regions gradually reduce in dimension and become almost $n-1$ dimensional regions, as the value of the control parameter is decreased. For example, the two dimensional region *A2* becomes thinner as $\epsilon$ approaches to *zero*. Formally, we call a regions a *stratus* if the region becomes arbitrarily thin as the control parameter becomes small, and the region does not contain any *zero* vector, and some flow is oriented in parallel to the longer boundaries of the region. For example, the phase portrait shown in figure 2 contains two strata, *A2* and *C2*, which can be recognized in this stage. We will refer to the flow in a stratus as a *stratified* flow.

In order to compute the order of magnitude of each vector, we use conventional symbolic formula manipulation methods, though sometimes it is possible to obtain an answer solely by symbolic manipulation.

As a result of phase space division, each region is represented symbolically as a set of attribute-value pairs. For example, regions *A2* and *A3* are represented as follows:

$$
\begin{array}{lll}
\text{region } A2: & \text{type : } \textbf{stratus} & \\
& \text{characterization : } & [-2y+2-x] = -\epsilon \sim +\epsilon \\
& & [y-1/2] = +\epsilon \sim +\infty \\
& \text{adjacent-regions : } & \text{A3} : [-2y+2-x] \searrow -M \\
& & \text{A1} : [-2y+2-x] \nearrow +M \\
& & \text{B6} : [y-1/2] \searrow 0 \\
& \text{flow-vectors : } [\dot{x}] = +M, [\dot{y}] = -M \sim +M & \\
& \text{flow-rate : } \textbf{moderate} & \\
& \ldots &
\end{array} \tag{3}
$$

$$
\begin{array}{lll}
\text{region } A3: & \text{type : } \textbf{2-dimensional open region} & \\
& \text{characterization : } & [-2y+2-x] = -\infty \sim -M \\
& & [y-1/2] = +\epsilon \sim +\infty \\
& \text{adjacent-regions : } & \text{A2} : [-2y+2-x] \nearrow -\epsilon \\
& & \text{B7} : [y-1/2] \searrow 0 \\
& \text{flow-vectors : } [\dot{x}] = +M, [\dot{y}] = -\infty & \\
& \text{flow-rate : } \textbf{fast} & \\
& \ldots &
\end{array} \tag{4}
$$

The characterization attribute specifies the region as a set of linear equations and inequalities. The adjacent-region attribute indicates adjacent regions immediately accessible from the current region, together with conditions of the transition.

## 4.2 Transition Analysis

The purpose of transition analysis is to enumarate possible fragments of trajectories by considering possible state transitions in each region. This is done based on local information. When possible, we use quantitative information specifying the orientation of vectors or boundaries to make prediction precise.

Several rules for transition analysis, partly shown below, refer to order of magnitude information, contributing to further resolving ambiguity. For example,

**Rule 1** *Each component of a vector should change smoothly in transition.*

By "smooth change" we mean that order of magnitude should change to an adjacent level at a transition. For example, a state is allowed to directly transition from a fast region to a moderate region, while it is not so to a slow region.

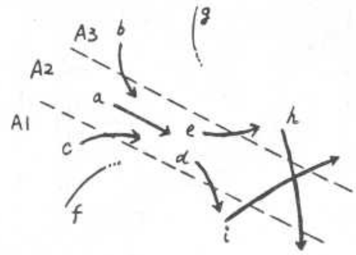


Figure 3: Effects of Rules that Make Use of Order of Magnitude Information

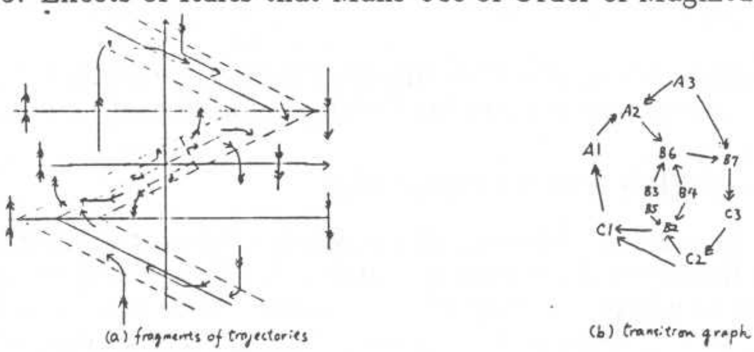


Figure 4: The Result of Transition Analysis for our Example

**Rule 2** **If** *a vector field is fast and does not allow vectors of all orientations,* **then** *the state will transition to one of adjacent regions in an infinitesimal period of time.*

**Rule 3** *At boudaries between a fast region and a moderate region, the orientation of trajectories is governed by that of vectors in the fast region.*

Figure 3 how they work for our example. Of nine possible fragments of trajectories, four fragments *d*, *e*, *h* and *i* are eliminated due to **rule 3**, and *f* and *g* are eliminated due to **rule 2**. As a result, we have only three fragments left. Figure 4 shows portion of the result of transition analysis for our example.

## 4.3 Trajectory Tracking

This stage connects fragments of trajectories, extends them to the "past" and "future", identifies their type, and enumerates types of trajectories with qualitatively different nature. Global constraints such as non-intersection constraints [5,10] are taken into account to eliminate spurious transitions.

Two issues are crucial here: use of properties of stratified flows, and subdivision and merge of regions. For convenience of further reference, we first show the result of trajectory tracking in figure 5. Though we mainly refer to the utility of topological information below, we equally keep track of metric information in trajectory tracking. In particular, left-right relation of two trajectories is important, since it remains invariant in regions

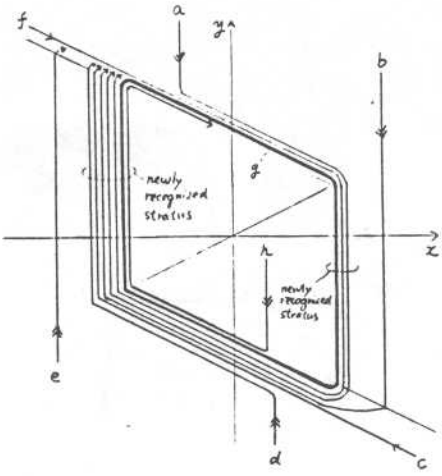


Figure 5: Result of Trajectory Tracking

where the orientation of vector is uniform. For example, we predict that trajectory a runs in the left side of trajectory e after they first meet together at stratus *A2*.

### 4.3.1 Stratified Flow and its Properties

It is important to grasp the behavior of trajectories before and after long enough period in time. Let us introduce the notion of a limit set. Given a trajectory $T$, let $L_\alpha$ and $L_\omega$ be sets of points to which $T$ can approach arbitrarily near when $t$ becomes $-\infty$ and $\infty$, respectively. $L_\alpha$ and $L_\omega$ are called a *$\alpha$ limit set* and *$\omega$ limit set*, respectively. (see chapter 9 and 11 of [3] for more formal formulation.) We classify trajectories into four categories by whether or not their $L_\alpha$ and $L_\omega$ are bounded. We indicate the category by placing a prefix consisting of two characters: each is either **u** (for unbounded) and **b** (for bounded), and the first character is for $L_\alpha$ and the second for $L_\omega$. For example, trajectories with bounded $L_\alpha$ and unbounded $L_\omega$ are called **bu**-trajectories. A closed orbit forms a special subcategory in **bb**-trajectories; a trajectory itself, and its $L_\alpha$ and $L_\omega$ concord to each other.

It is also useful to introduce the notions of *positively* and *negatively invariant* sets in phase space. Set $P$ of points in a phase space is *positively invariant* if a half trajectory starting from starting from any point in $P$ is totally contained in $P$. Negative invariant set is defined similarly, except that a half trajectory coming to the point is argued. The following theorem is important:

**Theorem 1** *A non-empty compact set which is either negatively or positively invariant contains in it either an equilibrium point or a limit cycle.* ... derived from Poincaré-Bendixon's theorem (see chapter 11 of [3])

The stratus plays an important role in identifying the type of trajectories. The following rule is useful in determining the type of trajectories, though its applicability is limited:

**Rule 4** **if** *the phase space is two dimensional*
**and** *a trajectory enters a stratus more than once*
**then** *the trajectory is not of* **uu** *category.*

Roughly speaking, validity of this rule follows from **theorem 1** plus the following argument: if a trajectory enters the same stratus twice, then it is either the case that *(a1)* there exists a positively invariant set $P$ surrounded by the edge of the stratus and the fragement of the trajectory, each of which is delimited by the two entrances, and *(a2)* the trajectory enters $P$, or that *(b1)* there exists a negatively invariant set $N$ surrounded by the local section of the stratus and the fragement of the trajectory, each of which is delimited by the two entrances, and *(b2)* the trajectory has exited from $N$.

We can identify the category more precisely using the following rule:

**Rule 5** **if** *the phase space is two dimensional*
**and** *a trajectory enters a stratus in which it entered previously*
**and** *the flow is oriented* towards *the area surrounded by the the local section of the stratus and the fragement of the trajectory, each of which is delimited by the two entrances*
**then** *the trajectory is of either* **bb** *or* **ub** *category.*

Similar rule exists for the remaining case. These rules have enabled to conclude that trajectories a to *f* in figure 5 are either of **bb** or **ub** category.

The following rule is useful in finding peculiar trajectories.

**Rule 6** **if** *the phase space is two dimensional*
**and** *the orientation of flow at the two opposit boundaries is in*
**then** *there exists a trajectory in parallel to the stratus; if the stratus is infinitely long, the correspoding end of the trajectory is unbounded.*

This rule has allowed to find trajectories c and *f* in figure 5. The following rule is a direct conclusion of Poincaré-Bendixson's theorem.

**Rule 7** **if** *a non-empty limit set* S *in two-dimensional phase space of* $C^1$ *does not contain any equilibrium point*
**then** S *is a closed orbit*

This rule is powerful enough to enable to find a closed orbit *g* in figure 5.

### 4.3.2 Subdivision and Merge of Regions

In the course of trajectory tracking, we subdivide those regions which have more than one region as a destination of immediate transition. Since unmotivated forward application of subdivision may create endless loop of subdivision, we do it only when it is necessary. One important case is when a flow in parallel to the stratus enters a fast region, for such a flow usually forms a stratified flow in the fast region, too. Thus, we identify a stratified flow in regions *B1* and *B7* in figure 2, resulting from outflow from regions *A2* and *C2*, respectively. As a result, regions *B1* and *B7* each are divided into three.

On the contrary, any adjacent regions with the same characteristic on vector are merged into one. As to our example, *A3* and one of subdivided region of *B7* are merged, since vectors in both of these regions are characterized as $(+\epsilon, -\infty)$ by its component.

## 4.4 Abstration

In the abstraction stage, we neglect regions where trajectories stay only for an infinitesimal period. Figure 6 illustrates the final result of stratification applied to our example. This allows us to capture the internal structure of discontinuous changes rather concisely.

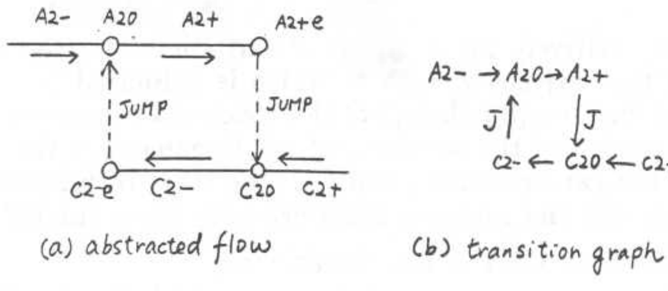


Figure 6: Result of Abstraction

# 5 Extension to General Cases

In the previous section, we presented a limited version of stratification by concentrating on two-dimensional phase portraits for piecewise linear approximations. In fact, these two limitations are very useful to make the problem tractable.

The first limitation does not seem hard to remove, whereas it is harder to remove the second. The key to the solution seems to be development of powerful representation of spatial concepts and classification of shape of trajectories on the representation.

# 6 Related Work

This work was inspired by a number of predecessors in qualitative reasoning: phase space analysis [9,5,10,12] for the basic framework of stratification, qualitative kinematics [2] for qualitative geometric analysis of space, order of magnitude reasoning and exaggeration [8,11]. However, no previous work in qualitative reasoning had provided a sufficently general framework of reasoning about discontinuous change. The approach presented in this paper is novel as a qualitative and computational bifurcation theory.

The theoretical work related to this work is dynamical systems theories [3], in particular, marvelous work by E. Zeeman [13] on modeling heart beat. This work is characterized as an attempt to making mathematical reasoning which underlies dynamical system theories computational.

# 7 Concluding Remarks

Stratification is a technique for reasoning about discontinuous change without paying much extra cost, for the stratification takes advantage of the property that each $n$-dimensional stratus is almost an $n-1$ dimensional region. In this paper, we have described computational techniques underlying stratification and demonstrated how it is applied to grasp dynamical behavior caused by Van der Pol's equation. Further work is remaining as to extending the work to higher order dimensional phase spaces and general nonlinear problems.

# References

[1] Johan de Kleer and John Seely Brown. Qualitative physics based on confluences. *Artificial Intelligence*, 24:7–83, 1984.

[2] Kenneth D. Forbus, Paul Nielsen, and Boi Faltings. Qualitative kinematics: a framework. In *Proceedings IJCAI-87*, pages 430–435, 1987.

[3] Morris W. Hirsch and Stephen Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, 1974.

[4] Benjamin Kuipers. Causal simulation. *Artificial Intelligence*, 26:289–338, 1986.

[5] Wood W. Lee and Benjamin J. Kuipers. Non-intersection of trajectories in qualitative phase space: a global constraint for qualitative simulation. In *Proceedings AAAI-88*, pages 286–290, American Association for Artificial Intelligence, 1988.

[6] Toyoaki Nishida and Shuji Doshita. Reasoning about discontinuous change. In *Proceedings AAAI-87*, pages 843–848, American Association for Artificial Intelligence, 1987.

[7] Toyoaki Nishida, Atsushi Takeshita, and Shuji Doshita. A heuristic method of guiding model inference. In *Proc. of International Computer Conference '88*, 1988.

[8] O. Raiman. Order of magnitude reasoning. In *Proceedings AAAI-86*, pages 100–104, American Association for Artificial Intelligence, 1986.

[9] Elisha Sacks. Piecewise linear reasoning. In *Proceedings AAAI-87*, pages 665–659, American Association for Artificial Intelligence, 1987.

[10] Peter Struss. Global filters for qualitative behaviors. In *Proceedings AAAI-88*, pages 275–279, American Association for Artificial Intelligence, 1988.

[11] Daniel S. Weld. Exaggeration. In *Proceedings AAAI-88*, pages 291–295, American Association for Artificial Intelligence, 1988.

[12] Kenneth Man-kam Yip. Generating global behaviors using deep knowledge of local dynamics. In *Proceedings AAAI-88*, pages 280–285, American Association for Artificial Intelligence, 1988.

[13] E. C. Zeeman. Differential equations for the heartbeat and nerve impulse. In M. M. Peixoto, editor, *Dynamical Systems*, pages 683–741, Academic Press, 1973.