February 10, 2003

A Survey of Co-Design Ideas and Technology

(draft)

by

G. Bosman

Supervisors:

Dr. Ir. A. M. Bos (Chess-iT)

Ing. P. G. C. Eussen (Chess-iT)

Dr. R. Lämmel (Vrije Universiteit)
Contents

Chapter

1 Introduction 1
  1.1 Traditional design ................................. 2
  1.2 Programmable logic ................................ 3
  1.3 Outline of this thesis ............................... 4

2 Assignment 5

3 Background 6
  3.1 History of Co-Design research ....................... 6
  3.2 Related work ........................................ 6
  3.3 Related approaches ................................ 6
    3.3.1 Software in the loop ........................... 7
    3.3.2 Reconfigurable hardware ....................... 7
    3.3.3 Adaptive computing .............................. 7
    3.3.4 Hybrid systems ................................. 8

4 Co-Design 9
  4.1 Abstraction ......................................... 9
  4.2 Specification and modelling ........................ 10
  4.3 Modelling approaches ............................... 11
    4.3.1 Homogenous modelling ......................... 11
# Computational Models

## 5.1 Dataflow vs. Control flow oriented

## 5.2 Common computational models

- **5.2.1 Finite-state machines**
- **5.2.2 Imperative models**
- **5.2.3 Differential Equations**
- **5.2.4 Difference Equations**
- **5.2.5 Process networks and dataflow**
- **5.2.6 Discrete-event models**
- **5.2.7 Petri Nets**
- **5.2.8 Synchronous models**
- **5.2.9 Rendezvous models**

## 5.3 Comparison

# Heterogeneous systems

## 6.1 Example: internal combustion engine

## 6.2 Old tools revived

## 6.3 New heterogeneous tools

## 6.4 Ptolemy II

## 6.5 COSYMA

## 6.6 Simulink

## 6.7 *charts

## 6.8 ForSyDe
Chapter 1

Introduction

Designing embedded systems is becoming more and more complex due to the increasing size of integrated circuits and time-to-market requirements. This increasing complexity of embedded systems is the driving motivation for a high-level system design approach. A high-level system design notation could also offer substantial performance benefits as there can be a smarter partitioning of the system’s parts onto hardware and software. High-level system design is currently a major research area by academics world-wide and it is strongly rooted in traditional areas such as control theory and digital design theory. For Chess such a design approach could mean a faster and better design process. This paper investigates a number of aspects of high-level system design and the current state of the art in existing tools and methods that allow high-level system design.

A system design notation should be at an abstract level that allows reasoning at a high level of abstraction. However, to be useful it must also be possible to synthesis code from this high-level language for hardware and software. How to divide the functionality of the system in hardware and software parts is a very important aspect of this synthesis. Traditionally the choice what to implement in hardware and what to implement in software is made early in the design process. Both parts are then made in separate tracks by separate design teams.

When the emphasis lies on the hardware-software partitioning problem, system-level design methods are also called "Co-Design" methods. These terms are often used mixed; in this paper Co-Design will be mainly used.
The goal of Co-Design is to explore the whole design-space to be able to make well-informed decisions and to be able to make this decision in a later phase in the design process. This would lead to a more optimal partitioning and more flexibility in the process from design to implementation. To really understand the improvement a high-level system design paradigm would make, it is important to see how traditional system design works and what the problems are with such an approach.

1.1 Traditional design

Traditionally complex embedded systems are designed around a microprocessor in a Von Neumann architecture. A Von Neumann system is fundamentally a sequential system. There is heavy optimisation inside the CPU itself (for instance using pipelining) but ultimately the commands are executed one at a time. Systems based on microprocessors have several benefits. Microprocessors are very well optimized and they allow design of families of products that can be built. Such a family of products can provide various feature sets at a different price point and can be extended to keep up with rapidly changing markets[49].

However, the fact that each command is executed sequentially leads to a fundamental limitation. Backus [1978] calls this the ”von Neumann bottleneck”. He points out that this bottleneck is not only a physical limitation, but has also been an ”intellectual bottleneck” in limiting the way we think about computation and how to program.

[Guus: explain difference between hardware and software as related to concurrency.] This same aspect is a major motivation for Chess to do this investigation: how to prevent the ‘paradigm-shift’ that often occurs in designing systems. [Guus: this paragraph should be better. Explain more what the paradigm shift is.]

Not all embedded systems are designed around microprocessors. It is possible to design embedded chips that compute in a parallel way. Application Specific Integrated Circuits (ASICs) are specialized chips that are used for example to implement encryption algorithms. They are similar to processors in the sense that they are also ’hardwired’ solutions. It is very expensive to
design an ASIC, and a very slow process. Therefore, customizing an ASIC for a single application is only feasible when the project is reasonably big. In any case an ASIC can only be produced when the layout is final: it is not possible to experiment with the layout and try several versions.

Although you’ll loose the optimalizations found in microprocessors there is a huge potential gain when designing hardware that is parallel in nature because the layout of the hardware can then be tailored exactly to the functional requirements. This can be extremely profitable, especially when the problem to be solved is mainly parallel in character. ASICs are therefore often used in areas such as compression or encryption which are parallel “in nature”.

An important motivation for the rising interest in Co-Design is the introduction of another type of chip that is much more flexible than ASICs: Programmable logic devices.

### 1.2 Programmable logic

Programmable Logic Devices (PLDs) are computer chips that can be programmed to implement circuits requiring both combinational and sequential logic. Reconfigurable logic devices are a class of programmable logic devices that may be reprogrammed as often as desired. FPGA’s are reconfigurable logic devices that are becoming very popular [9]. Three direct benefits of the reconfigurable approach can be recognized: specialization, reconfigurability and parallelism [41] [Gunns: use more of Tessier on future developments [41]]. FPGA’s shorten the development cycle dramatically and are much cheaper to use than ASICs. This allowed research in Co-Design to increase a lot. It also made researchers consider more fine-grained approaches.

Early approaches in Co-Design started with components of high granularity, such as ASICs and microprocessors. In Co-Design methods a mixture is often used. Parts of the hardware are designed from scratch, but a microprocessor is also used. This allows the system to benefit from both the microprocessor’s specialization, and for the parts of the system that benefit most of this a parallel implementations. This difference in level of granularity is an important feature to discriminate on various Co-Design approaches.

Obviously designing hardware and software for a system in an integral manner is an
extremely complex task with many aspects. There is a wide range of architectures and design methods, so the hardware/software Co-Design problem is treated in many different ways.

In this thesis it is investigated what methods exists that comply with the Chess business. I'll compare them with each other. I'll also investigate the problem of the paradigm shift (or the prevention of this) when designing parallelism in a high-level language. Ideally the resulting hard- and software description should be parallel in nature where possible.

In this paper I'll give an overview of the field and indicate the open issues. I'll try to answer to question which type of language is suitable for which problem.

1.3 Outline of this thesis

First we'll look at what Co-Design is and the problems it tries to solve. In Chapter 4.3 different types of Co-Design approaches are described, and the classifications that can be made. It turns out that the paradigms, the computational models, are very important. They are directly related to the paradigm shift. Chapter 5 describes computational models that are commonly used to model (parts of) systems.

A single paradigm approach has serious disadvantages so various hybrid models have been proposed in literature. In Chapter 6 existing projects and models will be described and analyzed using the classifications and ideas found in the first part. Per method we'll look at some case studies to get a better view of how the multi-paradigm modelling works and how well it can be applied. In Chapter 7 the pro’s and contra’s of the investigated heterogeneous methods will be discussed.
Chapter 2

Assignment

"To investigate various development methods and to investigate how an integrated approach of hardware-software design can improve system development at Chess."

I'll give an introduction to the field of Co-Design. I will investigate the relevance and relationships between these specification and design languages in the track from specification to implementation. I'll research to which degree existing development and implementation methods support compatible paradigms.

The focus in this paper will be on the shift of paradigms when traversing through the various levels of detail.
Chapter 3

Background

3.1 History of Co-Design research

The field of Co-Design is about 10 years old now. [Guus: about Gupta paper etc]

In 1998 a paper was published that described a method to synthesize code for both hardware and software, for a specific type of data-flow programs[14].

Research in Co-Design is a big field. Three institutes worth mentioning are the Royal Institute of Technology (Stockholm, Sweden), the Indian Institute of Technology (Delhi, India), University of California (Berkeley, USA).

3.2 Related work

The SAVE project of the Linköping University in Sweden did a survey on Co-Design representation models in 1999[7]. At the Leiden University in the Netherlands a survey was done comparing 8 tools and their underlying methodologies[48]. O’Nils presented ComSys, an approach to the generation of interfaces between application software and hardware IP components. In his PhD thesis from 1999 he reviewed several Co-Design methods[35]. [Guus: about other comparative papers; the one comparing 6 systems for example.]

3.3 Related approaches

A number of areas that are related to Co-Design have been researched.
3.3.1 Software in the loop

Some of the issues Co-Design tries to solve are also handled by ‘Software-In-The-Loop’. This is developing software in a virtual hardware environment. Although this eases the design of software for hardware it does not allow the full improvements made possible by Co-Design because the partitioning between hardware and software is not flexible. Within Chess a chip has been developed (the SHAM) that allows software testing for onboard software[45].

3.3.2 Reconfigurable hardware

Reconfigurable systems exploit FPGA technology, so that they can be personalized after manufacturing to fit a specific application.

"A promise of reconfigurable hardware is that it should allow the logic and memory resources in a chip to be used more efficiently, especially in applications that need massive computing power. But there is a further commercial advantage. It could turn finished products into a source of service revenue. Imagine a music player that includes programmable logic. When a new music-compression format emerges to replace MP3, owners of the player could download, for a fee, a new decompression algorithm for their player from the maker’s website”[42].

[Guus: Add reference to Viales or ITA project here (reconfigurable parts)]

The operation of reconfigurable systems can either involve a configuration phase followed by an execution phase or have concurrent (partial) configuration and execution.[32]. The major problem in this type of systems consists of identifying the critical segments of the software programs and compiling them efficiently to run on the programmable hardware. This is a different field and will not be treated in this thesis.

3.3.3 Adaptive computing

The field of adaptive computation is closely related to reconfigurable hardware. According to Neema the primary challenge of the Adaptive Computing approach is in system design[34].
3.3.4 Hybrid systems

Most traditional Co-Design methods explore ways of modelling digital systems. Embedded systems however, often interact with an analog environment. Traditionally, this is the domain of control theory and related engineering principles. Because of the way models are often treated (digital) the analog environment is often abstracted to a digital translation by computer scientists. This way traditional Co-Design is unable to guarantee safety and/or performance of the embedded device as a whole.

To address this issue hybrid embedded system models have been designed[1, 40]. The issues that Co-Design faces, such as combining various models of computations and making sure properties are valid throughout the whole design phase, can of course also be found in this hybrid system modelling. In fact, the difference between the two is not always very sharp. [TBD: on differential equations].
Chapter 4

Co-Design

Co-Design is "A design methodology supporting the concurrent development of hardware and software in order to achieve system functionality and performance goals. In particular, Co-Design often refers to design activities prior to the partitioning into Hardware and Software and the activity of design partitioning itself." [47].

4.1 Abstraction

A goal of all design methods is to allow systems to be designed on a higher level than the implementation level. The ultimate goal is to allow a very high-level design, that then automatically can be converted into the implementation level. This is a fundamental notion in computer science. Examples are programming languages, that allow humans to reason about variables or flow-of-control on an much higher level than machine code allows. There has also been a lot of research into finding languages to design hardware from a higher level. Nowadays it is very common to use a language as VHDL to define hardware. There are compilers available to generate netlists (hardware descriptions) from languages like VHDL. Although VHDL is probably considered by software engineers to be low-level, it is a major step forward compared to the arcane art of programming cells and gates directly.

This looking for a higher level of abstraction is an ongoing quest, and as so it can also be found in Co-Design methods. Ultimately the goal is to be able to design a system in a textual or graphical way, in such a manner that there will be an automatic compiler from this high-level
representation into the implementation level: hardware, software or (often) a combination of both. Sometimes such an approach is called model compilation[38].

A higher level of abstraction in modelling decreases the gap between functional requirements of a system and the modelling process, leading to a better fit between these.

The Co-Design system design process for embedded systems includes modelling, validation, and implementation[32].

4.2 Specification and modelling

Modelling is the process of conceptualizing and refining the specification. The result of the modelling phase is a model, which is specified in an internal design representation. There are several tasks that must be performed to create a system-level design model. To comprehend the benefits of various Co-Design technologies it is important to understand how the design process works.

'There is a subtle relationship between the specification of a system and the modelling of a system. An executable specification, for example, is also a model of an implementation. The difference is in emphasis. A specification describes the functionality of a system, and may also describe one or more implementations. A model of a system describes the functionality. In a specification it is important to avoid over-specifying the design, to leave implementation options open. In a model, often the key criteria are precision, simplicity and efficient simulation. A model should be the most abstract model that represents the details being tested.'[6]. [Guus: what to do with the paragraph?]

Specification is closer to the problem level, at a higher level of abstraction, and uses one or more models of computation. A specification undergoes a synthesis process (which may be partly manual) that generates a model of an implementation. That model itself may contain multiple models of computation. The outcome of the modelling process is the Internal Design Representation (IDR). There is a trade-off between scalability and expressiveness in this IDR[46].
4.3 Modelling approaches

Approaches to hardware/software Co-Design of embedded systems can be differentiated in several ways. One way is to consider the system-level specification, which is either homogeneous (in a single specification language) or heterogeneous (involving multiple specification languages)[33]. Another way is to distinguish how the design methods deal with the SW/HW partitioning: fine-grained or coarse-grained. Modelling in the context of Co-Design is sometimes called cospecification.

4.3.1 Homogenous modelling

Homogeneous modelling implies the use of single specification language for the modelling of the overall system. Lee[25] calls this the 'grand unified method'. Co-design starts with a global specification given in a single language. This specification may be independent of the future implementation and the partitioning of the system into hardware and software parts. In this case Co-Design includes a partitioning step aimed to split this initial model into hardware and software. The outcome is an architecture made of hardware processors and software processors. The is generally called a virtual prototype and may be given in a single language or different languages[20]. Lee sees as a big problem that a homogenous approach imposes a model of computation which might be good for a subset of systems but bad for others[25].

This approaches is called compositional by Coste[8]. It aims at integrating the partial specification of sub-systems into a unified representation which is used for the verification and design of the global behavior. Examples are Polis, COSYMA and SpecC.

4.3.2 Heterogeneous modelling

Heterogeneous modelling allows the use of separate languages for the hardware and software parts. The Co-Design starts with a virtual prototype where the hardware/software partitioning is already made. Here the emphasis is on the integral designing of the parts to make sure the overall system has the required properties. The key issues are validation and interfacing
A lot of research is done on the integration of different system parts that enables system optimization across language boundaries. The cosimulation-based approach consists in interconnecting the design environments associated to each of the partial specifications. Like its name suggests, with co-simulation the software parts and the hardware components of a system and their interactions are simulated in one simulation. It does not provide such a deep integration as compositioning does however it does allow for modular design. Communication is often done using a cosimulation bus, that is in charge of transferring data between the different simulators.

The Cosimulation-field is reasonably well established. Sometimes cosimulation is used to simulate the behavior of a system consisting of 2 models: the hardware and the software, and sometime it is used to model on a more abstract level where the hardware vs software decision has not been made yet. A typical example of a Co-Design approach is Ptolemy II.

4.4 Validation

Through the validation process, the designer achieves a reasonable level of confidence about how much of the original embedded system design will be in fact be reflected in the final implementation. There are 3 three methods for validation:

(1) Simulation

(2) Prototyping

(3) Formal Verification

There has been a lot of research in the simulation of heterogeneous hardware/software systems. Formal verification allows for a more thorough test of the embedded system behavior (maximum behavioral coverage) by means of logics.

It is good to note that in simulation the hardware is often simulated (although often not real-time, as it’s just a simulation). However there has also been some research in replac-
ing the hardware simulator with an FPGA (or multiple FPGAs) that simulate the real target hardware[24].

4.5 Implementation

The model found throughout the modelling phase has to be implemented into hardware and software. It is an important notion in a Co-Design approach that there is no continuity problem. That is: the steps from model to the synthesis should be all in the design process[38].

In the phase architectural information is taken into account. Varea[46] calls this a merger between the IDR and the technology library. It is important that the intermediate IDR or specification is not too operational (influenced by the current technology) or it will bias the design towards a specific architecture. Steps in the implementation phase include:

(1) Hardware/Software partitioning

(2) Synthesis of the code for software and hardware

4.5.1 Hardware/software partitioning

"The partition of a system into hardware and software is of critical importance because it has a huge impact on the cost/performance characteristics of the final design. In the case of embedded systems, a hardware/software partition represents a physical partition of system functionality into application-specific hardware and software executing on one (or more) processor(s)."[32]. When considering general purpose computing systems, a partition represents a logical division of system functionality, where the underlying hardware is designed to support the software implementation of the complete system functionality. This division is elegantly captured by the instruction set. Thus instruction selection strongly affects the system hardware/software organization.

Obviously is important to look at the architectural organization of the system. Although it is possible to generate a complete system using only FPGA’s, it is very common to use a
combination of 1 (or more) processors with dedicated hardware. This is called coprocessing.

On the lowest level, FPGA’s can be used to implement SM’s, datapaths and nearly any digital circuit. The outcome of the synthesis process is a final implementation of the embedded system.

It is clear that the choice of an IDR is a very important aspect of a Co-Design method, therefore in the next chapter we’ll go into the various types of IDR’s there are.
Chapter 5

Computational Models

Modelling is a the heart of development methods. The computational models can be found in the Immediate Representation Language. All Co-Design systems are based on a computational model, or combine a few of them.

A computational model is a formal, abstract definition of a computer. It describes the components in a system and how they communicate and execute. Several models exist. There are a number of authors who made an overview of various development methods, i.e. [50], [20].

[Guus: here explain what I want to find about of the models]. Ie, timing, hierarchy.]

An essential difference between concurrent models of computation is their modelling of time. [25] (page 11). Lui[27] states that the different notions of time make programming of embedded systems significantly different from programming in desktop, enterprise or Internet applications. Lee[26] proposed a mathematical framework to compare certain properties of models of computation. This allows for a precise definition of the various computational models.

In the Chapter a few 'basic' Models of Computations are described.

5.1 Dataflow vs. Control flow oriented

[Guus: insert examples here]. Fundamental to embedded software is the notion of concurrency. There is a lot of research done on compiling concurrent languages into sequential code that can be run on a microprocessor, see for example [21]. For this thesis however it is more interesting to see what happens when this paradigm-shift does not have to be made.
There are models that are designed to describe dataflow oriented systems (ie DSPs) and there are models more suitable for control-flow systems. However this approach lacks generality as most systems are not easily put in either one category. It should be noted that the difference between a control-flow or data-flow oriented computational model is important for both control and data-oriented systems. Many control-systems use complex sensors or subsystems such as image processing algorithms that are best specified using a type of data-oriented computational model[28].

5.2 Common computational models

5.2.1 Finite-state machines

The finite-state machine model has been widely used in control theory and is the foundation for the development of several models for control-dominated embedded systems. The classical FSM consists of a set of states, a set of inputs, a set of outputs, a function which defines the outputs in terms of input and states and a next-state function.[7].

FSMs model systems where the system at any given point in time can exist in one of finitely many unique states. This makes them excellent for control logic in embedded systems. They can very well be formally analyzed and it is relatively straightforward to synthesis code from this model[25].

FSM can be visualized very well using a state transition graph (see Figure 5.2.1). FSM have a number of weaknesses. They are not very expressive, and the number of states can get very large even in the face of only modest complexity. Is intended for control-oriented
systems with relatively low algorithmic complexity. A number of variations has been proposed to overcome weaknesses of the classical FSM model. Using FSMs in a hierarchical model was first made popular by Harel[28]. He proposed StateCharts, which combine hierarchical FSMs and concurrency. In Chapter 6.7 we’ll see some examples of heterogenous models based -partly-on FSMs.

5.2.2 Imperative models

In an imperative model of computation, modules are executed sequentially to accomplish a task. This is the most trivial computational model there is and most authors don’t even mention it. However in Chang’s paper[6] it is mentioned because it can be used in combination with other models in a hierarchical system. See also Chapter 7.3.

5.2.3 Differential Equations

These are often used to model mechanical dynamics, analog circuits, chemical processes and many other physical systems. [15].TBD. When using real numbers as time model, continuous-time systems are active over the entire time axis processing their input and producing output.[40].

5.2.4 Difference Equations

Like differential equations, but discrete. TBD. These two are very important as they deal with a very common type of signal from the outside world. Discrete-time systems can only react to their input and produce new output at distinct, equidistant time instances. [40]

5.2.5 Process networks and dataflow

In a process network model of computation the arcs represent sequences of data values (token) and the bubbles represent functions that map input sequences into output sequences. Certain technical restriction are necessary to ensure determinacy.[25]. It is a common represen-
Figure 5.2: A dataflow graph for \((a \times b) + \frac{c}{d} - \sqrt{e \mod f}\).

tation formalism for modeling algorithms. The graph representation can be interpreted asynchronous (ADF) or synchronous (SDF). [Guus: uitwerken wat voor soort data-flow modellen er allemaal bestaan; split it into two?]

5.2.6 Discrete-event models

In a discrete-event system, modules react to event that occurs at a given time instant and produce other events either at the same time instant or at some future time instant. Execution is chronological[6]; time is an integral part of the model. Events will typically carry a time stamp, which is an indicator of the time at which the event occurs within the model. A simulator for Discrete-Event models will typically maintain a global event queue that sorts events by time stamp. This sorting can be computationally costly. [Guus: hard to simulate, nice in hardware].

5.2.7 Petri Nets

In the classical approach a Petri net is composed of 4 basic elements: a set of places, a set of transition, an input function that maps transitions to places, and an output function which is also a mapping from transition to places. This is an well-understood modelling tool. Two important features of Petri nets are its concurrency and asynchronous nature.[7]. [Guus: expand a little bit: where are petri nets good for?]
5.2.8 Synchronous models

In synchronous languages, modules simultaneously react to a set of input events and instantaneously produce output events. If cyclic dependencies are allowed, then execution involves finding a fixed point, or a consistent value for all events at a given time instant.\[6\]

Very often real-time systems are specified by means of concurrent processes, which communicate asynchronously \[37\].

The synchrony hypothesis forms the base for the family of synchronous languages. It assumes, that the outputs of a systems are synchronized with the system inputs, while the reaction of the system takes no observable time. So time is abstracted away. The synchrony hypothesis abstracts from physical time and serves as a base of a mathematical formalism. All synchronous languages are defined formally and system models are deterministic.

In synchronous languages, every signal is conceptually (or explicitly) accompanied by a clock signal. The clock signal has meaning relative to other clock signals. It defines the global ordering or events. Thus, when comparing two signals, the associated clock signals indicate which events are simultaneous and which precede or follow others. A clock calculus allows a compiler to reason about these ordering relationships and to detect inconsistencies in the definition.\[6\].

This model serves as a good implementation model.

5.2.9 Rendezvous models

In a rendezvous model, the arcs represent sequences of atomic exchanges of data between sequential processes, and the bubbles presents the processes. \[25\]. Examples are Hoare’s CSP and Milner’s CCS. This model of computation has been realized in a number of concurrent languages, like Lotos and Occam. [Guus: based on algebra? How is timing handled?]

5.3 Comparison

\[7\] also made a comparison of various computational models.
Table 5.1: The main characteristics of the models of computation described in this Chapter.

<table>
<thead>
<tr>
<th>Computational model</th>
<th>Main application</th>
<th>Timing</th>
<th>Modus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential Equations</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Difference Equations</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Kahn Process Networks</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>SDF</td>
<td>?</td>
<td>No explicit timing</td>
<td>Synchronous</td>
</tr>
<tr>
<td>ADF</td>
<td>?</td>
<td>No explicit timing</td>
<td>Asynchronous</td>
</tr>
<tr>
<td>Discrete-event models</td>
<td>?</td>
<td>Globally sorted events with time tag</td>
<td>?</td>
</tr>
<tr>
<td>Petri Nets</td>
<td>?</td>
<td>No explicit timing. Just order of</td>
<td>Asynchronous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>transitions[7]</td>
<td></td>
</tr>
<tr>
<td>Synchronous/reactive models</td>
<td>?</td>
<td>No explicit timing</td>
<td>Synchronous</td>
</tr>
</tbody>
</table>
Chapter 6

Heterogeneous systems

Heterogeneous systems are systems that allow more than one computational model in a system to be used. Experience suggest that several MoC are required for the design of a complete system.[6]. A nice illustration of the need for multiple computational models can be found in [27]:

[Guus: how can you determine system behaviour for these types of systems? Things like deadlocks, lifelock, timing...]

6.1 Example: internal combustion engine

Figure 6.1: An internal combustion engine is made up of several parts that should be modelled by different MoCs.
A cylinder of an internal combustion engine has four working phases: intake, compress, explode, and exhaust. The engine generates torque that drives the power train and the car body. Depending on the car body dynamics, the fuel and air supply, and the spark signal timing, the engine works at different speeds, and thus makes phase transitions at various time transitions. The job of the engine controller is to control the fuel and air supplies as well as the spark signal timing, corresponding to the drivers demand and available sensor information from the engine and the car body.

The engine and the car body in this example are mechanical systems, which are naturally modelled using differential equations. The four phases of the engine can be modelled as a finite state machine, with a more detailed continuous dynamics for the engine in each of the phases. While all the mechanical parts interact in a continuous-time style, the embedded controller, which may be implemented by some hardware and software, works discretely.

Additionally, sensor information and driver’s demands may arrive through some kind of network. The controller receives this information, computes the control law, controls the air and fuel values, and produces spark signals, discretely. So, we want to use a model that is suitable for handling discrete events for the network and the controller.

In this very common example, we have seen both continuous-time models and several discrete models: finite state-machines, discrete events, and real-time scheduling.

It's common for a specification language to allow more than 1 model of computation. However this does not always mean that this allows for suitable high-level mixture of 2 models. An imperative language can be used to implement for example a dataflow MoC[6]. It is obvious that low-level languages such as VHDL are able to implement different models of computation. However, their lack of abstraction disqualifies them as candidates for modelling combined computational model-systems because it leaves programmers no freedom to make trade-offs between programmability, utilization of resources and silicon area.
6.2 Old tools revived

Many languages and tools that were developed based on a single model start to embrace other models [15]. The downside of such large languages with multiple MoC’s is (according to [6]) that formal analysis may become very difficult. It compromises the ability to generate efficient implementations or simulations and makes it more difficult to ensure that a design is correct. It precludes such formal verification techniques as reachability analysis, safety analysis and liveness analysis.

Most complex system are a combination of data- and control-flow oriented parts. Varea[46] proposes a classification according to the following taxonomy:

(1) Models originally developed for control-dominated embedded systems and later expanded to include data-flow (these models will be called \( M_{CD} \)).

(2) Models developed in a data-dominated basis extended to support also control flow (referred to as \( M_{DC} \))

(3) Unbiased model developed specifically to deal with combined control/data-flow interactions (\( M_{b} \))

[Guus: use this classification!]

As noticed in Chapter 3.3.4 the modelling of hybrid digital-analog systems is a related field that is gaining more attention too. Also in this field there are existing tools that are extended with functionality to deal with hybrid modelling. Examples of this are VHDL-AMS.

6.3 New heterogeneous tools

However also new frameworks have been developed that took multiple models of computations into their design from the beginning. A framework is a software architecture that specifies the possible interactions of components, provides a set of services that components can use and may have a set of formal properties for the system.
In this Chapter we’ll see a few of the most famous modelling methods and a few that have been selecting because they are special. [Yes, this should be a different sentence].

6.4 Ptolemy II

The Ptolemy project studies heterogeneous modelling, simulation and design of concurrent systems, where the focus is on embedded systems.[11].

The primary investigator of the Ptolemy project is Edward A. Lee. In 1991 he presented a paper[5] that described the Ptolemy system. This system has been in use for many years, and it’s now succeeded by a new version, Ptolemy II.

The Ptolemy II software environment provides support for hierarchically combining a large variety of models of computation and allows hierarchical nesting of models.[15]. It combines the wish for a homogeneous and thus predictable model with the desire to mix partial models of different kinds in a common heterogeneous model by hierarchically nesting sub-models of potentially different kinds.

A very good description of how this hierarchical mixed approach works in practice can be found in [28].

![Figure 6.2: Example of a hierarchical specification of a systems using two (possibly different) Models of Computation. The directors controls the flow of control and data in such a MoC.](image)

Ptolemy II is a component-based design methodology. The components in the model are
called actors. A model is a hierarchial composition of actors. The atomic actors, such as A1, only appear at the bottom of the hierarchy. Actors that contain other actors, such as A2, are composite. A composite actor can be contained by another composite actor.

Atomic actors contain basic computation, from as simple as an AND gate to more complex as an FFT. Through composition, actors that perform even more complex functions can be built. Actors have ports, which are their communication interfaces. For example in Figure 6.2, A5 receives data from input ports P3 and P4, performs its computation, and sends the result to output port P5. A port can be both an input and an output. Communication channels among actors are established by connecting ports.

The possibility to have various MoC’s can be found in the director. A director controls the execution order of the actors in a composite, and mediates their communication. In figure 1, D1 may choose to execute A1, A2 and A3 sequentially. Whenever A2 is executed, D2 takes over and executes A4-A7 accordingly. A director uses receivers to mediate actor communication. As shown in figure 2, one receiver is created for each communication channel; it is situated at the input ports, although this makes little difference. When the produces actors send a piece of data (a token) to the output port, the receiver decides how the transaction is completed. Within a composite actor, the actors under the immediate control of a director interact homogeneously.

6.5 COSYMA

A bit older design method is COSYMA, "CoSynthesis for Embedded Architectures". It was developed at the IDA, Germany. It covers the entire design flow, from specification, to synthesis. The target architecture consists of a standard RISC processor, a fast RAM for program and data with single clock cycle access time and an automatically generated application specific co-processors. Communication between processors and co-processor takes place through shared memory.

The system is designed in Cx. This is a C-extension with support for parallel processes and timing constraints. The Cx specification is then converted into an Extended Syntax Graph
(ESG), the IDR. The ESG describes a sequence of declarations, definitions and statements and is overlayed with the Data Flow Graph (DFG) containing information about data dependencies.

Research using COSYMA has been discontinued in 1999.

6.6 Simulink

Existing (commercial) tool. [27] calls it a framework. A modelling and simulation environment for continuous-time dynamic systems with discrete events. It has been extended with StateFlow[28].

6.7 *charts

Statecharts are essentially a combination of FSMs with a SR. The tools Statemate from Ilogix uses statecharts as its control specification model.

A recent development is *Charts (pronounced Starcharts). TBD

[Guus: SOLAR here[8]]. Another model based on Finite State Machines is the CFSM model. The communication primitive is called event.[7].

6.8 ForSyDe

An interesting method has been development by Sander and Jantsch[36, 37]. In their model events are totally-ordered by their tags. Every signal has the same set of tags. Events with the same tag are processed synchronously. There is a special value \(\bot\) ("bottom") to indicate the absence of an event. These are necessary to establish a total ordering among events. A system is modelled by means of concurrent processes; it is a model based on the synchronous-assumption (see Chapter 5.2.8).

Lu[29] shows how to transform a system specification described in ForSyDe into its hardware and software counterparts. He does not provide a mixed implementation of HW and SW. [Guus: why not per module possible to make this decision?]
The hardware version of the Digital Equalizer that Lu makes is described using behavioral VHDL. The process are described using skeletons and these are then synthesized to VHDL code. The process described is manual. The Haskell code turns into behavioral VHDL quite easily.

To generate (naturally sequential) C code an analysis phase is done to create a PASS. [4] also investigated the design of a Digital Equalizer. They used a combination of SDL and Matlab as their design languages.

SDL is used to model the control parts, Matlab is used for the DSP parts.

6.9 Moses framework

6.10 Polis

The Polis research project started in 1988 by the UC Berkeley. It is a design environment for control-dominated embedded systems. It supports designers in the partitioning of a design and in the selection of a micro-controller and peripherals.

The system specification language is Esterel, but a graphical specification can also be given.

The generated software part consists of:

1. the application that has been modelled in CFSMs

2. a generated application-specific operating system for the selected processor

3. the I/O drivers

Hardware is synthesized as well. The Polis environment provides an interface for verification and simulation tools as well as an simulator.

A fundamental limitation of the Polis system is the MoC used, the Codesign Finite State Machine (CFSM). A CFSM is an extended finite state machine that communicates with other CFSMs asynchronously with unbounded delay and by means of events. The communication model between CFSMs is not efficient in representing systems with intensive data processing,
since CFSMs communicate over channels with one-place buffers and have non-blocking write communication semantics. Therefore, a buffer is overwritten every time the sender emits an event before the receiver has consumed the previous event. This can be avoided either by means of scheduling constraints or with a blocking write protocol: however, both mechanisms often result in a loss of performance. This means that POLIS is mainly useful for control-dominated embedded systems. Although the POLIS method allows performance-estimation for the simple controller that is generated, estimation techniques for more complex processor models are lacking[48].

The Polis project has led to a set of commercial tools called Felix in 1998. It allows integration with Ptolemy, which makes it quite a powerful tool. [Guus: expand relationship Felix, Polis, VCC and Metropolis].

6.11 Metropolis

6.12 Solar/Music

[8] describes a multi-language approach at the system level providing both system-level refinement and high-level interfaces synthesis.

6.13 FunState/SPI Workbench

FunState is an internal design model based on functions driven by state machines[43]. It is an enhancement of the older SPI Workbench[51].

The SPI Workbench [51] is based on intervals of system properties and is specifically targeted to cosynthesis. Made for performance estimations. [Guus: Funstate = new version of SPI workbench?]
Table 6.1: The main characteristics of systems described in this Chapter.

<table>
<thead>
<tr>
<th>Model</th>
<th>MoC based on</th>
<th>Main application</th>
<th>Target architecture</th>
<th>Specification languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>*charts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSYMA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ForSyDe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FunState</td>
<td></td>
<td>FSM with functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRSYD</td>
<td></td>
<td>Flow Charts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metropolis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polis</td>
<td>CFSM</td>
<td>Control-dominated</td>
<td>Esterel, graphical</td>
<td></td>
</tr>
<tr>
<td>Ptolemy II</td>
<td>Multiple MoC's</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SimuLink</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpecC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.14 IRSYD

IRSYD[16] is an acronym that stands for Internal Representation for System Description. It is thus an internal design representation language; not a complete method. However IRSYD is fully defined and implemented (in C++).

It is quite ambitious in the sense that it claims to be able to be used for synthesis, performance estimation, formal verification etc.

The basic idea behind IRSYD is a unified graph representation for control flow and data flow. It is a Flow Chart extended to handle hierarchy, both structural and behavioral, different data types and different communication mechanisms. A flow chart is an informal graphical description of an algorithm built as a directed graph.

Unfortunately the work on this representation seems to have stopped. The latest papers are from 1998. One of the authors of some papers about IRSYD, Axel Jantsch, is also working on ForSyDe (see Chapter 6.8).

6.15 Comparison

[Guus: Add Varea’s comparison to the paper.] Most approaches described are component based.

Not every approach that has been described support all the phases of Co-Design described
Table 6.2: The supported phases of the Co-Design development process.

<table>
<thead>
<tr>
<th>Model</th>
<th>Simulation</th>
<th>Formal analysis and verification</th>
<th>Synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>*charts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSYMA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ForSyDe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FunState</td>
<td></td>
<td>FSM with functions</td>
<td></td>
</tr>
<tr>
<td>IRSYD</td>
<td></td>
<td>Flow Charts</td>
<td></td>
</tr>
<tr>
<td>Metropolis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ptolemy II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SimuLink</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpecC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Chapter 4, the supported steps are indicated.
Chapter 7

Discussion of hybrid models

7.1 Design-space exploration

Many parts in the design of embedded system require manual decisions, this remains so when using Co-Design methods. The integration Co-Design offers is valuable because of validation and easier synthesis of code from a model that allows hard- and software to be generated.

Because of the complexity of most systems, optimal manual decisions are sometimes not feasible. There are simply too many possibilities to consider. To use all of the potential improvements that a later HW/SW partition decision allows, it is therefore very important to reduce the user-decisions as far as possible. This has been recognized and there are various methods for systematic design space exploration cite(cassidy). ”In order to perform rigorous analysis and synthesis it is essential to prune the design space retaining only the most viable alternatives.” In the past heuristics have been used to prune large design spaces. However, due to the complex behavior and interactions in multi-modal systems it is difficult to come up with effective heuristics. A better approach is to use constraints to explore and prune the design spaces; constraint satisfaction can eliminate the designs that do not meet the constraints. The pruned design space contains only the designs that are correct with respect to the applied constraints. These designs can then be simulated, synthesized and tested.” [34].
7.2 Target architecture

When a Co-Design methodology allows for the generation of both software and hardware, it must also generate the communication mechanisms between these two parts. This includes the operating system perhaps, and the device drivers of some sort.

O’Nils point out that very often off-the-shelf IP components are used in system design, and that often a major part of the work will be in interfacing these IP components. Tools like Polis are primarily designed for cases in which the whole design functionality is captured within the tool’s environment and communication refined during system synthesis. That is, the device drivers are generated together with the custom hardware and the operating system. However, if users want to use IP blocks and off-the-shelf operating systems they will face the same problems that occur in manual design[35]. This has important implications for the commercial use of Co-Design methods and design space exploration tools: if they don’t take into account off-the-shelf IP components they are not suitable for many types of projects.

An extension to most Co-Design tools would be an explicit mapping between the IDR and the target architecture. Keutzer et al. state: "We actually believe that worrying about HW-SW boundaries without considering higher levels of abstraction is the wrong approach. HW/SW design and verification happens after some essential decisions have been already made, and this is what makes the verification and the synthesis problem hard. SW is really the form that a behavior is taking if it is "mapped" into a programmable microprocessor or DSP. [...] The origin of HW and SW is in behavior that the system must implement."[22]. If this mapping is recognized together and used in design-space exploration, a typical approach would be the Y-Chart\(^1\) as proposed by Kienhuis[23] (see Figure 7.2).

\(^1\) This term is rather unfortunate, as there exists another Y-Chart in Co-Design. That one relates to the level of abstraction of a system [Guus: TBD]
7.3 Hierarchy/emergent behavior

‘Brute-force composition of heterogeneous models may cause emergent behavior. Model behavior is emergent if it is caused by the interaction between characteristics of different formal models and was not intended by the model designer.’[15]

A common way to prevent unwanted emergent behavior is isolating various subcomponents and letting these subcomponents work together in a hierarchical way. Hierarchical in the sense of a containment relation, where an aggregation of components can be treated as a (composite) component at a higher level. In general, hierarchies help manage the complexity of a model by information hiding – to make the aggregation details invisible from the outside and thus a model can be more modularized and understandable[27].

‘Note that in Ptolemy, models of computations are mixed hierarchically. This means that two MoC’s do no interacts as peers. Instead, a foreign MoC may appear inside a process. In the old version of Ptolemy, such was process is called a wormhole. It encapsulates as subsystem specified using one MoC within a system specified using another. The wormhole must obey the semantics of the outer MoC at its boundaries and the semantics of the inner MoC internally. Information hiding insulates the outer MoC from the inner one.’[6]. This approached of wormholes was a bit biased towards data-flow computational models. In Ptolemy II it was replaced

There are other ways to mix models of computation too. Statemate uses views. [Guus: are these views related to what Jantsch[19] calls analytical slicing into domains?]

7.4 Synthesizing code

A discrete-event model of computation is well suited for generating hardware. It is not very suitable to generate (sequential) software[6] (p. 131). This is for example why VHDL simulate the the designer by taking so long. A model that heavily uses entities communicating through signals will burden the discrete-event scheduler and bog down the simulation. Thus, a specification built on discrete-event semantics is a poor match for implementation is software.

By contrast, VHDL that is written as sequential code runs relatively quickly but may not translate well into hardware. The same goes for C: it runs very quickly and is well suited for software, but not for specifying hardware. [Guus: Expand this. What are the problems with C?]

It is possible to conclude that the two issues found in this paper are closely related. If you want a single specification language you’ll loose in the paradigm-shift. It is hard to imagine an (efficient) language that allows both control- and dataflow types to be presented and generate efficient code for it for all types of applications. On the other hand, if you don’t mind taking the HW/SW decision earlier there are very good integrated tools and frameworks that allow working with both parts of your system in a systematic way. Code generation (or hardware generation) is easier in this style.

[23] also realizes this. He says: "the refinement approach has proven to be very effective for implementing a single algorithm into hardware. The approach is, however, less effective for a set of applications. In general, the refinement approach lacks the ability to deal effectively with making trade-offs in favor of the set of applications."

There is also a mixed form possible. This mixed form would not be applicable for every type of system, but only for a subset. An example of such an approach is ForSyDe. They allow
the specification of both control- and dataflow parts in a single language. They have shown to be able to generate (reasonably) efficient code. A thing to investigate would be for what types of systems this kind of Co-Design approaches are suitable, and for which not.

Dataflow and finite-state models of computation have been shown to be reasonably re-targettable. Hierarchical FSMs such as Statecharts can be used effectively to design hardware or software. It has also be shown that a single dataflow specification can be partitioned for combined hardware and software implementation.'[6] [Guus: very interesting, expand this].

The difference made between heterogeneous and homogenous modelling seems to be a bit to blunt. Some authors make it look like heterogenous modelling is the only answer possible because that’s the only type of method that allows all kinds of systems to be modelled. However, some homogenous modelling tools allow for a subset of applications to be modelled. This way –for specific applications– the benefits of having a single language will be available without a drawback in the paradigm-shift. This relates to the conventional wisdom that high performance while minimizing resources needed (or time needed) can be obtained by matching the architecture to the algorithm[34].

The main difference in the two approaches in my view is that the true heterogenous style (like Ptolemy II) allow for more types of MoC’s, and are better capable of working with new MoC’s. The other option promises a closer integration though.

7.5 Paradigm shift

It has been recognized in literature that there is an important relationship between the model of computation and the target-architecture. Kienhuis et all.[23] speak in this context about a mapping between a model of computation and the architecture: "In mapping we say that a natural fit exist if the model of computation used to specify applications matches the model of architecture used to specify architectures and that the data types used in both models are similar."

Algebraic formal methods are not capable of dealing with the complexity of complete
(embedded) systems design technologies. The algebra is not sophisticated enough and the design technologies are not suited toward formal verification. For the description of MoC’s[26] and the interactions between them formal methods have been very valuable.
Chapter 8

Conclusions

A synergistic approach of hardware and software and taking design to higher level are
nowadays recognized as mandatory to keep up with the increasing complexity of embedded
systems design.

Most Co-Design tools make use of an internal representation for the refinement of the
input specification and architectures[20]. Many such internal representations exist. They are all
based on one or more Computational Models.

Experiments with system specification languages have shown that there is not a unique
universal specification language to support the whole design process for all kinds of applications.
[20].

The computation model to use is first of all dependent on the type of problem to be solved.
There is a definite trend towards heterogeneous modelling systems that allow more mixed types
of systems to be modelled.

Many researchers are doubting whether a grand unified approach will work. Specifically
the group of Edward A. Lee (who created the Ptolemy system) has doubt about this[6]. In order
to be sufficiently rich to encompass the varied semantic models of the competing approaches,
they become unwieldy, too complex for formal analysis and high quality synthesis. He claims
that generality can be achieved through heterogeneity, where more than model of computation
is used.

Ptolemy II seems to be a very mature tool, specifically the theoretical foundation of
mixing various computational models is well thought-out. However, it is mainly focussed towards simulation.

[Guus: etc etc... this will take a lot of work!]
Bibliography


41


# Appendix A

## Vocabulary

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASIC</td>
<td>Application specific Integrated Circuit</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field-Programmable Gate Array (a specific type of PLD)</td>
</tr>
<tr>
<td>FSM</td>
<td>Finite State Machine (see Chapter 5.2.1)</td>
</tr>
<tr>
<td>IDR</td>
<td>Internal Design Representation</td>
</tr>
<tr>
<td>IP</td>
<td>Intellectual Property. Used in the field of embedded systems to refer to existing modules (from other vendors) that can be used to build a system</td>
</tr>
<tr>
<td>MoC</td>
<td>Model of Computation, or computational model</td>
</tr>
<tr>
<td>PLD</td>
<td>Programmable Logic Device</td>
</tr>
<tr>
<td>VHDL</td>
<td>A language to describe layout and or behaviour of hardware. Comparable to a program-language for software</td>
</tr>
</tbody>
</table>