

Constructing Process Models Systematically

Heinz A Preisig

Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway
 Heinz.Preisig@chemeng.ntnu.no

Models are the starting point for any computation-based engineering activity and thus errors in the construction - instantiation - coding of models are beyond the most expensive ones. Any improvement of this situation requires a systematic approach that enables automatic generation of coded models.

We discuss on a simple example the systematic construction of process models based on system theoretical concepts that gives us the possibility to guarantee the completeness and structural correctness of the resulting equations. Whilst we explain the basic ideas on a very simple example, the method expands seamless to very large scale systems.

As part of the discussion we shall also explore the link to the behaviour theory as it was defined in the mathematical system theory community. This then provides a clear basis for the mathematical analysis in the abstract signal space whilst having a clear physical interpretation in the thermodynamic and fluid-dynamic state space.

1. Goals and Objectives

Models are essential, they are the core of computer-aided engineering. Having them right is essential and any fault in their definition is expensive, as the model is at the beginning of any of the computational processes. The projects in computer-aided modelling in our group thus focus on defining proper process models, where proper implies structurally sound, differential index 0 or 1, no obsolete equation or variable and no coding and transcript errors. Thus we aim at tools that can construct process models meeting these goals and make it easy and quick to generate and modify process models.

2. Systems View

The term "behaviour" is often used when explaining what a model is: The model mimics the behaviour of the system, for example. This though asks the question on what "behaviour" is, how it is defined. The term is intuitively interpreted as: a model shows the "same" response than the modelled system when excited the same way. So it kind of describes an input/output behaviour: given input, the output of the model is close to the (measured/observed) output of the modelled object, which closely fits the definition of the term model in the literature of philosophy of science, for example Apostel (1960). Mathematical system theory defines a domain called "Behaviour Theory" with Willems being the centre person. A collection of his papers can be found on the webpage (Willems, 2014). Willems defines behaviour by "The behaviour B of a model is constructed as a subset of the universe of conceivable events U , thus $B \subseteq U$."

The definition is illustrated with the ideal gas law : $p V = n R T$. The universe is every combination of values $U := [0, \infty) ^4$; so the model lives in this universe. The behaviour is $B := (p, V, n, T) \in [0, \infty) ^4 \mid p V = n R T$. The behaviour splits the universe of conceivable events into two parts: one that DEFINES the behaviour and the other which defines what is NOT part of the behaviour. If we solve for any of the variables we have a manifold description of the system. Willems defines a signal space in which the model "operates". One is then tempted to talk about inputs and outputs in the signal space without being very specific on what is an input and what an output. Physics on the other hand provides a very clear view on inputs, outputs and states, which we will use to further refine providing a convenient structure for the representation of physical / biological systems. We

view the process we model as one part of the universe with the other being its environment into which it is embedded (Preisig, 2010 and 2014) .

The physical system lives in the frame of time and space. The plant occupies a piece of the universe's space and in order to describe its behaviour we split it into smaller part, usually motivated by arguments being functional, like functional units, or mechanical properties such as different phase. The individual parts exchange extensive quantity as a consequence of a difference in the state. Overall the plant is driven as a function of a difference in the intensive properties of the reservoirs making up the environment of the plant. So flow of extensive quantity is driven by the difference in the effort variable, which are the conjugates to the potentials. As rightly is observed by Willems (1989), flow is possible in either direction, which makes it difficult to use the term input for a system, if one uses the intuitive view that flow into the a space is an input and an outflow is an output. This link does not provide us with the right picture. The change of the system is indeed driven by the conditions at its boundaries, globally the environment. There is though another element affecting the dynamics, namely the system-internal dynamics: transposition of one extensive quantity into another one. This mechanism models phase changes, chemical and biological reactions and conversion of different forms of energy to mention the main ones. The art of generating the model is primarily the definition of the basic breakdown of the space into smaller partitions, which each is described as a physical system having the capacity of storing extensive quantity and exchange such with its neighbours due to the difference in the state of the two neighbours and internal imbalance of its constituents. It is thus the granularity of the decomposition that determines the resolution with which the dynamic behaviour is being described by the model. Defining the volume elements that act as capacities for the conserved extensive quantities as nodes and the connections as arcs which have a defined directionality, one has as a base description of the containment a directed graph. Two elements describe the directed graph, namely the list of nodes and a list of arcs, whereby the latter is a list of tuples with source node and sink node in each tuple.

What is in this containment we describe with tokens (Preisig, 2015). The tokens in dynamic systems give raise for the definition of state being the conserved quantities in each node.

3. The Grand Scheme

Based on these considerations, we define a set of signals: *state* – the conserved extensive quantities, *transport* of extensive quantity, *transposition* of extensive quantity, and *secondary states*, which are state-dependent quantities being used to represent the transport and the transposition. The latter are the least obvious class of signals except that the effort variables are apparently members of this class. Since the effort variables are linked to the state through the equations of state, the latter are part of the relations that come to provide the necessary link between the state and the secondary state. From this point of view it also becomes evident that the definition of the intensive properties is part of the same set of equations and the link between mass and volume plays an important role. This brings geometry into the picture, which is provides the necessary link between mass, shape and volume of the system. The result of these considerations is summarised in Figure 1.

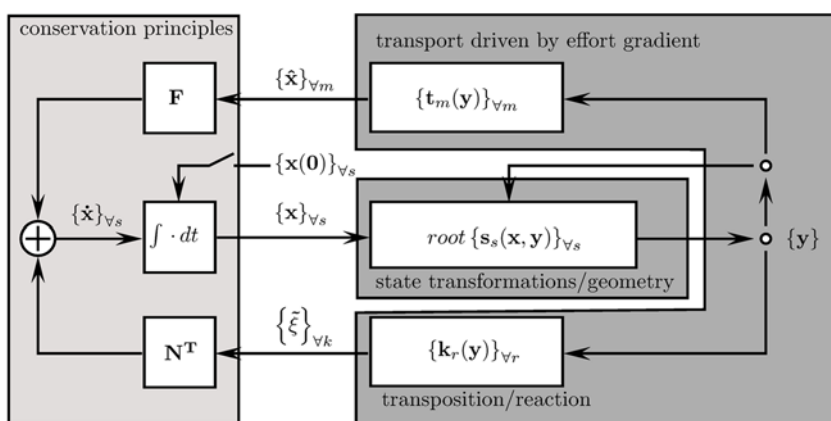


Figure 1: The Grand Scheme – the conservation principles are linear a reflection of the superposition that applies to the conserved quantities in the capacities. The F matrix is the incidence matrix of the directed graph of the network representing the containment. The N matrix is the transposition matrix, for reactions this is the stoichiometric matrix. The U-shaped block on the right

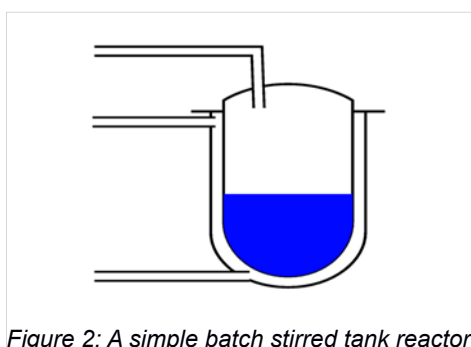
represents the second part, namely the transport modelling the transfer between capacities with the flows of extensive quantities being marked by a hat, and the transposition representation the capacity-internal dynamics marked by a tilde, here the extent of reaction ξ . Finally the centre box represents the link between the secondary states y and the primary state x . Here the material properties dominate, thus the equation of state supplemented by geometry and definition of intensive quantities.

4. Model Definition

The construction of the model, is done in counter-direction of the arrows starting with the state. It is obtained by integrating the differential balances, the change in the state per unit time. The incidence matrix is representing the structure of the network (Preisig, 2010, 2014) and the transposition matrix provides the ratio of the transposition. For reactions that is the stoichiometric matrix, for phase changes it is unity. The "input" to the balances are the transport of extensive quantity and the transposition. Thus the next stage in the definition is two fold, namely the definition of the transport equations and the definition of the transpositions, usually the kinetics. These two set of definitions are implicitly defining the set of secondary states that are required. The secondary states must be the result of a mapping from the fundamental state, the extensive quantities. Whilst this transformations are often not analytically available, they must exist and solvable by numerical methods. It should be noted that this class of equations includes implicit equations. Notably it is the temperature which appears usually in an implicit equation.

5. Illustrative Example

We use a simple example to demonstrate the approach to construct a complete proper process model. The motivation to choose a simple example is to provide an example that most can relate to and that is not resulting in an excessive set of equations.



The plant being modelled is a jacketed ideally stirred tank reactor, a classic. The assumptions for generating the process model are rigorous: (i) only the liquid phase is modelled – the gas phase is neglected (ii) ideally stirred liquid phase. (iii) Tank is initially charged with a load of fluid. (iv) The jacket is ideally mixed and connected to a source and ejecting fluid to a sink. Both not being further specified. (v) No heat and material losses.

We use a graphical approach to generating the model structure (Preisig, 2014). The resulting topology is shown in Figure 3.

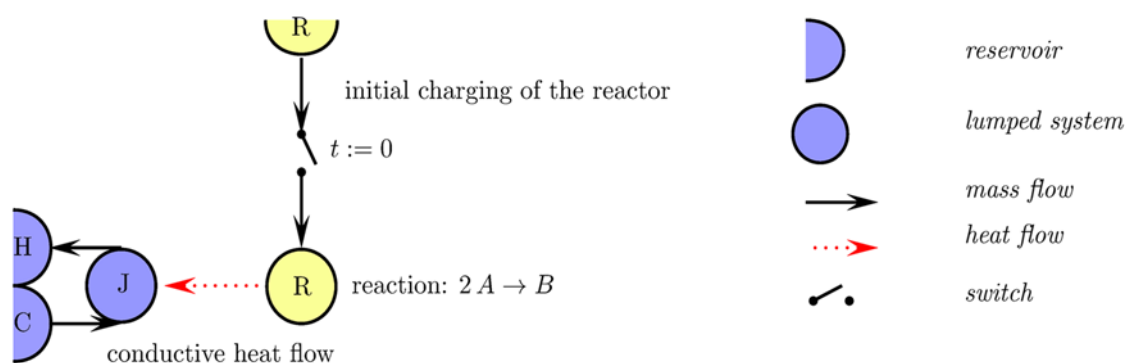


Figure 3: A simple topology of a jacketed batch reactor modelled as an ideally stirred tank reactor for the liquid phase, with no gas phase and a uniform jacket.

In the tank a reaction is going on, namely $2 A \rightarrow B$. The topology is extremely simple, which reflects into the incidence matrix of the graph. For the mass flow attached to the contents, it is really only one flow that is to be considered, namely the charging. The energy balance is affected by the same mass flow and in addition the heat exchange with the jacket.

<i>equations</i>	<i>res</i>	<i>vars</i>	<i>given</i>
integrals			
$\underline{\mathbf{n}}_R := \int_0^t \dot{\underline{\mathbf{n}}}_R dt + \underline{\mathbf{n}}_R^o$	$\underline{\mathbf{n}}_R$	$\dot{\underline{\mathbf{n}}}_R$	$\underline{\mathbf{n}}_R^o$
$H_R := \int_0^t \dot{H}_R dt + H_R^o$	H_R	\dot{H}_R	H_R^o
differential balances			
$\dot{\underline{\mathbf{n}}}_R = \tilde{\underline{\mathbf{n}}}_R$	$\dot{\underline{\mathbf{n}}}_R$	$\tilde{\underline{\mathbf{n}}}_R$	
$\dot{H}_R = -\hat{q}_{J R}$	\dot{H}_R	$\hat{q}_{J R}$	
flows			
$\hat{q}_{J R} := -k_{J R}^q (T_R - T_J)$	$\hat{q}_{J R}$	T_R	$k_{J R}^q, T_J$
production rate			
$\tilde{\underline{\mathbf{n}}}_R := \underline{\underline{\mathbf{N}}}^T \tilde{\underline{\underline{\xi}}}_R$	$\tilde{\underline{\mathbf{n}}}_R$	$\tilde{\underline{\underline{\xi}}}_R$	$\underline{\underline{\mathbf{N}}}^T$
$\tilde{\underline{\xi}}_R := V_R k_R^r g(\underline{\mathbf{c}}_R)$	$\tilde{\underline{\xi}}_R$	$k_R^r, g(\underline{\mathbf{c}}_R)$	V_R
$k_R^r := E^o \exp E_A/(RT_R)$	k_R^r	T_R	E^o, E_A, R
$g(\underline{\mathbf{c}}_R) := ([1, 0] \underline{\mathbf{c}}_R)^2$	$g(\underline{\mathbf{c}}_R)$	$\underline{\mathbf{c}}_R$	
state variable transformations			
$\underline{\mathbf{c}}_R := V_R^{-1} \underline{\mathbf{n}}_R$	$\underline{\mathbf{c}}_R$	$\underline{\mathbf{n}}_R$	V_R
$H_R := \int_{T_R^o}^{T_R} \frac{\partial}{\partial T} \left(\frac{\partial H(T)}{\partial \underline{\mathbf{n}}_R^T} \right)_p \underline{\mathbf{n}}_R dT$	T_R	$H_R, \underline{\mathbf{n}}_R, \frac{\partial}{\partial T} \left(\frac{\partial H(T)}{\partial \underline{\mathbf{n}}_R^T} \right)_p$	T_R^o
$\frac{\partial}{\partial T_R} \left(\frac{\partial H(T)}{\partial \underline{\mathbf{n}}_R^T} \right)_p := \frac{\partial}{\partial T} \left(\frac{\partial H(T^o)}{\partial \underline{\mathbf{n}}_R^T} \right)_p$	$\frac{\partial}{\partial T_R} \left(\frac{\partial H(T)}{\partial \underline{\mathbf{n}}_R^T} \right)_p$	$\frac{\partial}{\partial T} \left(\frac{\partial H(T)}{\partial \underline{\mathbf{n}}_R^T} \right)_p$	$\frac{\partial}{\partial T_R} \left(\frac{\partial H(T^o)}{\partial \underline{\mathbf{n}}_R^T} \right)_p$

Figure 4: Set of equations describing the behaviour of the reactor part, thus not the jacket.

Figure 4 shows the set of equations representing the behaviour of the liquid contents. The table has 3 columns, the first is the equation, the second what variable is computed, the third what is needed to be available and the last what is given. As suggested we start with the integration of the differential balances, shown as the first block of two equations. The variables being computed are the state variables, namely the molar species vector and the enthalpy (as we assume constant pressure in the tank). The initial conditions are given and take care of the initial charging with reactants at time zero. Required are the change of the state with respect to time, thus the differential balances. Those are in the next block. The next block is devoted to the transport. Here it is only one stream, namely the heat flow between jacket and liquid contents. The arrow in the graphical representation of the topology provides the reference co-ordinate for the flow. Thus positive flow is from the contents to the jacket, whilst a negative flow would model the inverse flow. The introduction of

the directions in the connections keeps the signs in the balances clean and the flow must only be defined once, a very logical request indeed.

The model we use for the heat flow is simply a discrete gradient with a conductivity as a parameter. This is assuming that the heat transfer through the inner reactor wall is very fast and has the consequence that the capacity effect of the wall is neglected. The quantity being computed is the heat flow, required is the temperature in the contents, whilst the temperature in the jacket is assumed to be known as well as the conductivity.

The next block is about the reaction ongoing in the liquid phase. The production rate vector is required for the differential species balances. The stoichiometry defines the ratio and contribution of the different extent of reactions, which in turn is given by a kinetic law. The latter is normed by the volume and the reaction constant, which is not constant, is given as a function of the temperature using the Arrhenius model in the next line. The composition dependency is captured in the $g(\cdot)$ function. In all cases, the result is the quantity on the left-hand-side, so all equations are explicit. This changes in the next block.

Here we close the two cycles by defining a mapping of the state into the secondary state. This part involves the definition of the effort variables, thus the equations of state (Caratheodory, 1909; Peterson, 1979; Rajeev, 2008), all the intensive quantities and geometry. The first is the concentration, an example of defining an intensive quantity defining the need for the volume. In this case we assume the volume to be given. In many applications though the volume has to be obtained through an equation of state. Lastly we have an implicit equation. In this case it is the enthalpy being the integral over the derivative with respect to temperature of the partial molar enthalpies multiplied by the molar masses in the reactor, thus the state. The quantity to solve for is the upper limit of the integral. Often the heat capacity is given as an empirical model that was fitted to experimental results. Here we assume this relation is given.

We construct the equations assuming a simulation, that is, the initial conditions and the parameters must be known to define a simulation problem. The reader should though be aware of the fact that the assumption of defining a simulation model is only used to make sure we generate a consistent algebraic structure. We do not initiate the problem, so once the algebraic structure is generated, one is absolutely free to define any mathematical/numerical problem that uses the model.

6. Conclusions

We take a view of mathematical behaviour theory to construct a complete representation of chemical-physical systems that are embedded in their environment of constant reservoirs. So we bring together the mathematical view and the physical view of thermodynamics, fluid flow and reaction kinetics all under the umbrella of a state-space representation of such systems. The definition of the model is done in reverse order than the computations starting with the initial conditions, we define the integrals, the differential balances that are dependent on flow and kinetic descriptions, one for the exchange of extensive quantities between parts of the system and the kinetics describing the transposition of the species internal to the capacities.

So the definitions are done starting from the state towards the secondary states, which then are tied to the state thereby closing the loop.

Writing the equations in this order assuming that one defines a simulation problem, provides the means to construct the model such that one can check on consistency and completeness. Latter is simply a degree of freedom argument that must be satisfied, namely the number of equations must match the number of variables. This is easily done by following the said procedure and using a table for the model equations. The degree of freedom argument becomes visible as the variables appearing in the third column, namely what is needed must match what is being computed, thus the second column.

The construction mechanism is chosen such that the consistency of the equations can be guaranteed. Taking the view of defining a simulation problem is the most convenient choice. The resulting algebraic structure can though be used for any problem formulation – it is by no means bound to a simulation problem. If indeed one has a simulation problem and thus instantiates what is in the above example shown as given, then also the computational sequence is defined, as one simply needs to invert the definition sequence of the equations: one starts with the states, solves for the mapping into the secondary states, which then enables the computation of the transport and the transposition, which when linearly combined proved the differential balances.

Acknowledgements

The various discussions with students in the group are greatly appreciated.

References

- Caratheodory, C., 1909. Untersuchung ueber die Grundlagen der Thermodynamik. *Mathematische Annalen*, 67(1):355–386. One of the key references on axiomatic thermodynamics.
- Peterson, M A , 1979. Analogy between thermodynamics and mechanics. *Am J Phys*, pages 488–490.
- Preisig, H. A. , 2010. Constructing and maintaining proper process models *Comp & Chem Eng*, 2010, 34(9), 1543-1555
- Preisig, H. A. , 2014. Visual modelling *FOCAPD 8 (Foundation of Computer-Aided Process Design)*, 730-734
- Preisig, H A, 2015. Constructing an ontology for physical-chemical processes, *PSE 12 \& ESCAPE 25*.
- Rajeev, S G, 2008. A Hamilton-Jacobi formalism for thermodynamics. *Annals of Physics*, 323:2265-2285.
- Willems, J C, 2014. personal webpage. lists and links to publications and lectures. URL <http://homes.esat.kuleuven.be/~jwillems/> last accessed 2014-11-20.
- Willems, J C, 1989. Some thoughts on modelling Newton to Aristotle: Toward a theory of models for living systems. In J Casti and A Karlquist, editors, *Proceedings of a Workshop held in Abisko, Sweden*, Birkhuser, pages 91–120.
- Willems, J C, 2006. Thoughts on system identification. In B A Francis, M Smith, and J C Willems, editors, *Control of Uncertain Systems: Modelling, Approximation, and Design -- Festschrift on the occasion of the 60-th birthday of Keith Glover*, volume 329, pages 389–416. Springer Lecture Notes in Control and Information Sciences.
- Willems, J C, 2007. The behavioral approach to open and interconnected systems. *IEEE Control System Magazine*.