CRITICAL HYPERSURFACES AND THE QUANTITY SPACE

Mieczyslaw M. Kokar
Northeastern University
360 Huntington Avenue
Boston, Massachusetts 02115
KOKAR%NUHUB.ACS.NORTHEASTERN.EDU@RELAY.CS.NET

ABSTRACT[1]

Qualitative reasoning about physical processes is based on the notion of "quantity space" [Forbus, 1984a, 1984b]. The question is how to construct a quantity space for a particular physical process. One line of research is to establish a set of so called "landmark points" by selecting some values of the continuous physical parameters characterizing the physical process under consideration [Kuipers, 1985a, 1985b]. The landmark points are to delimit operating regions of qualitative processes. In most practical situations it is impossible to find a finite set of such points. This is because the operating regions of physical processes are delimited not by some specific values of physical parameters but by some hypersurfaces in the cross-product of the parameters, they are called here "critical hypersurfaces". The paper presents a relatively complete methodology for establishing critical hypersurfaces.

1. INTRODUCTION

Qualitative reasoning about physical processes gained much of attention in the last several years. Not only this reflects a general spirit of AI, which is symbolic reasoning, but it complements the existing methodology of modelling and simulation of physical processes, which was limited to quantitative analysis of numerical models. Experience shows us that in many cases quantitative simulations are not feasible because of a very high complexity of the quantitative models. Qualitative simulation can then come to the rescue with methods that generate some less specific results, but which are feasible. This is not the only reason for using qualitative simulations. In many situations quantitative simulations are feasible, but not necessary. If we are interested whether a particular physical parameter is going to stay within its allowable range following some changes in the controls, then we should try to answer just this question and not try to calculate the exact value of this parameter. This exact value would be discarded anyway and only the qualitative information that "the parameter will (or will not) stay within the range" will be utilized to make a control decision. In such a case why should we perform the full quantitative analysis when several simple logical operations might do it.

On the other hand, one should not go into another extreme and try to solve all the problems with qualitative methods only. Trying to resolve all inequalities in predicting behavior of a physical process would inevitably lead to simulations of differential equations and real numbers. And this would definitely lead into higher complexity problems than when using classical quantitative methods, even though defining a quantitative parameter requires a measurement method that is defined in terms of qualitative operations. Ultimately, there is a need for understanding relations between qualitative and quantitative methodologies. Integration of quantitative and qualitative methodologies is one of the key issues of this paper.

In the qualitative simulation methodology states of processes are characterized by some parameters which can take on a limited number of nominal values. These values are usually related to some quantitative parameters. The relationships among the qualitative parameters are described in terms of "quantity space" [Forbus, 1984a, 1984b]. When the relations among these parameters change, some "processes" are started or stopped. Kuippers [Kuipers, 1985] uses terms like "critical points", "landmark points", or "characteristic points" to describe some specific values of physical parameters. Qualitative simulation methodology involves moving from one qualitative state to another, or from one set of critical points characterizing a given physical system to another. The question is, however, how do we establish such a quantity space. In the literature on qualitative simulations the critical points are selected on the base of some semantics related to the values of physical parameters. The examples of those are boiling temperature, melting temperature, etc. The qualitative processes start or stop when the inequalities between the physical parameters, like temperature, and those critical points, change their signs. In fact, this is a very restricted approach. These processes usually depend not only on the particular physical parameters but on the relationships among them. For instance the water temperature 100 C does not necessarily mean the

[1] This research was supported in part by the Center for Electromagnetics Research, Northeastern University. The Center is supported by NSF and a number of industrial sponsors.

water is boiling, it depends what is the pressure, too. Therefore, to be able to predict the behavior of a process we need to know not the critical points of the physical parameters, but the sets of points fulfilling a specific constraint. We will call these sets "critical hypersurfaces". A more precise definition of critical hypersurface will be given in Section 2. Intuitively, a critical hypersurface will be understood as a subset in a cross-product of continuous physical parameters. A critical hypersurface constitutes a "distributed landmark". The position of a particular state with respect to such hypersurfaces determines whether a particular process is active or not.

Given a description of a physical process and a semantics for qualitative reasoning, one can establish critical hypersurfaces in many ways. For instance, one can use differential equations for this purpose (provided such models are available). Another possibilty is to try finding the critical hypersurfaces experimentally. This would involve an amount of effort far beyond the advantages from the qualitative simulation. This paper presents a methodology for finding critical hypersurfaces which does not require such strong models as differential equations. This methodology is based on the theory of dimensional analysis.

Critical hypersurfaces can be utilized in a similar way critical points are used in qualitative simulations. This problem is beyond the scope of this paper.

2. CRITICAL HYPERSURFACES vs. CRITICAL POINTS

In this section we show an example of a physical process in which the notion of critical hypersurfaces plays a significant role. Then we formally define this notion.

As the example we consider the process of viscous fluid flow in a pipe. It is a known fact in fluid mechanics ([Young, 1964], [Monin and Yaglom, 1973]) that the flow can be in one of at least three different qualitative states: laminar, turbulent, or transitional (unstable). In each of these states the process of fluid flow is described by qualitatively different mathematical models. Intuitively, the flow is laminar for small velocities, turbulent for large, and transitional for some interval of velocity in between. Therefore, one could try to find two critical velocities, v1, and v2, delimiting the three flow regions. Unfortunately, this can work only for a very limited range of situations, namely when the pipe diameter D, fluid density ρ, and fluid viscosity η, are constant. For instance, for water in room temperature in a 1-cm-diameter pipe, turbulent flow occurs above about 0.2 m/s, while for air the critical speed is of the order of 4.0 m/s [Young, 1964]. Thus the critical velocity in the latter case is 20 times higher than in the former one. This clearly shows that to reason about qualitative states of physical processes one cannot restrict the set of notions to critical points only.

In hydromechanics the transition from laminar to turbulent flow is characterized in terms of so called "Reynolds number". Reynolds number (usually described as R) is defined as a function of the above referenced parameters of density ρ, average velocity v, pipe diameter D, and viscosity η as:

$$R = \rho \cdot v \cdot D/\eta .$$

It is found experimentally that laminar flow occurs whenever R is less than about 2000. When R is greater than about 3000, the flow is nearly always turbulent, and in the region between 2000 and 3000 the flaw is unstable, changing from one form of flow to the other (here we call it "transitional").

The relationships

$$R = 2000$$

and

$$R = 3000$$

describe two hypersurfaces in the space defined by the cross-product of the continuous parameters ρ, v, D, and η. The two hypersurfaces divide the space into three operational regions: laminar, turbulent and transitional.

This example shows that critical hypersurfaces play an important role in qualitative reasoning about physical processes. There are however three major problems with this approach:

- how to know which physical parameters should be considered in the search for critical hypersurfaces,
- how to know what kind of relationships should be taken into account in the search for critical hypersurfaces,
- how to determine what are the critical points of the function that describes critical hypersurfaces (for instance, how do we know that the critical values of R are 2000 and 3000).

These three problems will be referenced to as the complete relevance problem, the relationship problem, and the semantics problem respectively. They are discussed in the following sections.

3. SEMANTICS OF CRITICAL HYPERSURFACES

In this section we concentrate on the third of the problems listed in Section 2. Assume for a while that we know both the parameters and the form of the relationship that describes the hypersurfaces. We need to find a set of critical points of the function describing the relationship.

In the literature on qualitative reasoning there are known several semanics rules: sign semantics, derivatives semantics [deKleer and Brown, 1982], [Forbus, 1984a, 1984b], Kuipers' QSIM semantics [Kuipers, 1985a, 1985b], order of magnitude semantics [Raiman, 1986]. For the example presented in the previous section none of these seem to be right. Neither Reynolds number R or its derivative change their directions of growth but flow changes from laminar to turbulent. Therefore some other approach is required.

One possibility is to find a relationship between a parameter, like Reynolds number, and some observable qualitative parameter for which the semantics of qualitative states is explicit. In our example the character of flow (laminar, turbulent, or transitional) can be determined by directly observing the profile of the speed of flow. As a matter of fact, that is how the relationship defining Reynolds number has been discovered. The bad side of this approach is that the semantics is specific to a particular physical process. It is hard to apply this approach as a general method.

Another approach would be to find a continuous physical parameter functionally dependent on the value of the function defining the hypersurfaces, for which we can utilize one of the general semantics for qualitative states (signs, derivatives, or orders of magnitudes). Once critical points have been established for the new parameter we can transform them to the critical points of the value of the function defining the hypersurfaces. In the case of critical points of Reynolds number we could utilize the dependency of the rate of heat transfer on Reynolds number. It is a known fact that the rate of heat transfer from/to the liquid in a pipe to/from the outside strongly depends on Reynolds number, it is much more intensive for $R > 3000$ than for $R < 2000$. If we take the heat exchange rate as an indicator then we can utilize, for instance, the semantics of sign of derivative, or we can apply Kuipers' semantics to the derivative of the heat transfer rate.

In all of these cases we suggested to use a parameter which depends on the value of function characterizing hypersurfaces. The choice of the parameter depends on which process we are interested in. When we were interested in fluid flow, we took the speed profile. In the second example our attention was concentrated on heat transfer, or more precisely, on the rate of heat transfer. Therefore we selected the characteristic parameter for this process. This seems to be a reasonable methodology.

Mathematically, we are looking for a hypersurface in the cross-product of some parameters x1, ..., xm, described as

$$f(x1, \ldots, xm) = C.$$

A critical hypersurface is defined by fixing the value of C. If we are interested in a parameter Z, which depends on C, i.e., $Z = g(C)$, and if we have some semantics (critical points) for Z, then we can determine critical points for C by applying the inverse of the dependency g to the critical points. In our example, Z represents the heat transfer rate and C represents Reynolds number. If we are able to determine some critical points for the heat exchange rate, then we can transfer them onto Reynolds number.

4. THE RELATIONSHIP PROBLEM

In this section we concentrate on the forms of the relationships that describe critical hypersurfaces. We assume now that all the relevant physical parameters are known, the question is how to combine them into a formula that defines a critical hypersurface in the cross-product of the parameters.

The interpretation of critical hypersurfaces is such that they describe boundaries between two regions of qualitatively different behaviors of a physical system. In other words, the physical system behaves similarly within the region, even though the quantitative parameters characterizing the system take on different values for different points within the region.

Similar behaviors of physical systems are the subject of the similarity (or similitude) theory (eg., [Birkhoff, 1960], [Drobot, 1953]). The similarity theory gives rules for combining physical parameters into monomials called similarity numbers (similarity modules, dimensionless numbers). A physical system is characterized by one or more similarity numbers. Two states of a system are called similar if all the similarity numbers for the two states are equal. The Reynolds number referenced in previous sections is one example of such a similarity number. It characterizes the process of flow of liquids. A set of states for which all the similarity numbers are equal is a hypersurface. The cross-product of physical parameters describing a particular process is subdivided by the relation of similarity into classes (hypersurfaces). We are concerned only with some special of the classes, the ones that we call critical hypersurfaces.

In this paper we present a method for determining similarity numbers from only the knowledge of physical parameters uniquely characterizing the process, and their dimensions. The rules for doing this are given by the theory of dimensional analysis ([Birkhoff, 1960], [Drobot, 1953], [Whitney, 1968]).

In order to figure out the forms of similarity numbers one needs to analyze dimensions of all physical parameters involved. Dimensions are expressed in terms of units of measure and exponents, eg., $5kg^1m^2s^{-2}$. A number of physical parameters from the set of relevant parameters for the particular process are chosen as so called "dimensional base". A dimensional base can consist of at most as many parameters as many units of measure are used (mostly three). A set of parameters can constitute a dimensional base if the determinant of the exponents' matrix for these parameters is not null. Usually several subsets of the parameters involved satisfy this condition. After selecting a base we take each parameter and combine it with the base by multiplying or dividing it by elements of the base raised to a real-number-power. Those powers are selected in such a way that the resulting monomial is dimensionless. This procedure is repeated for each parameter not in the dimensional base, which means that we can generate as many similarity numbers as many parameters remain in the set after removing from it the ones that constitute the base.

In our example the relevant parameters are expressed in terms of units of mass (kg), length (m), and time (s) as: $v = x_v kg^0m^1s^{-1}$, $\eta = x_\eta kg^1m^{-1}s^{-1}$, $\rho = x_\rho kg^1m^{-3}s^0$, $D = x_D kg^0m^1s^0$. The numbers x_v, x_η, x_ρ, and x_D represent numerical scales of the particular parameters.

Suppose ρ, η, and D have been selected as dimensional base. It can be done because the determinant of the following matrix of exponents is not null:

$$\begin{matrix} 1 & -3 & 0 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \end{matrix}$$

The remaining parameter, v, can be combined with the base into a dimensionless monomial in the following way:

$$v \cdot \rho^{1} . D^{1} . \eta^{-1}.$$

Note that this is the only possible combination of the exponents, given a dimensional base. This is how we derive the functional formula that represents Reynolds number. Selecting a different dimensional base would result in a little different formula. Fortunately, for the purpose of similarity it does not really matter which base has been chosen. The states that are similar in one dimensional base remain similar in any other base that we choose. More formally it is expressed as: similarity of the states of a physical process is invariant with respect to the choice of dimensional base.

5. THE COMPLETE RELEVANCE PROBLEM

The last main problem with establishing a set of critical hypersurfaces is how to determine all relevant physical parameters characterizing a given physical phenomenon. This seems to be the most difficult problem. This problem is beyond the scope of qualitative simulations. But it is closely related to qualitative reasoning as it is one of the steps that lead to establishing operating regions for qualitative processes. This problem has been intensively investigated in modelling and simulation, in systems science, and in the area of artificial intelligence as well.

In any approach to this problem one must accept some kind of a "closed world assumption". In this paper we take a goal-oriented approach in which the closed world assumption is closely related to the goal of the process under consideration. This is manifested in the way of selecting the relevant parameters: only those parameters are selected which can uniquely determine the value of the parameter that characterizes behavior of the system. Mathematically, we are looking for a functional dependency between a parameter explicitly chosen as a characteristic of the behavior of the system and a set of the arguments of this function. The set of parameters that satisfy the requirement of functionality will be complete set of relevant parameters.

Essentially, there are three main ways of determining what are the relevant parameters for a given physical process. One of them is the existing knowledge of the process being modelled. In many cases this knowledge is available, all the relevant parameters can be listed by experts in the given field, the only problem is the qualitative analysis of the possible behaviors of the process.

In more difficult cases the list of the relevant parameters is not readily available, or at least the experts in the given field cannot come to a cosensus on this. The approach in such cases consists of listing all the suspected hypothetical parameters as the candidates, collecting some experimental evidence, and selecting those for which the supportive evidence is the strongest. In such a situation usually statistical methods are applied (factor analysis).

The third class of problems is such in which there is no knowledge about what the suspects might be. It calls for the help of inductive methods. This problem falls into the machine learning category. The BACON learning system [Langley, 1981] postulates some new relevant parameters, arguments of physical/chemical laws. ABACUS [Falkenhainer and Michalski, 1986] subdivides the domain of a functional dependency into qualitatively different subsets. The COPER system [Kokar, 1986a, 1986b] analyzes a set of physical parameters from the point of view of completeness; it utilizes observational data for this purpose. COPER can efficiently determine whether all the relevant parameters have been taken into account in the observations. If incompleteness is found it generates a hypothesis on which parameter is missing. The hypothesis is communicated to the user as a dimensional formula in terms of which the missing parameter is expressed. The user then should assign an interpretation to this description. For instance, if the dimensional formula $[m/s^2]$ is generated, the user is able to recognize that the missing parameter is 'acceleration'. It is perhaps worthwhile to notice that COPER can draw all these inferences (i.e., detect incompleteness of the set of parameters, and generate dimensional formulas of the missing parameters) even when the missing parameter was not varied in the observational data.

6. A FRAMEWORK FOR ESTABLISHING CRITICAL HYPERSURFACES

In this section we summarize the results of the previous three sections by describing several steps in which critical hypersurfaces can be established.

1. Determine the characteristic parameter (goal parameter, output parameter, dependent parameter) of the process which is to be modelled.
2. Determine what are the parameters known to be relevant to this particular characteristic parameter. Use some expert knowledge to this aim.
3. Collect some experimental data - measurements of the characteristic parameter for a number of combinations of values of the relevant parameters.
4. Analyze completeness of the set of relevant parameters. If the completeness condition is fulfilled then analyze redundancy of some of the parameters.

5. If the set of parameters is not complete then use one of the methods for generating descriptions of relevant parameters.
6. Using methods of dimensional analysis derive a set of similarity numbers as functions of the physical parameters.
7. Using one of the available semantics (e.g., signs of derivatives) determine critical values of the derived functions.
8. The expressions equating the monomials representing the similarity numbers with the critical values constitute descriptions of the critical hypersurfaces.

7. CONCLUSIONS

One of the most important problems in qualitative reasoning is how to establish an appropriate quantity space to conduct the simulations in, or in other words, how to establish operating regions for qualitative processes. One possible way is to apply some semantics to quantitative parameters characterizing the physical system under consideration, resulting in a set of critical points for every physical parameter involved. A set of critical points is applicable for some limited situations only, namely when some relevant parameters remain constant. The objective of this paper is to extend this approach to a wider domain of situations, such that would account for variability of all relevant parameters. The generalization of the methodology is achieved through introduction of the notion of critical hypersurfaces in place of critical points. These critical hypersurfaces delimit operating regions of qualitative processes. The paper presents a relatively complete methodology for finding critical hypersurfaces. The methodology is based on the theory of dimensional analysis. A part of this methodology is implemented in the system for discovery of physical parameters, called COPER.

REFERENCES

[Birkhoff, 1960] Birkhoff, G. *Hydrodynamics. A study in logic, fact and similitude*. Princeton University Press, Princeton, 1960.

[deKleer and Brown, 1982] deKleer, J., Brown, S. Foundations of Envisioning. *Proceedings of the National Conference on Artificial Intelligence AAAI-82*, 209-212, 1982.

[Drobot, 1953] Drobot, S. On the foundations of dimensional analysis. *Studia Mathematica, 14*, 84 - 89, 1953.

[Falkenhainer and Michalski, 1986] Falkenhainer, B. and Michalski, R., S. Integrating quantitative and qualitative discovery: The ABACUS system. *Machine Learning, 4*, 1986.

[Forbus, 1984a] Forbus, K., D. Qualitative Process Theory. *Artificial Intelligence*, 24, 85-168, 1984.

[Forbus, 1984b] Forbus, K., D. Qualitative Process Theory. *Technical Report*, 789, MIT Artificial Intelligence Laboratory, 1984.

[Forbus, 1986] Forbus, K., D. Interpreting measurements of physical systems. *Proceedings of AAAI-86, Fifth National Conference on Artificial Intelligence*, Philadelphia, 113-117, 1986.

[Kokar, 1986a] Kokar, M., M. Determining Arguments of Invariant Functional Descriptions. *Machine Learning*, 1, 1986.

[Kokar, 1986b] Kokar, M., M. Discovering functional formulas through changing representation base. *Proceedings of the Fifth National Conference on Artificial Intelligence*, Philadelphia, PA, 1986.

[Kuipers, 1985a] Kuipers, B., J. Qualitative Simulation of Mechanisms. *MIT Laboratory for Computer Science, TM-274*, Cambridge, MA, 1985.

[Kuipers, 1985b] Kuipers, B., J. The Limits of Qualitative Simulation. *Proceedings of the Ninth Joint Conference on Artificial Intelligence*. Los Angeles: 128-136, 1985.

[Langley, P. 1981] Langley, P. Data Driven Discovery of Physical Laws *Cognitive Science*, 5, 31-54, 1981.

[Monin and Yaglom, 1973] Monin, A., S., and Yaglom, A., M. *Statistical Fluid Mechanics: Mechanincs of Turbulence*. The MIT Press, Cambridge, MA, and London, England, 1973.

[Raiman, 1986] Raiman, O. Order of Magnitude Reasoning. *Proceedings of the National Conference on Artificial Intelligence*, Philadelphia, PA, 100-104, 1986.

[Whitney, 1968] Whitney, H. The Mathematics of Physical Quantities, part I and II. *American Mathematical Monthly*, pp. 115-138 and 227-256, 1968.

[Young, 1964] Young, H., D. *Fundamentals of Mechanics and Heat*. Second Edition, McGraw-Hill Book Company, 1964.