Using a Domain Theory to Guide Automated Modeling of Complex Physical Phenomena

Sui-ky Ringo Ling'*
National Supercomputing Center for Energy and the Environment
4505 Maryland Parkway, Box 454028
Las Vegas, Nevada 89154-4028, U.S.A.
email: ling@nye.nscee.edu

Abstract

Choosing appropriate models is crucial in analyzing complex physical phenomena, especially when supercomputing resources and complex partial differential equations are involved. This paper presents an approach to formulating mathematical models guided by the structure of a domain theory and the gross behavior of a physical problem. The approach is motivated by the observation that many physical domains, though complex and computationally expensive to analyze, have strong domain theories based on a few fundamental conservation laws and well-defined physical processes. Furthermore, modeling decisions have to be guided by the behavior specific to a physical problem that the system is trying to model. By exploiting a domain theory and using problem specific behavior, the approach offers an uniform and efficient way of formulating models of various complexity, ranging from algebraic, ordinary to partial differential equations. The approach has been implemented in a computer program, MSG, and tested in the heat transfer domain.

1 Introduction

Modeling is an important first step in scientific computation and engineering analysis. To understand a physical phenomenon requires a representation of the phenomenon in the form of a mathematical model, carrying out simulations of the model, and interpreting the data from the simulations. The form of the model has a dramatic impact on the cost and the accuracy in understanding the phenomenon. The impact is even more significant in analyzing complex physical phenomena that require high performance and supercomputing resources. A few crucial approximations can transform complex partial differential equations to ordinary differential equations or algebraic equations, with the differences in resource requirements between supercomputers and workstations.

This paper describes an approach to formulating models. The models include not only algebraic and ordinary differential equations for modeling lumped phenomena, but also *partial* differential equations for modeling distributed phenomena, i.e. values of physical phenomena vary with time and spatial dimensions. This approach is based on two principles:

- the use of the structure of a domain theory
  - to focus the approximations on choosing the types and the number of basic physical processes, and
  - to sequence the modeling process around the structure of the conservation laws.
- the use of the gross behavior of a physical phenomenon, obtained through simple approximate models, to guide its modeling decisions.

The approach is targeted to domains where strong domain theories exist and simple approximate models for estimating gross behavior are available. It is motivated by the observation that many physical domains have strong and well-defined domain theories from which models are derived. These domain theories are based on several fundamental conservation laws and well-defined physical processes that are governed by domain-specific laws. Examples of these domains are heat transfer, fluid mechanics and structural analysis where the laws of conservation of mass, of energy, and of momentum are the fundamental laws and the domain-specific laws relate the basic processes of mass, energy and momentum to the properties of materials in their respective domains. By organizing the modeling process around the structure of a domain theory, we can provide an uniform approach to formulating models of various complexity.

The approach is also motivated by the need of incorporating problem specific behavior of physical phenomena into modeling decisions. A model reflects the behavior

"Most of the work was done at Rutgers UniverBity and partially supported by NSF under Grant Number DMC-8610507, by the Rutgers Center for Computer Aids to Industrial Productivity, and by ARPA under Grant N00014-81-K-0394 and by ARPA and NASA under ARPA-funded NASA grant NAG2-645. I would like to thank Lou Steinberg for his guidance and support, and Saul Amarel, Andrew Gelsey, Yogesh Jaluria, Elaine Kant, Bob Vichnevetsky and members of the CAP project for their comments. Also thanks to Bob Boehm, Richard Marciano and three anonymous reviewers for their feedback to the earlier versions of this paper.

of a physical problem at hand and the decisions behind the models have to be made within the context of the behavior. But how do we obtain the problem specific behavior to guide the modeling decisions? Simple and approximate models exist in a domain that can be used to obtain the gross behavior of a problem. These models only give the estimates of the gross behavior of a problem, e.g., the maximum temperature drop instead of the temperature distribution in a solid. However, these models are simple to compute, and the gross behavior is often adequate for making the modeling decisions. Using the gross behavior of a problem, obtained from these approximate models, allows a system to quickly focus on a likely candidate model for the problem at hand.

The approach has been implemented in a computer program, MSG (Model Selection and Generation), which formulates mathematical models in the heat transfer domain. The models in this domain involve algebraic, ordinary and partial differential equations, for modeling lumped and distributed heat transfer phenomena.

The paper is structured as follows. Section 2 introduces the domain theory of heat transfer and its implications for organizing the modeling process. Section 3 presents the gross behavior of heat transfer and the methods for obtaining them. Section 4 describes the program MSG implementing this approach and an example problem. Section 5 describes several features of this approach. Section 6 compares this approach with related work, followed by the summary in Section 7.

2 Modeling Process Structured by Domain Theory

Many complex physical domains have well-developed domain theories that are based on several fundamental conservation laws and a set of physical processes defined by domain specific laws. The basic physical processes are the underlying concepts and entities. They are related to other physical properties in the domain through the domain specific laws. The conservation laws govern the interactions of these basic processes, and dictate how these processes are composed into a model. Heat transfer is one of these domains. In the domain of heat transfer, the fundamental law is the conservation of energy and the three basic processes are conduction, convection and radiation heat transfer, defined by the Fourier's law of conduction, the Newton's law of convection and the Stefan-Boltzmann law of radiation. When a temperature gradient is imposed on an object from its environment, instances of the basic heat transfer processes occur and their interaction is governed by the law of conservation of energy. This law says that the net change of energy stored within any bounded region of space, i.e., a control region[1], is equal to the net heat flow into the region plus any internal heat generation within the region. All mathematical models, regardless of their complexity, are based on this conservation law.

One way to view this conservation law is in the form of

$$\sum_{i=1}^{n} \dot{Q}_i = \dot{E}$$

where *Q,s* represent instances of basic heat transfer processes acting on surfaces of the region or internal heat generation processes within the region, and *E* represents internal energy change or energy storage within that region. Regardless of the complexity of heat transfer models and the various forms in which the models appear, all of them can be traced to this form.

This general and yet simple form has two implications for the modeling process. The first implication is for the approximation decisions. Since a model is a representation of this conservation law and the physical processes instantiated for a region in space, the choice of these processes will affect the final form of the model. Various types and number of processes can be chosen. A model representing a few physical processes is less complex. Similarly, a model representing lumped processes, i.e. processes which are assumed constant with respect to time and space, is less complex than a model representing differential processes, i.e. processes varying with respect to time and space. In other words, the complexity of a model depends on the choices of processes, and all the approximations in this domain are focused on *choosing the types and the number of processes.* A corollary to the focus of choosing physical processes is that the modeling process is *compositional* at the level of physical processes. The conservation law provides a template to put an arbitrary number of processes together once they are chosen. Processes are put in the right or left hand sides of the equation depending on their types.

The form of a model depends on the choices of physical processes. However, these processes depend on the type of a control region where the conservation law is instantiated and the type of query that a model is supposed to answer. This dependency constrains how modeling choices should be made. In other words, the modeling decisions and approximations are not totally independent of one another. Some of the decisions are constrained by earlier decisions. This dependency suggests a way to organize the modeling process in the following sequence.

1. First, decide on the control regions of an object.
2. For each control region, identify the potential active heat transfer and energy storage processes, and make approximations on them.
3. Finally, turn the processes into symbols in equations, simplify the equations, and provide relevant initial and boundary equations to make the models well-formed.

3 Modeling Decisions Guided by Gross Behavior

The domain theory provides a set of modeling decisions, and a way of organizing those decisions. How are these modeling decisions made? In particular, how can a system choose the control regions and the physical processes

[1] Also called a control volume

such that the final model is a candidate representing a physical problem with certain accuracy?

To choose a candidate model, the system must be guided by the behavior specific to the problem that the system is trying to model. For example, if the temperature distribution of a problem is relatively constant in the *x* dimension, the system will choose heat transfer processes lumped in the *x* dimension, resulting in a model with the temperature function independent of x. The question is: how does a system know the temperature is constant in x? To know the behavior requires a model, but a model is what the system is looking for in the first place. The situation becomes a vicious circle, where intelligent modeling requires knowing a lot about a physical problem at hand. To know the problem requires a model, which in turn begs the question of where the model comes from.

One way to avoid this vicious circle is to use simple models to estimate the behavior of a problem. These models are simple to compute. They are domain specific. They are mostly expressed in terms of ratios and dimensionless numbers. They are approximate in the sense that they only predict gross behavior of a problem. However, the gross behavior is often adequate for guiding modeling decisions. These simple and approximate models exist in the heat transfer domain and allow this approach to incorporate problem-specific behavior into its modeling decision. For example, a control region can be lumped in the *x* dimension in an object if the temperature distribution is constant in that dimension. Instead of finding the temperature distribution T(x), the approach estimates the maximum temperature drop $_{\Delta x}T$ in the object in the *x* dimension. $_{\Delta x}T$ is the gross behavior of *T(x)* because it only predicts the maximum drop. It does not give the temperature value at every point in x, as *T(x)* does. If that maximum temperature drop is smaller than a threshold, then the approach infers that the temperature function is constant in the *x* dimension[2].

Instead of using a complex model to estimate T(x), the approach can use some simple models to estimate $_{\Delta x}T$. For example, $_{\Delta x}T$ in a single-component solid can be estimated from the Biot number [incropera and DeWitt, 1990], *hL/k,* where *h,k,L* are respectively the convection coefficient, the conductivity and the thickness of an object in *x*. This ratio is simple to calculate and the values of *h,* L, *k* can be obtained from the input description of a problem. The rule for choosing a lumped control region in the *x* direction for a single component object then becomes:

> If the biot number *hL/k* is less than the threshold of 0.1, choose a lumped control region in x.

For a multi-components object, the maximum temperature drop in each component can be estimated by its thermal resistance over the total thermal resistance of the object, and the control region for each component can be chosen using similar rules. Other simple models for estimating the relative magnitude of heat transfer processes and the rules for choosing them are obtained from domain experts, and the textbooks [incropera and DeWitt, 1990; Jaluria and Torrance, 1986].

Besides using simple and approximate models to estimate the gross behavior of a problem, the approach also relies on a set of thresholds to determine the conditions under which approximations can be made, e.g. the value of the Biot number less than 0.1. These thresholds are established based on the empirical and experimental studies of previous heat transfer problems. These thresholds are calibrated against accuracy requirements, and they establish the conditions for making approximations such that the output result of the model would meet certain accuracy. For example, the Heisler charts [Heisler, 1947] present the relations between temperature distributions and the Biot number for various regular shape objects. From these charts, the maximum error for assuming uniform temperature distributions when the Biot number is less than 0.1 is less than 5%. Other studies have also been done to calibrate other thresholds for choosing various types and number of heat transfer processes [Ling, 1994].

4 MSG: the Program

This section presents MSG, the computer system that has been implemented based on this approach for modeling heat transfer. MSG is focused on choosing mathematical models for typical heat transfer engineering analysis where accurate results with a typical error of less than 10% would be expected. The mathematical models are for modeling lumped and distributed heat transfer behavior in regular shape objects of single and multiple components.

The input to MSG is

- A description of an object in terms of its geometry, components, connections, material properties, initial and boundary conditions.
- A query that a model is required to answer, and its spatial and temporal attributes, e.g., *T(x,t)* would ask the temperature varied in the x spatial dimension and time.

The output is:

- A mathematical model expressed in terms of a set of equations, and
- A set of assumptions which the system makes for that model.

The top level description of MSG follows the sequence of choosing control regions, then heat transfer processes, and finally transforming processes into mathematical forms, as shown in Figure 1. The approximate methods for estimating gross behavior and the thresholds calibrated to the accuracy requirement of less than 10% error are encoded within MSG.

4.1 An example: 1-D Transient Model

The example involves heat transfer analysis of a composite made of two plates of brick and gypsum, having the same width and height dimensions, 0.3 m on all sides, and the same thickness, 0.02 m, shown in Figure 2. Plate

[2] subject to some assumptions about the temperature profile.

1. **Preprocessing:**
   - identify common surfaces between components and environments.
   - compute the thermal resistance R_i of each component C_i in an object in each spatial dimension.
   - compute the total thermal resistance R_{total} for the object in each dimension by summing R_i.
2. **Choices of Control Regions:**
   - if the object has one component, use the Biot number $\leq$ the lumping threshold to choose the control region along each spatial dimension
   - if the object has multiple components, choose the control region CV_i for each component i by comparing the ratio $R_i/R_{total} \leq$ the lumping threshold along each dimension,
3. **Choices of Physical Processes:**
   - instantiates processes $\{\dot{E}, \dot{Q}_i..\}$ for each CV_i
   - eliminate $\dot{E}$ if the response time ratio or $\Delta T/T_{ref} \leq$ the transient threshold
   - eliminate $\dot{Q}_x$s if $\dot{Q}_x \ll \dot{Q}_y$ or $\dot{Q}_x = 0.0$ or aspect ratio $\ll$ the spatial threshold or ...
   - approximate linear properties P if $\Delta P/P_{average} \leq$ the linear threshold
4. **Mathematical Transformations:**
   - For each CV_i, transforms $\{\dot{E}, \dot{Q}_i..\}$ into $\dot{E} = \dot{Q}_i + ...$
   - instantiates an initial condition equation for each $\dot{E}$,
   - instantiates boundary condition equations for $\dot{Q}_{diffusion}$ in each CV_i,
   - expand $\dot{E}$ or $\dot{Q}$ into expressions containing the temperature distribution T if the query is T,
   - instantiates temperature continuity $T_j = T_i$ along the common surfaces between the components, i, j.

Figure 1: Top level Description of the MSG Algorithm

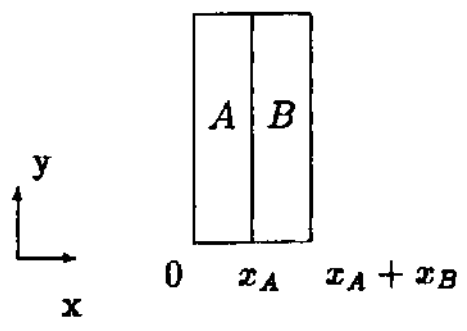

Figure 2: Two Plates Object

A has conductivity of 0.72 $W/m.K$, specific-heat of 835 $J/kg.K$, and density of 1920 kg/m^3. Plate B has conductivity of 0.22 $W/m.K$, specific-heat of 1085 $J/kg.K$, and density of 1680 kg/m^3. Their initial temperature is 273.15 K. They are exposed to an environment of temperature 473.15 K with convection heat coefficient of 50 $W/m^2.K$. The query involves the temperature of both plates with respect to all three spatial dimensions and time, i.e., $T(x, y, z, t)$. The following is a trace of major steps that generates a one-dimensional transient heat transfer model for the composite plates.

1. **Preprocessing:**
   - MSG first identifies a common surface between Plates A and B.
   - Thermal Resistances for Plate A, Plate B are 0.091 $K.m^2/W$, and 0.028 $K.m^2/W$. Thermal resistance for each environment is 0.02 $K.m^2/W$.
   - Total resistance is 0.159 $K.m^2/W$.
2. **Choices of Control Regions:**
   - Since the object is a two-plates object, MSG uses thermal resistance ratios to determine the type of control region in each plate. R_A/R_{total} is equal 0.57 and R_B/R_{total} is equal to 0.18. Both of them are greater than the lumping threshold of 0.1. Differential control regions, CV_A, CV_B, are chosen for both plates.
3. **Choices of Physical Processes:**
   - For the differential control region CV_A, MSG instantiates a set of processes $\{\dot{E}, \dot{Q}_x, ...\dot{Q}_{x+dx}\}$, which contains the conduction processes $\dot{Q}$ in the x, y, z dimensions and the energy storage process $\dot{E}$.
   - MSG uses the aspect ratios to estimate the magnitudes of conduction processes in the y, z dimensions. Both the aspect ratios are 0.07, smaller than the spatial threshold of 0.1, MSG removes $\dot{Q}_y, \dot{Q}_{y+dy}, \dot{Q}_z, \dot{Q}_{z+dz}$ from the set and reduces the final set of processes to $\{\dot{E}, \dot{Q}_x, \dot{Q}_{x+dx}\}$.
   - MSG retains $\dot{E}$ since $\Delta T/T_{ref}$ = (473.15-273.15)/473.15 is larger than the transient threshold.
   - Linear properties are chosen for their input values which are constant.
   - Similar choices are made with respect to the control region CV_B.
4. **Mathematical Transformations:**
   - MSG maps $\{\dot{E}, \dot{Q}_x, \dot{Q}_{x+dx}\}$ into a mathematical equation $\dot{Q}_x - \dot{Q}_{x+dx} = \dot{E}$ for CV_A.
   - Since the query involves temperature, MSG expands the equation $\dot{Q}_x - \dot{Q}_{x+dx} = \dot{E}$ into the governing equation shown in Figure 3.
   - MSG instantiates an initial condition equation, $T_A(x, 0) = 273.15$ for CV_A.
   - MSG instantiates two boundary condition equations at $x = 0, x = xA$ for CV_A.

Governing Equation:

$k_A \frac{\partial^2 T_A}{\partial x^2} = \rho_A C_A \frac{\partial T_A}{\partial t}$

$k_B \frac{\partial^2 T_B}{\partial x^2} = \rho_B C_B \frac{\partial T_B}{\partial t}$

Initial conditions:

$T_A(x,0) = 273.15$

$T_B(x,0) = 273.15$

Boundary conditions:

$x = 0 : h_{e1}(T_{e1} - T_A) = -k_A \frac{dT_A}{dx}$

$x = xA : -k_B \frac{dT_B}{dx} = -k_A \frac{dT_A}{dx}$

$x = xA : T_A = T_B$

$x = xA + xB : -h_{e2}(T_{e2} - T_B) = -k_B \frac{dT_B}{dx}$

Figure 3: 1-D Transient Model

```
    /   2          \
    |  d           |              /  d          \
k_a |----- temp_a(x)| = rho_a cv_a |---- temp_a(t)|
    |   2          |              \ dt          /
    \ dx           /

          init_eqn_a := 273.15

      /  d          \
- k_a |---- temp_a(x)| = - h_e1 (temp_a - temp_e1)
      \ dx          /
```

Figure 4: Part of MSG's Actual Output

- Because of a common surface between the plate A and B, MSG instantiates $T_A = T_B$ at $x = xA$.
- Similar steps are repeated for CV_B except only one boundary condition is instantiated at $x = xA + xB$. The boundary condition at $x = xA$ has already been instantiated in the previous step for CV_A.

The final model, shown in Figure 3, has two governing equations (PDE) describing the temperature distribution in the plates in the x direction, four boundary equations for the surfaces of the plates and two initial conditions. Part of the actual output from MSG is shown in Figure 4.

4.2 Implementation and Testing

MSG has been implemented in Common Lisp and CLOS and interfaced with MAPLE, a mathematical manipulation system. The system is able to formulate heat transfer models for single and multiple components of rectangular shape geometries, including composite plates, concentric shell and L-shaped blocks. The system has been tested with 27 problems taken from two heat transfer textbooks [incropera and DeWitt, 1990; Jaluria and Torrance, 1986] and 16 hand-crafted problems. The textbook problems come from heat transfer analysis in walls, in ovens, in cooling of electronic circuits and in materials bonding. The list of examples are listed in [Ling, 1994]. The examples cover three kinds of queries. They are: energy storage, heat flux and temperature distribution varying with respect to spatial and time dimensions. The models range from coupled nonlinear three dimensional partial differential equation models to algebraic equations.

5 Discussion

This section describes several issues and consequences of using a domain theory and problem specific behavior to guide modeling decisions.

5.1 Models Consistent with Domain Theory

Models are not just symbols and equations to produce numerical results. They represent physical realities of a domain. Engineers and scientists use models to obtain numerical data and use the data to guide their activities in physical world, such as planning physical experiments and building artifacts. To be able to use the numerical data and to have confidence in the data, they must be able to interpret their models as representing valid physical realities according to the theory of a domain. While they are interested in the accuracy of a model, they would also like the model to be consistent with the basic domain theory. Given a choice, they do not want models which just curve-fit the data without relationships to the underlying physics of a domain. The issue is especially important in formulating models in a domain with a well-established theory, where models are seen as instances of the theory applied to physical phenomena. However, some approximations, such as the order of magnitude approximation, can lead to approximate models that are inconsistent with a domain theory, if they are applied to mathematical equations syntactically. This is illustrated in the following two examples.

In the order of magnitude approximation, a mathematical expression $A - B$ is simplified to A when $A >> B$, i.e., the numerical value of A is much greater than the numerical value of B where the notion of much greater is defined by some numerical measure.

First Example The first example involves a model representing the radiant heat exchange between a surface and its surroundings, shown in the left side of Figure 5. The model is $\sigma\epsilon(T_1^4 - T_2^4)$ where σ represents the Stefan-Boltzmann constant, ϵ represents emissivity of a surface, T_1 the temperature at surface and T_2, the temperature at its environment. Assuming that $T_1^4 \gg T_2^4$, then the order of magnitude approximation can simplify the original model to a simplified model $\sigma\epsilon T_1^4$, representing a single radiant heat flux emitted from the surface at temperature T_1. The simplified model is consistent with the heat transfer theory.

Second Example The second example involves a model representing the convection heat exchange between a surface and its surroundings, shown in the right side of Figure 5. The model is $hA(T_1 - T_2)$

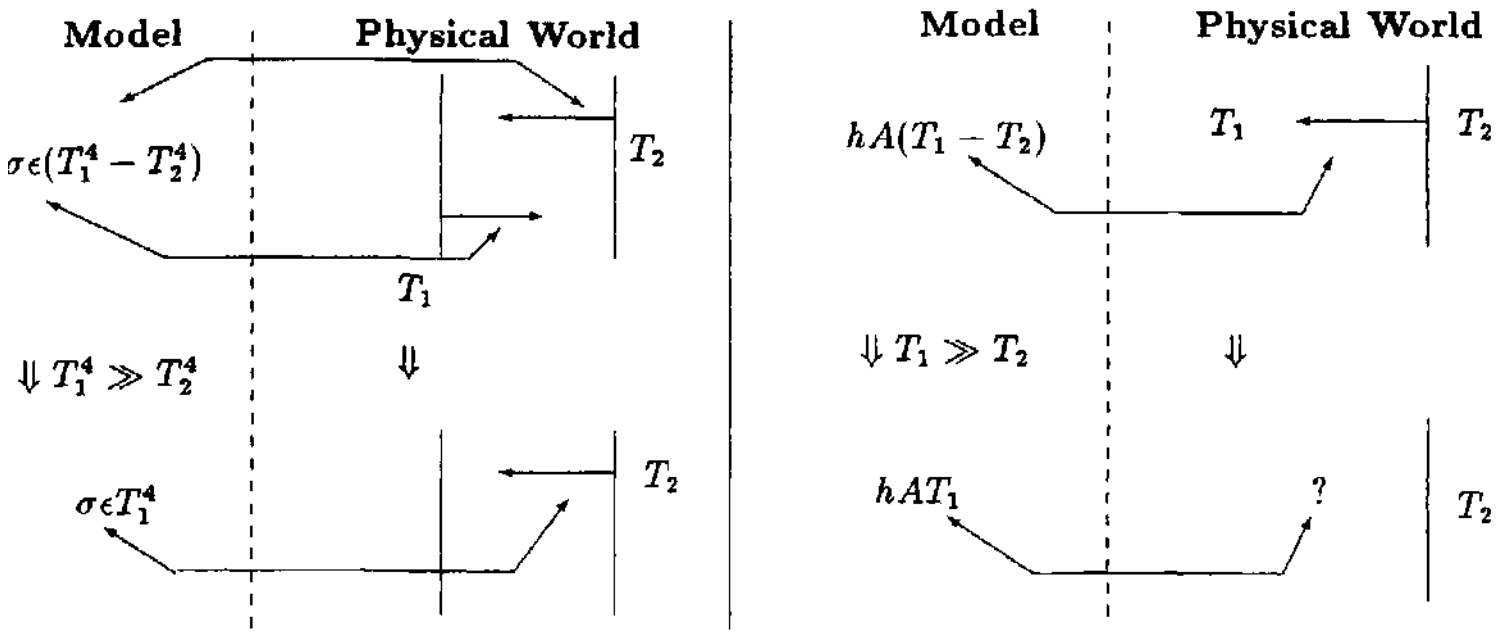

Figure 5: Two examples of the Order of Magnitude Approximations

where *h* represents the convection heat coefficient, *A* represents the area of a surface, T_1 the temperature at surface and T_2, the temperature at its environment. Assuming that $T_1 > T_2$, then the order of magnitude approximation can simplify the model to a simplified model hAT_1. However, that simplified model is *not* consistent with the heat transfer theory as the expression does not represent any valid heat transfer process. The original model represents a convection heat flux and eliminating T_2 from the original model destroys its definition, i.e., convection heat transfer depends on temperature difference instead of temperature only.

The two heat transfer examples illustrate the case where the syntactic use of the order of approximation can lead to approximate models inconsistent with a domain theory. To ensure models consistent with a domain theory, the order of magnitude approximation cannot be applied to any arbitrary mathematical expression. It can only be applied to a set of physical processes in a way that represents the action of eliminating one or more physical processes. By focusing the approximations on choosing the types and the number of physical processes in a domain, MSG avoids generating models inconsistent with the heat transfer theory. The constraint of ensuring consistent models is already built in MSG.

5.2 Constraining Modeling Process

MSG demonstrates that the process of modeling is compositional at the level of choosing physical processes. Furthermore, it demonstrates that a domain theory can constrain how modeling decisions are made. Some modeling decisions can only be made after other decisions have been made. For example, formulating boundary conditions has to be made after choosing physical processes, and choosing physical processes has to be made after choosing control regions. A consequence of constraining modeling decisions is that modeling choices are also constrained by the earlier decisions, leading to a reduced search space for MSG. If a totally lumped control region is chosen, MSG does not need to choose heat transfer processes since a totally lumped model will only involve an algebraic or ordinary differential equation, either of which can be efficiently solved with many processes. Furthermore, MSG does not need to choose boundary equations for the totally lumped model. The reduced search space from constraining modeling decisions, and the use of simple approximate model to guide modeling decisions contribute to the efficiency of MSG. MSG currently takes less than 4 seconds on a SparcStation 2 to formulate a model involving coupled partial differential equations.

MSG demonstrates an *uniform* approach to formulating models of various complexity in a physical domain, from simple algebraic to complex partial differential equations models. While partial differential equations are complex to analyze and difficult to comprehend, they are formulated in the same way as other simple models. They do not need a different modeling approach. The complexity of a model lies not in the way it is formulated, but in the choices made to the physical processes.

5.3 Limitations

This approach works because of a strong domain theory, where conservation laws exist and the domain specific laws are well defined, and approximate numerical methods and thresholds calibrated against accuracy are available. These features are present in other domains as well, such as fluid mechanics [Panton, 1984], based on textbook analysis and discussions with domain experts. However, the requirement of a strong domain theory and the availability of approximate methods and thresholds limit this approach to well-established physical domains.

6 Related Work

Yip [Yip, 1993] uses the asymptotic order of magnitude approximation to simplify partial differential equation models in fluid mechanics. His system accepts a detailed mathematical model and the constraints and values of terms in the model as inputs, and then carries out the asymptotic order of magnitude approximation on the input model. MSG differs from his system by using approximate models to infer the order of magnitude values of physical processes and by carrying out approximations on physical processes to ensure models consistent with a domain theory.

Most of the other modeling work is based on the compositional approach pioneered by Falkenhainer and Forbus [Falkenhainer and Forbus, 1991]. The compositional approach has a library of model fragments representing various phenomena in a domain. Given a query and a physical situation, the approach composes relevant model fragments into a complete model. Other work [iwasaki and Levy, 1994; Nayak *et al.,* 1992] has proposed ways to improve the process of selecting model fragments. MSG differs from the compositional approach by focusing its modeling decisions on physical processes and using a domain theory to structure its modeling process. By organizing its modeling decisions around a domain theory, MSG can formulate models for lumped and distributed phenomena in an uniform and efficient way. The compositional approach is aimed at a very broad class of phenomena, without making any assumption of a domain theory. MSG relies on a strong domain theory, which makes MSG less broadly applicable than the compositional approach, especially for a domain with no well developed theory. The compositional approach can handle qualitative and quantitative ordinary differential equations. MSG is focused on quantitative models involving algebraic, ordinary and partial differential equations. Finally, the compositional approach has paid less attention to the issue of getting the right level of numerical approximation than to getting a consistent and coherent model.

Both Addanki *et al* [Addanki *et* al., 1991], Ellman *et al* [Ellman *et* a/., 1993] and Weld [Weld, 1992] focus on model selection. The work of Addanki *et al* uses graphs to represent a set of models of a physical domain, and domain specific rules to select models that can resolve the conflicts between predictions and observation. Weld's work uses the domain independent technique of inter-model comparative analysis. Ellman *et al* use gradient magnitudes to select models. In all cases, model equations are explicitly input to the system. MSG differs from them by constructing a model from the description of a physical phenomenon.

7 Summary

This paper describes a modeling approach to formulating complex mathematical models involving partial differential equations. The approach is targeted to domains with strong domain theories. The approach organizes its modeling process based on the structure of a domain theory, and uses simple approximate models to obtain problem specific behavior to guide its modeling decisions. By exploiting these ideas, the approach can offer an uniform and efficient way to formulate models consistent with a domain theory. The approach has been implemented in a computer program, MSG, which formulates algebraic, ordinary and partial differential equation models in the domain of heat transfer.

References

[Addanki *et al.,* 1991] S. Addanki, R. Cremonini, and J. S. Penberthy. Graphs of models. *Artificial Intelligence,* 51(1-3):145-177, 1991.

[Ellman *et al.,* 1993] Thomas Ellman, John Keane, and Mark Schwabacher. Intelligent model selection for hill-climbing search in computer-aided design. In *Proceedings of the Eleventh National Conference on Artificial Intelligence,* 1993.

[Falkenhainer and Forbus, 1991] Brian Falkenhainer and Ken Forbus. Compositional modeling: Finding the right model for the job. *Artificial Intelligence,* 51(I-3):95-144, 1991.

[Falkenhainer, 1993] Brian Falkenhainer. Ideal Physical Systems. In *Proceedings of the Eleventh National Conference on Artificial Intelligence,* pages 600-605, Washington, DC, 1993. AAAI.

[Heisler, 1947] M. P. Heisler. Temperature charts for induction and constant temperature heating. *Transaction of ASME,* 69, 1947.

[incropera and DeWitt, 1990] Frank P. Incropera and David P. DeWitt. *Fundamentals of Heat and Mass Transfer.* John Wiley and Sons, 3rd edition, 1990.

[Iwasaki and Levy, 1994] Yumi Iwasaki and Alon. Y. Levy. Automated model selection for simulation. In *Proceedings of the Twelfth National Conference on Artificial Intelligence,* pages 1183-1190. AAAI, 1994.

[Jaluria and Torrance, 1986] YogeBh Jaluria and Kenneth E. Torrance. *Computational Heat Transfer,* Hemisphere Publishing Corporation, New York, 1986.

[Ling, 1994] S. Ringo Ling. *Automated Modeling of Lumped and Distributed Phenomena in Heat Transfer.* PhD thesis, Dept. of Computer Science, Rutgers University, 1994.

[Nayak *et al.,* 1992] Pandurang Nayak, Leo Joskowicz, and Sanjaya Addanki. Automated model selection using context-dependent behaviors. In *Proceedings of Tenth National Conference on Artificial Intelligence,* 1992.

[Panton, 1984] Ronald L. Panton. *Incompressible Flow.* John Wiley and Sons, 1984.

[Weld, 1992] Daniel S. Weld. Reasoning about model accuracy. *Artificial Intelligence,* 56, 1992.

[Yip, 1993] Kenneth M. Yip. Model simplification by asymptotic order of magnitude reasoning. In *Proceedings of the Eleventh National Conference on Artificial Intelligence,* pages 634-640. AAAI, 1993.