Budapest University of Technology and Economics
Faculty of Electrical Engineering and Informatics
Department of Measurement and Information Systems

Formal modeling and verification of process models in component-based reactive systems

Scientific Students' Association Report

Author:

Ármin Zavada

Advisor:

dr. Vince Molnár
Bence Graics

Contents

Kivonat

A biztonságkritikus rendszerek komplexitása folyamatosan növekedett az elmúlt években. A komplexitás csökkentése érdekében a modellalapú paradigma vált a meghatározó módszerré ilyen rendszerek tervezéshez. Modellalapú rendszertervezés során a komponensek viselkedését általában állapotalapú, vagy folyamatorientált modellek segítségével írjuk le. Az előbbi formalizmusa azt írja le, hogy a komponens milyen állapotokban lehet, míg az utóbbié azt, hogy milyen lépéseket hajthat végre, valamint milyen sorrendben. Gyakran ezen modellek valamilyen kombinálása a legjobb módja egy komplex komponens viselkedésének leírásához.

Formális szemantikával rendelkező modellezési nyelvek lehetővé teszik a leírt viselkedés (kimerítő) verifikációját. Formális verifikáció használatával már a fejlesztés korai fázisaiban felfedezhetőek a hibák: a módszer ellenőrzi, hogy a rendszer egy adott (hibás) állapota elérhető-e, és amennyiben elérhető, ad hozzá egy elérési útvonalat. A formális verifikációs eszközök emiatt gyakran csak alacsony szintű, állapotalapú modelleken működnek, melyek messze vannak az emberek által könnyen érthető nyelvektől. Ezért, hogy magas szintű viselkedési modelleket tudjunk verifikálni, implementálnunk kell egy olyan modell transzformációt, mely megtartja a folyamat- és állapotalapú modellek szemantikáját azok kombinációja után is.

Ebben a dolgozatban megvizsgálom a folyamatalapú modellek szemantikáját, valamint a kapcsolatukat egyéb hagyományos állapotalapú modellekkel. Emellett megoldásokat vetek fel a potenciális konfliktusokra a kombinált alacsonyszintű modellben. Munkám során a Gamma állapotgép kompozíciós keretrendszerre építek, mellyel komponensalapú reaktív rendszereket modellezhetünk és verifikálhatunk. Mivel a Gamma még nem támogatja az aktivitásokat, bevezetek egy új aktivitás nyelvet, melyhez a SysMLv2 szolgál inspirációként. Ezzel együtt implementálom hozzá a szükséges transzformációkat a Gamma alacsony szintű analízis formalizmusára. Végezetül pedig kiértékelem a koncepcionális és gyakorlati eredményeket esettanulmányokon és méréseken keresztül, valamint felvetek lehetséges fejlesztéseket és alkalmazásokat.

Abstract

The complexity of safety-critical systems has been increasing rapidly in recent years. To mitigate said complexity, the model-based paradigm has become the decisive way to design such systems. In model-based systems engineering, we usually define the behaviour of system components using state-based or process-oriented models. The former formalism describes what states the component can be in, while the latter describes what steps it can perform and in what order. Oftentimes, the best way to model the behaviour of a complex component is to combine these models in some way.

Modelling languages with formal semantics enable the (exhaustive) verification of the described behaviour. Formal verification may be used to detect errors early during development by checking if a given (erroneous) state of the system can be reached, and if so, providing a way to reach it. Formal verification tools often require low-level state-based mathematical models, which are far from human-understandable languages. Thus, to enable the verification of high-level behavioural models, a model transformation must be implemented that preserves the semantics of both process-oriented and state-based models, even when combined.

In this report, I analyse the semantics of process-oriented models, as well as its relation to traditional state-based models, and propose solutions for the possible conflicts in a combined low-level model. In my work, I build on the Gamma Statechart Composition Framework, which is a tool for modelling and verifying component-based reactive systems based on statecharts. Since Gamma does not support activities yet, I introduce a new activity language inspired by SysMLv2, and implement the necessary transformations to Gamma's low-level analysis formalism. Finally, I evaluate the conceptual and practical results through case studies and measurements then propose potential improvements and applications.

Chapter 1

Introduction

Add introduction

Chapter 2

Background

In this chapter I present the foundations of this work. In Section 2.1 I introduce the concept of model-based systems engineering, which is a well known approach for complex system design. This section also goes into detail about SysML as well (Section 2.1.1), which is a widely used *systems modeling language*. Next, I talk about the concept of formal verification in Section 2.2, and introduce Petri nets (Section 2.2.2) and a mapping between activities and Petri nets (Section 2.2.3). Lastly, I introduce the Gamma Statechart Composition Framework, which is a tool for modeling and verifying composite systems using statechart behaviours Section 2.3, and a low-level formalism called XSTS (Section 2.4), used as an intermediary language for model checking by Gamma.

2.1 Model-based Systems Engineering

The INCOSE SE Vision 2020 [1] defines Model-based systems engineering (MBSE) as:

> the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases. MBSE is part of a long-term trend toward model-centric approaches adopted by other engineering disciplines, including mechanical, electrical and software. In particular, MBSE is expected to replace the document-centric approach that has been practiced by systems engineers in the past and to influence the future practice of systems engineering by being fully integrated into the definition of systems engineering processes.

Applying MBSE is expected to provide significant benefits over the document centric approach by enhancing productivity and quality, reducing risk, and providing improved communications among the system development team [2].

In MBSE, one of the most important concepts is the term „model" itself. Literature gives various definitions for models:

1. A physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process.[5]
2. A representation of one or more concepts that may be realized in the physical world [7].

3. A simplified representation of a system at some particular point in time or space intended to promote understanding of the real system [4].

4. An abstraction of a system, aimed at understanding, communicating, explaining, or designing aspects of interest of that system [6].

5. A selective representation of some system whose form and content are chosen based on a specific set of concerns. The model is related to the system by an explicit or implicit mapping [10].

As one can see, choosing a definition is very much dependent upon how we wish to use our „model"; in this work I will use the 1st definition.

2.1.1 Systems Modeling Language

Systems Modeling Language (OMG SysML [9]) is a general-purpose modeling language that supports the specification, design, analysis, and verification of systems that may include harware and equipment, software, data, personnel, procedures, and facilities. SysML is a graphical modeling language with a semantic foundation for representing requirements, behaviour, structure, and properties of the system and its components [7].

This work focuses only on the *behavioural* modeling tools SysML provides. In the following section, I present two of the most used concepts; State Machines and Activity Diagrams.

2.1.1.1 State Machine

Reactive systems are all around us in our daily lives; in smart phones, avionics systems or event our calculators. Frequently, reactive systems appear in areas, where safety-critical operation is crucial, as even the slightest misbehaviour can have catastrophic consequences. This makes the verification of these systems a must during their design process.

The defining characteristic of reactive systems is their event-driven nature, which means that they continuously receive external stimuli (*events*), based on which they change their internal *state* and possibly react with some output [13]. Reactive systems can be verified using model checking techniques (see Section 2.2). Statecharts [14] are a popular and intuitive language to capture the behaviour of reactive systems [16, 24], while also being formally defined.

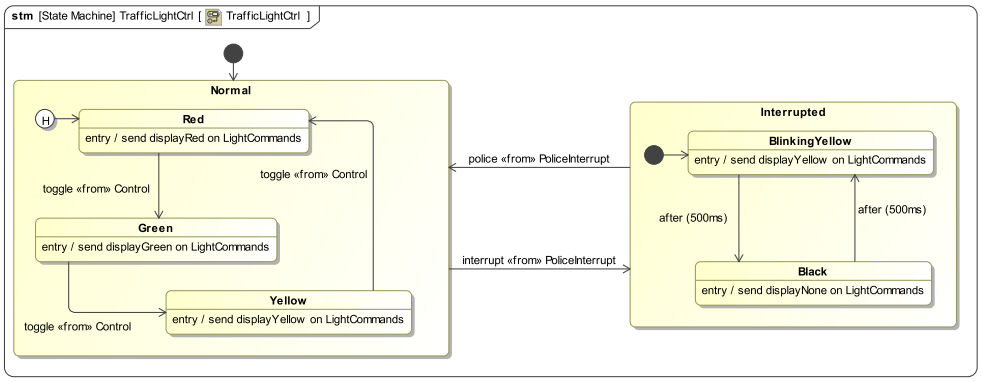


Figure 2.1: SysML State Machine describing the behaviour of a traffic light controller.

SysML state machines extend the concept of statecharts with hierarchical state-refinement, orthogonal regions, action-effect behaviour and state machine composition. These advanced abstraction constructs make state machines easy to use for engineers, but hinder their formal verification. This abstraction gap can be bridged using a transformation tool, such as Gamma, of which I will be talking in Section 2.3.

Figure 2.1 shows a state machine modeling the behaviour of a traffic light controller.

explain controller behaviour in words

2.1.1.2 Activity Diagram

State machines cannot describe the complicated semantics of distributed systems with concurrent, parallel behaviour, where the *interesting* thing is what the system *does* step-by-step. SysML activity diagrams are a primary representation for modeling process based behaviour [9] for distributed, concurrent systems. Figure 2.2 shows the set of interesting modeling elements used in this work. In the following, I will introduce the different artifacts and show an example of a SysML activity diagram.

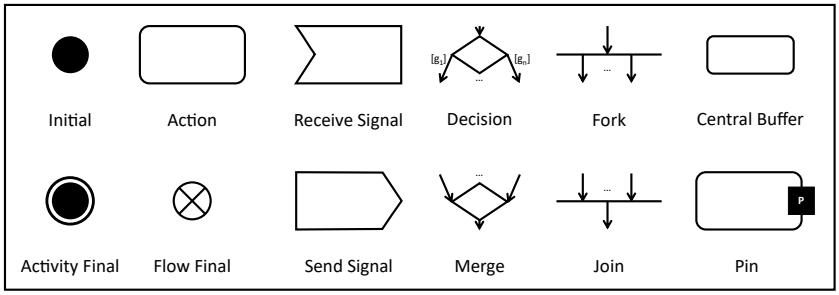


Figure 2.2: Artifacts of SysML activity diagrams.

SysML activity diagram is a graph based model, where the nodes are connected via flows. The dynamic behaviour of activity diagrams comes from *tokens* travelling from node to node; based on the given node's semantics, a connected flow removes tokens from the source node and puts them onto the target node. Flows can also have *guards*, which are expressions specifying when the given flow is *enabled* or not; only enabled nodes can transfer tokens.

Tokens are a way of creating limitations over which node can run, and which cannot; a given node is considered *running*, only when it contains a token. Tokens *flow* between nodes, carrying with them a given value - this value can be of type *void*, which makes it a *control* token.

The different nodes represent the different semantic „tools" at our disposal; they can represent different actions, or introduce interesting token flow semantics. Simple **actions** represent a single step of behaviour that convert a set of inputs to a set of outputs. Both inputs and outputs are specified as pins, which get their data from connected flows - making that flow a *data flow*. The flow starts from the initial node, and ends with a *Flow Final* or *Activity Final* node. *Fork* nodes generate tokens on all of their output flows, and *Join* nodes forward the tokens, only when all input flows contain one - thus, the two nodes come in a *pair*. On the other hand, *Merge* nodes do not wait for all flows, they forward any token they receive instantly. Likewise, *Decision* nodes take one token from its input flows, and sends it out on its single enabled output flow.

The detailed specification for SysML Activity Diagrams can be found in the OMG specification [9].

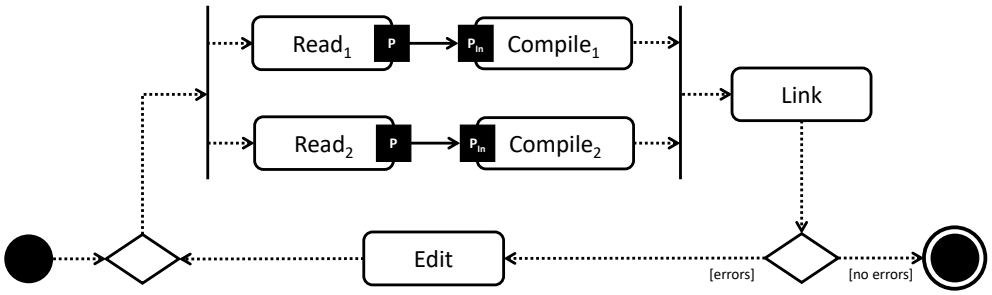


Figure 2.3: The activity of editing, compiling and linking two files.

Figure 2.3 shows an example activity diagram, modeling the process of editing, compiling and linking two files. First, the files have to be read, after which they are *transfered* to the compiler module. We want to compile the two different files in *parallel*, thus we split the control flow using a *fork* node. Once the files are compiled, we link both of them - since linking requires both files, it is preceded by a *join* node. Finally, if the resulting code contains errors, we *edit* the source files and start over - otherwise we are done.

2.2 Formal Verification

> In order to raise the reliability of system analysis, a system analysis technique is required that can have the precision of paper-and-pencil based mathematical proofs, and thus does not rely upon computer-arithmetic, and utilizes the computers for bookkeeping, to be able to handle complex systems without having to worry about human-errors. Formal verification methods, which are primarily based on theoretical computer science fundamentals like logic calculi, automata theory and strongly type systems, fulfil these requirements. The main principle behind formal analysis of a system is to construct a computer based mathematical model of the given system and formally verify, within a computer, that this model meets rigorous specifications of intended behaviour. Due to the mathematical nature of the analysis, 100% accuracy can be guaranteed. [15]

2.2.1 Model Checking

Model checking is a formal verification method to verify properties of finite systems, i.e., to decide whether a given formal model M satisfies a given requirement γ or not. The name comes from formal logic, where a logical formula may have zero or more models, which define the interpretation of the symbols used in the formula and the base set such that the formula is true. In this sense, the question is whether the formal model is indeed a model of the formal requirement: $M \models \gamma$?

Model checker algorithms (see Figure 2.4), such as the ones used in UPPAAL[1] [3] or Theta[2] [25] can answer this question, and can even return a *proof* (i.e., a part of the model) that M indeed does satisfy said requirement[3].

[1] https://uppaal.org/
[2] https://inf.mit.bme.hu/en/theta
[3] These proofs usually come in the form of an execution trace

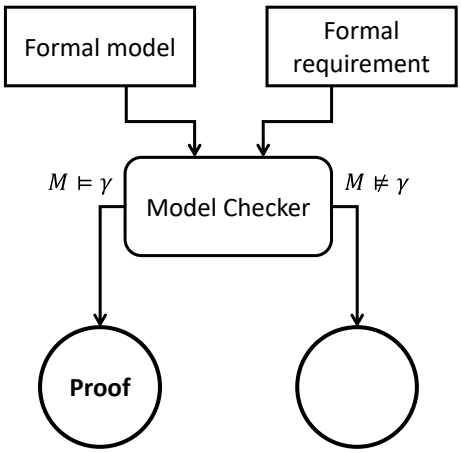


Figure 2.4: An illustration of model checking.

2.2.2 Petri Nets

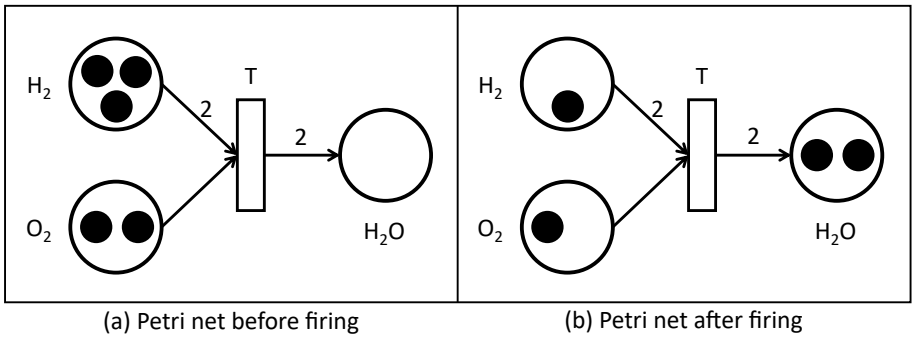


Figure 2.5: An example Petri net modeling the process of H_2O molecule creation.

Petri nets are a widely used formalism to model concurrent, asynchronous systems [22]. The formal definition of a Petri net (including inhibitor arcs) is as follows (see Figure 2.5 for an illustration of the notations).

Definition 1 (Petri net). A Petri net is a tuple $PN = (P, T, W, M_0)$

- P is the set of *places* (defining state variables);
- T is the set of *transitions* (defining behaviour), such that $P \bigcap T = \emptyset$;
- $W \subseteq W^+ \bigcup W^-$ is a set of two types of arcs, where $W^+ : T \times P \to \mathbb{N}$ and $W^- : P \times T \to \mathbb{N}$ are the set of input arcs and output arcs, respectively ($\mathbb{N}$ is the set of all natural numbers);
- $M_0 : P \to \mathbb{N}$ is the *initial marking*, i.e., the number of *tokens* on each place. ▪

The state of a Petri net is defined by the current marking $M : P \to \mathbb{N}$. The behaviour of the systems is described as follows. A transition t is enabled if $\forall p \in P : M(p) \in W(p, t)$. Any enabled transition t may fire nondeterministically, creating the new marking M' of the Petri as follows: $\forall p \in P : M'(p) = M(p) - W^-(p, t) + W^+(t, p)$.

In words: W describes the *weight* of each flow from a transition to a place, or from a place to a transition. Firing a transition t in a marking M consumes $W^-(p_i, t)$ tokens from each

of its input places p_i, and produces $W(t, p_o)$ tokens in each of its output places p_o. One such transition t is *enabled* (it may *fire*) in M if there are enough tokens in its input places for the consumptions to be possible, i.e., if and only if $\forall p : M(p) \geq W(s, t)$.

2.2.3 Activities as Petri Nets

Formal verification requires models to be specified using *mathematical* precision, however SysML does not have precise semantics [21, 23, 18]. Huang et al. in [17] propose a way to *partially* map SysML activity diagrams to Petri nets. In the following, I will summarise their work.

2.2.3.1 Constrained Subset of SysML

Since SysML Activity Diagrams do not have precise execution semantics, the Petri net mapping can only be done for a limited subset of the modeling elements: *actions*, *initial nodes*, *final nodes*, *join nodes*, *fork nodes*, *merge nodes*, *decision nodes*, *pins* and *object/control flows*, which have precise execution semantics as defined in the Foundational Subset for Executable UML Models [11].

The paper also assumes the following constrains:

1. The value of tokens is not considered.
2. Control flows with multiple tokens at a time are not considered.
3. Optional object/control flows are not considered, i.e., multiplicity lower bounds are strictly positive.

These constraints allow the mapping between activities and Petri nets, however, the constructed Petri net will not be semantically equivalent - data cannot flow between nodes. This fact is the motivation behind formalising a more-complete mapping (see Chapter 3).

2.2.3.2 Mapping Rules

More description

Activity elements can be grouped into two sets: *load-and-send* (LAS) and *immediate-repeat* (IR).

LAS nodes are fired when all their inputs have tokens. When an LAS node fires, the number of tokens associated with the input arcs/pin is consumed and the number of tokens associated with an output arc/pin is added. In SysML activity diagrams, the execution semantics of fork nodes, join nodes, and basic actions are also LAS because these nodes are fired when all their input nodes have at least one token. As a result, these nodes can be mapped to *transitions* in the resulting Petri net.

In contrast, as soon as an IR node receives a token from any input, it immediately adds a token to its output nodes. For UML/SysML activity diagrams, activity final nodes, merge nodes, decision nodes and pins are IR nodes because they are fired immediately when any token is received. As a result, these nodes can be mapped to *places* in the resulting Petri net.

Given the set of assumptions in Section 2.2.3.1, control flows and object flows in an activity diagram can be mapped to arcs in a Petri net.

Finally, after mapping the elements, the resulting Petri net may contain transition-transition and place-place arcs, which are not valid; as final step, these arcs have to be split in two by inserting a transition or a place in the middle, making the model conform to the formalism.

2.2.3.3 Example Mapping

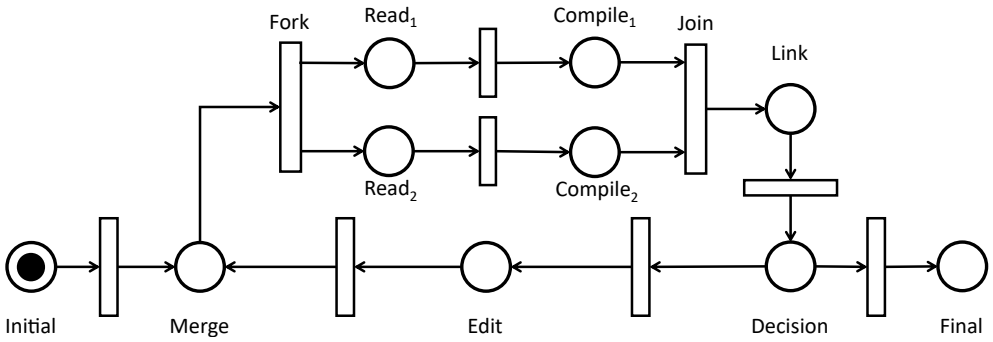


Figure 2.6: Example mapping from activity diagram to Petri net.

Figure 2.6 shows an example mapping from the activity Figure 2.3. The resulting elements are annotated with the names of their counter parts in the activity diagram.

For more information, please refer to [17]!

2.3 The Gamma Statechart Composition Framework

The Gamma Statechart Composition Framework[4] [8] is an integrated tool to support the design, verification and validation as well as code generation for component-based reactive systems. The behaviour of each component is captured by a statechart, while assembling the system from components is driven by a domain-specific composition language[5]. Gamma supports formal verification by mapping composite statecharts to a back-end model checker. Execution traces obtained as witnesses during verification are back-annotated as test cases to replay an error trace or to validate external code generators [19].

The workflow of Gamma builds on a model transformation chain depicted in Figure 2.7, which illustrates the input and output models of these model transformations as well as the languages in which they are defined, and the relations between them. The modeling languages are as follows.

- The **Gamma Statechart Language (GSL)** is a UML/SysML-based statechart language supporting different semantic variants of statecharts.
- The **Gamma Composition Language (GCL)** is a composition language for the formal hierarchical composition of state-based components according to multiple execution and interaction semantics.
- The **Gamma Genmodel Language (GGL)** is a configuration language for configuring model transformations.

[4] https://inf.mit.bme.hu/en/gamma

[5] The composition language be the *ibd* model in SysML

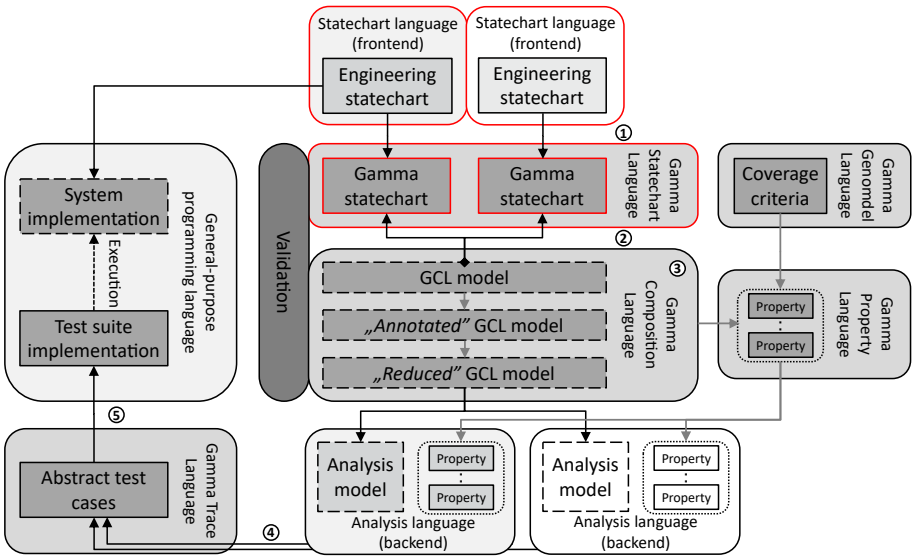


Figure 2.7: The overview of model transformation chains and modeling languages of the Gamma framework [8]. The parts relevant to this work have been marked with red outline.

- The **Gamma Property Language (GPL)** is a property language supporting the definition (CTL*) properties and thus, the formal specification of requirements regarding (composite) component behavior.
- The **Gamma Trace Language (GTL)** is a high-level specification language for execution traces of (composite) components.

Optionally, statechart models defined in supported modeling tools (front-ends) can be imported into Gamma (Step 1), which can be integrated according to well-defined execution and interaction semantics (Step 2). The resulting composite model is processed and transformed into the input formalisms of integrated model checker back-ends (Step 3). The model checker back-ends provide witnesses (diagnostic traces) based on specified properties, which are back-annotated, resulting in abstract traces (Step 4). Finally, the abstract traces are mapped into concrete (executable) traces tailored to the targeted execution environment (Step 5). For a more detailed description, see [8].

2.3.1 Example Statechart

Listing 2.1 shows the Gamma Statechart representation of the State Machine introduced in Figure 2.1.

more description

2.4 Extended Symbolic Transition System

The high-level nature of engineering models means they are easy-to-use for engineers, but leads to difficulties during the formal verification process. SysML state machines and acitivty diagrams for example contain high-level modeling constructs that make the modeling workflow more intuitive and enable the modeling of complex systems, however they are difficult to process using formal methods that are defined on low-level mathematical formalism. In this section, I introduce the Extended Symbolic Transition System [20]

```
package TrafficLightCtrl
import "Interfaces"
statechart TrafficLightCtrl [
    port Control : requires Control
    port PoliceInterrupt : requires PoliceInterrupt
    port LightCommands : provides LightCommands
] {
    timeout BlinkingYellowTimeout3
    timeout BlackTimeout4
    transition from Yellow to Red when Control.toggle
    transition from Normal to Interrupted when PoliceInterrupt.police
    // ...
    transition from BlinkingYellow to Black when timeout BlinkingYellowTimeout3
    transition from Black to BlinkingYellow when timeout BlackTimeout4
    region main_region {
        state Normal {
            region normal {
                shallow history Entry2
                state Green {
                    entry / raise LightCommands.displayGreen;
                }
                state Red {
                    entry / raise LightCommands.displayRed;
                }
                state Yellow {
                    entry / raise LightCommands.displayYellow;
                }
            }
        }
        state Interrupted {
            region interrupted {
                initial Entry1
                state Black {
                    entry / set BlackTimeout4 := 500 ms; raise LightCommands.displayNone;
                }
                state BlinkingYellow {
                    entry / set BlinkingYellowTimeout3 := 500 ms; raise LightCommands.
    displayYellow;
                }
            }
        }
        initial Entry0
    }
}
```

Listing 2.1: The traffic light controller state machine in the Gamma textual representation.

language, which is a low-level modeling formalism designed to bridge the abstraction gap between engineering models and formal methods.

2.4.1 Formal definition

Definition 2 (Extended symbolic transition system). An *Extended symbolic transition system* is a tuple $XSTS = (D, V, V_C, IV, Tr, In, En)$, where:

- $D = \{D_{v_1}, D_{v_2}, \ldots, D_{v_n}\}$ is a set of value domains;
- $V = \{v_1, v_2, \ldots, v_n\}$ is a set of variables with domains $D_{v_1}, D_{v_2}, \ldots, D_{v_n}$;
- $V_C \subseteq V$ is a set of variables marked as *control variables*;
- $IV \in D_{v_1} \times D_{v_2} \times \cdots \times D_{v_n}$ is the *initial value function* used to describe the initial state. The initial value function IV assigns an initial value $IV(v) \in D_v$ to variables $v \in V$ of their domain D_v;
- $Tr \subseteq Ops$ is a set of operations, representing the *internal transition relation*; it describes the internal behaviour of the system;
- $In \subseteq Ops$ is a set of operations, representing the *initialisation transition relation*; it is used to describe more complex initialisation, and is executed once and only once, at the very beginning;
- $En \subseteq Ops$ is a set of operations, representing the *environmental transition relation*; it is used to model the system's interactions with its environment. ▪

In any state of the system, a single operation is selected from the sets introduced above (Tr, In and En). The set from where the operation can be selected depends on the current state: In the initial state - which is described by the initialization vector IV - only operations from the In set can be executed. Operations from the In set can only fire in the initial state and nowhere else. After that, En and Tr are selected in an alternating manner.

Operations $op \in Ops$ describe the transitions between states of the system, where Ops is the set of all possible transitions. All operations are atomic in the sense that they are either executed in their entirety or none at all. XSTS defines the following operations:

Basic operations Basic operations contain no inner (nested) operations.

- *Assignments* assign a given value v from domain D_n to variable V_n.
- *Havocs* behave likewise, except the value is not predetermined; giving a way to assign random value to a variable.
- Lastly, *assume* operations check a condition, and can only be executed if their condition evaluates to *true*.

Composite operations Composite operations contain other operations, and can be used to describe complex control stuctures.

- *Sequences* are essentially multiple operations executed one after the other.
- *Parallels* are like sequences, however they execute all operations at the same time.
- And lastly, *choices* model non-deterministic choices between multiple operations; one and only one branch of the choice operation is selected for execution.

Note that while these are composite operations, their execution is still atomic; i.e., a potential false evaluation of a containing assume operation prevents the execution of all operations in that particular branch. E.g., if a sequence's 2nd operation is an assume operation, which cannot execute, then the whole sequence operation will be prevented from execution.

2.4.2 Traffic Light Controller Example

Listing 2.2 is the textual representation of the traffic light controller statechart (Figure 2.1) transformed to XSTS[6]. For a more exact presentation please see Section A.1.

2.5 Related Work

Add related work

[6]The transformation was done using Gamma.

```
type ActivityNodeState : { __Idle__, __Running__, __Done__ }
// ...
type Operating_Controller : { __Inactive__, Priority, Init, PriorityPrepares, Secondary,
     SecondaryPrepares }
var PoliceInterrupt_police_In_Controller : boolean = false
// ...
ctrl var main_region_Controller : Main_region_Controller = __Inactive__
ctrl var operating_Controller : Operating_Controller = __Inactive__
var SecondaryTimeout2_Controller : integer = 0

trans {
    // ...
    choice {
        assume (((main_region_Controller == Interrupted) &&
        // ...
        choice {
            assume (operating_Controller == Priority);
            SecondaryTimeout2_Controller := 0;
            PriorityControl_toggle_Out_Controller := true;
            SecondaryControl_toggle_Out_Controller := true;
            // ...
        }
    } or {
        // ...
    }
    PoliceInterrupt_police_In_Controller := false;
}
init {
    SecondaryTimeout2_Controller := (2 * 1000);
    main_region_Controller := __Inactive__;
    // ...
    PriorityPolice_police_Out_Controller := false;
    choice {
        assume (operating_Controller == __Inactive__);
        operating_Controller := Init;
    } or {
        assume (!((operating_Controller == __Inactive__)));
    }
    // ...
}
env {
    havoc PoliceInterrupt_police_In_Controller;
    PriorityPolice_police_Out_Controller := false;
    SecondaryControl_toggle_Out_Controller := false;
    PriorityControl_toggle_Out_Controller := false;
    SecondaryPolice_police_Out_Controller := false;
}
```

Listing 2.2: Gamma XSTS Language representing the traffic light controller statechart.

Chapter 3

Gamma Activity Language

The high-level nature of SysML activities means they are easy to model complicated behaviours of distributed system, but leads to an increasingly hard time trying to run formal verification on them. As discussed above in Section 2.2.3, we can define a mapping between activities and Petri nets (which have exact semantics), however, that mapping is not exhaustive, as it disregards the data contained in tokens.

The pre-existing Gamma Statechart Composition Framework (Section 2.3) implements this transformation-verification pipeline, however, it does not include activity diagrams. For this reason, I propose (Chapter 3) the *Gamma Activity Language*, and integrate it (Chapter 4) into the pipeline shown in Figure 2.7.

3.1 Language Design

Every language starts with defining exactly what we want our language to accomplish. With the Activity Language, this was the following: it should support as much features from SysML activity diagrams as possible, all the while being simple to implement and transform to low-level models. As you could see from the Section 2.2.3, SysML activity diagrams cannot be exactly transformed to a formal model; we have to constrain our language in order to offer formal semantics. However, as you will see, these constrains still allow us to write *almost* all activities.

3.1.1 Constrains

Compared to SysML activity diagrams, Activity Language has the following constrains:

- No *flow final* nodes.
- No *receive signal* nodes.
- No *interrupt* flows.
- Flows may only contain *one* token at a given time.
- Activities can only be inside *statecharts*, cannot describe a component's behaviour in itself

As you can see, most of these constrains can be substituted with other constructs, and just serve as *shorthands*. The absence of *flow final* nodes can be resolved by merging all flows into one single *activity final* node. These steps could be done automatically using model transformation from SysML to Activity.

3.1.2 Improvement over Petri net mapping

Compared to the Petri net mapping introduced in Section 2.2.3, the Activity Language supports guards, valued tokens, actions and composite activities.

3.2 Formal Definition

Finalise formalism

In order to offer mathematical precision, formal verification methods require formally defined models with clear semantics. In this section I present the formal definition of the novel GAL formalism.

Definition 3 (Gamma Activity Language). A Gamma Activity is a tuple of $GA = (N, P, F, G, D, F_{Action})$, where:

- $N = N_{Initial} \bigcup N_{Final} \bigcup N_{Pseudo} \bigcup N_{Decision} \bigcup N_{Merge} \bigcup N_{Action}$ is a set of *Nodes*, where $N_{Initial}$ contains the *Initial* nodes, N_{Final} contains the *Final* nodes, N_{Pseudo} contains the *Pseudo* nodes, $N_{Decision}$ contains the *Decision* nodes, N_{Merge} contains the *Merge* nodes and N_{Action} contains the *Action* nodes;
- $P \subseteq P_{In} \bigcup P_{Out}$ is a set of two types of pins, where $P_{In} : N_{Action} \to \{p_1, \ldots, p_n\}$ and $P_{Out} : N_{Action} \to \{p_1, \ldots, p_n\}$ are the set of *InputPins* and *OutputPins*, respectively, with domains $\{d_1, \ldots, d_n\} \subseteq D$;
- $F \subseteq F_C \bigcup F_D$, where $F_C = \{f_{C_1}, \ldots, f_{C_n}\}$ and $F_D = \{f_{D_1}, \ldots, f_{D_n}\}$ are the control and data flows, respectively. Let us denote the input/output flows of node n as $\delta(n)$ and $\Delta(n)$, and the source/target pins of flow f as $\phi(f)$ and $\Phi(f)$. For any given node n, $\delta(n) \neq \Delta(n)$ and for any given action node n_a $\Phi(f) \in P_{In}(n_a) \forall f \in F_D \bigcap \delta(n_a)$ and $\phi(f) \in P_{Out}(n_a) \forall f \in F_D \bigcap \Delta(n_a)$ shall always hold. This means, that a flow cannot be input and output to the same node at the same time, and for a given action node, all input/output flows shall be associated with an input/output token, respectively;
- $G : F \to \{True, False\}$ is a function determining whether a given flow is enabled;
- $D = \{d_1, d_2, \ldots, d_n\}$ is a set of value domains;
- $F_{Action} : N \to D$ is a function running the contained action, and returning it's result values. ▪

Informally, Gamma Activities are composed of *nodes* (*Initial/Final*, *Decision/Merge*, *Action* and *Pseudo*) and flows (*Control* and *Data*) in between them. A given action can also have any number of *Pins* with a domain, for which, there must be one and only one *data* flow connected to the node. An action node also has a special F_{Action}, which implements its specific behaviour, and returns the values of its output pins.

For the most part, these elements have a similar behaviour as their SysML counter parts; however, there is a crucial difference regarding how a flow transmits a token. Figure 3.1 shows a simple data flow between two nodes, connected via their pins. When this flow is fired, the data inside the pins are transferred instantly, without any pseudo-state in between.

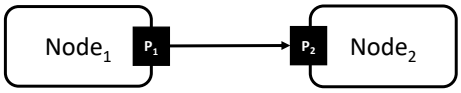


Figure 3.1: An example SysML data flow

Contrarily, in Gamma Activity Language, the equivalent data flow does indeed contain the given node, creating an intermediate state, where the token is in neither of the nodes. This

would look like a *Central Buffer* in SysML (Figure 3.2). The reason for this is simplicity; by creating this intermediate state (and others), it is easier to state the set of transitions necessary to formally define the semantics of the language.

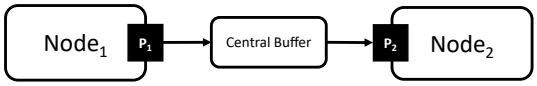


Figure 3.2: An example SysML data flow with central buffer

Definition 4 (Gamma Activity State). The state of a GA is defined the following:

- $S_N : N \to \{Idle, Running, Done\}$ is a function returning the *state* of a node;
- $S_F : F \to \{Empty, Full\}$ is a function returning the *state* of a flow;
- $V_N : N \to D$ is a function returning the current *value* contained in a node;
- $PV_N : N \times P \to D$ is a function returning the current *value* contained in a node's pin;
- $V_F : F \to D$ is a function returning the current *value* contained in a flow. ▪

Informally, the state of Gamma Activities are determined by the nodes' *state* (*Idle*, *Running*, *Done*) and their (or their pins') values, the flows' *state* (*Empty*, *Full*) and their values. For example, given an action node n and one of it's pins p_1, $PV_N(n, p)$ would give us the exact value that pin contains in this instance. This gives us the power - contrary to the Activity-PN mapping introduced in Section 2.2.3 - to formally define the values contained in tokens and nodes.

Definition 5 (State Transition Functions). The behaviour of the system is described by three functions (F_{In}, F_{Run} and F_{In}). At any one time, a node $n \in N$ is selected, and an *enabled* function is executed non-deterministically. These functions define the state-transitions for our system:

$$F_{In} : N \to \begin{cases} \text{select } f_s f_s \in \delta(n) \wedge S_F(f) = Full & \\ V'_N(n) = V_F(f_s) & \\ S'_F(f_s) = Empty & \\ S'_N(n) = Running, & \text{if } n \in N_{Merge} \\ PV'_N(n, \Phi(f)) = V_F(f) : \forall f \in \delta(n) \bigcap F_D & \\ S'_F(f) = Empty : \forall f \in \delta(n) & \\ S'_N(n) = Running, & \text{if } n \in N_{Action} \\ V'_N(n) = V_F(f) : f \in \delta(n) \bigcap F_D & \\ S'_F(f) = Empty : \forall f \in \delta(n) & \\ S'_N(n) = Running, & \text{otherwise} \end{cases} \tag{3.1}$$

In words, the function F_{In} takes the tokens from the input flows, and puts it in the given node, thus starting it's execution. The rule how it selects when is enabled, and which flows to empty is determined by the node's type:

- if it is a *Merge* node, then it must only forward one, and only one token,
- if it is an *Action* node, then it must forward all tokens, along with their values into the associated pins,
- otherwise, it must forward all tokens.

This way we can ensure the behaviours of the simple 'LAS' nodes, as well as the 'IR' nodes.

$$F_{Run} : N \rightarrow \begin{cases} V'_N(n) = F_{Action}(n, V_N(n)) & \\ S'_N(n) = Done, & \text{if } n \in N_{Action} \\ S'_N(n) = Done, & \text{otherwise} \end{cases} \tag{3.2}$$

In words, the function F_{Run} stochastically sets the given node's state to *Done*. If the given state is of type *Action*, then it must have a deeper semantic associated with it, thus the associated F_{Action} function must also be called, which implements it's specific behaviour (more detail below).

$$F_{Out} : N \rightarrow \begin{cases} \text{select } f_s \in \Delta(n) \wedge S_F(f_s) = Empty \wedge G(f_s) = True & \\ V'_F(f_s) = V_N(n) & \\ S'_F(f_s) = Full & \\ S'_N(n) = Idle, & \text{if } n \in N_{Decision} \\ V'_F(f) = PV_N(n, \phi(f)) : \forall f \in \Delta(n) \bigcap F_D & \\ S'_F(f) = Full : \forall f \in \Delta(n) & \\ S'_N(n) = Idle, & \text{if } n \in N_{Action} \\ S'_F(f) = Full : \forall f \in \Delta(n) & \\ S'_N(n) = Idle, & \text{otherwise} \end{cases} \tag{3.3}$$

The function F_{Our} takes the token from node, and puts it in its output flows, thus ending it's execution. The rule how it selects when is enabled, and which flows to fill is determined by the node's type:

- if it is a *Decision* node, then it must only forward the token on one of its *enabled* flows,
- if it is an *Action* node, then it must forward all tokens, along with their values from the associated pins,
- otherwise, it must forward all tokens. .

Figure 3.3 depicts these three functions, and how they change the state of a given node.

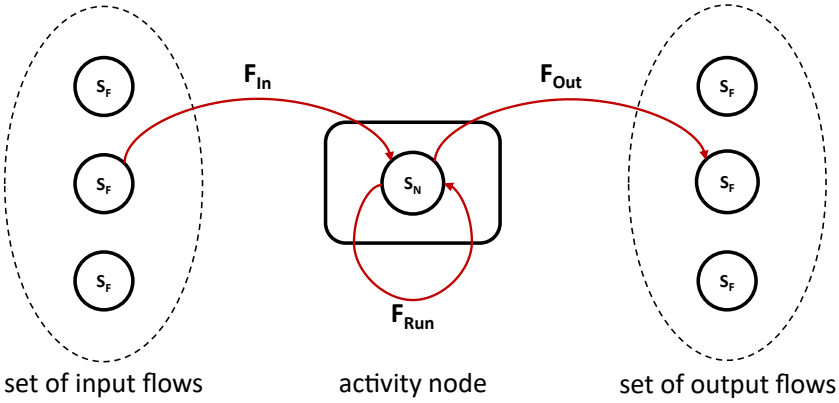


Figure 3.3: An illustration of activity node state and change functions

3.3 Domain-specific Language

Like the Gamma Statechart Language, the Gamma Activity Language is intended to be a first-class citizen in the Gamma Framework, thus it must have a domain-specific language to represent in a textual way. This implementation draws highly from the SysMLv2 [12] language design, while also fitting into the already existing language family of Gamma (Section 2.3).

3.3.1 Meta-model

Due to the complexity of the final meta-model of the language, I have split it into multiple parts for easier understanding.

3.3.1.1 Pins

There are two kinds of pins, *InputPins* and *OutputPins*. All pins have a *Type* associated with them from the Gamma grammar, which defines it's domain. Figure 3.4 shows this section of the meta-model.

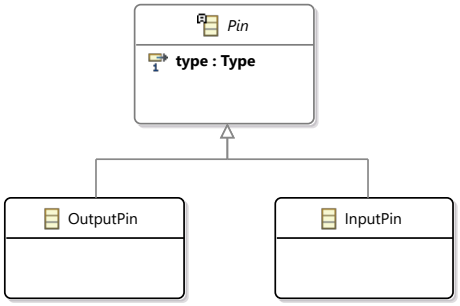


Figure 3.4: The Pins structure

3.3.1.2 Flows

There are two kinds of flows, *ControlFlows* and *DataFlows*. All flows have a guard of type *Expression*, which can be evaluated to a *boolean* value. If this guard evaluates to *True*, the flow is considered *enabled*. Figure 3.5 depicts the relation of flows in the meta-model.

Control Flow *ControlFlows* have a reference to their source and target nodes.

Data Flow *DataFlows* contain a *DataSourceReference* and a *DataTargetReference*. The reason for the reference, is because said reference can either be a *Pin*, or a *DataNode*. A data token may contain data (or value) of any kind, but that token can only travel to and from data *sources* and *targets*. This will be explained in more detail below.

3.3.1.3 Activity Nodes

In the following, I will talk about the different kinds of *ActivityNodes* and their special meanings. The meta-model described in this section can be seen in Figure 3.6.

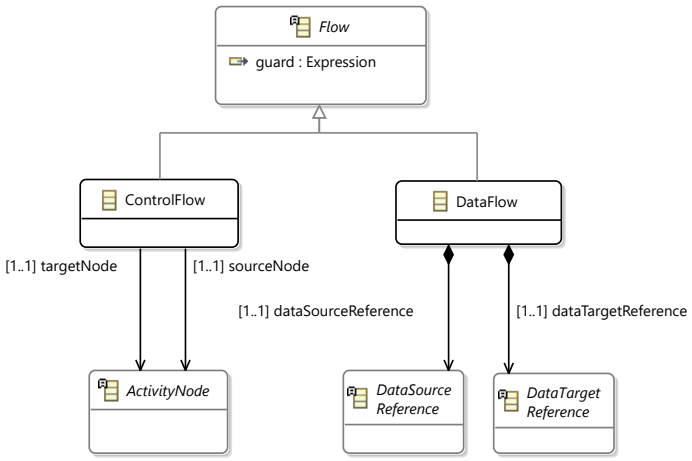


Figure 3.5: The Flows structure

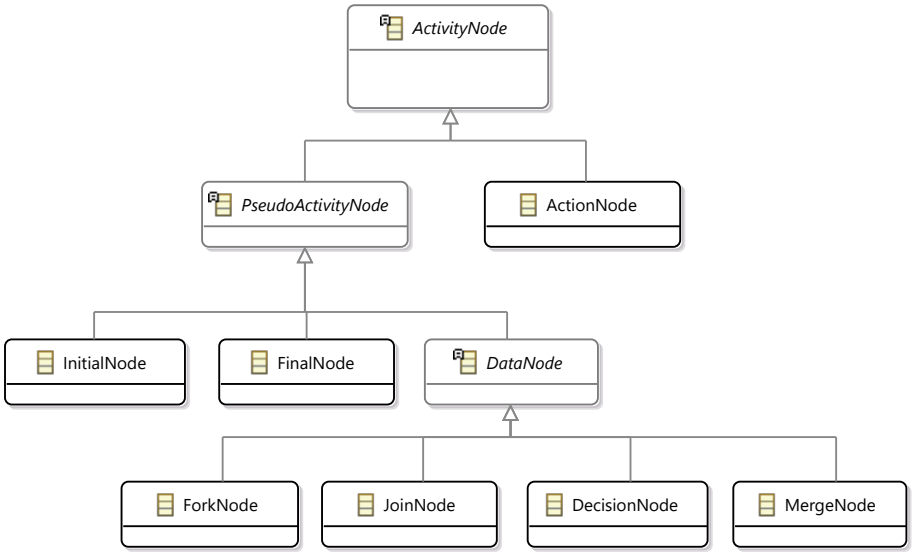


Figure 3.6: The Activity Node structure

Action Node *ActionNodes* represent a specific action the activity may execute. This action can be defined in multiple ways, see Section 3.3.1.5 for more detail.

Pseudo Activity Node *PseudoActivityNodes* are nodes, that do not represent a specific action, however are needed to convey specific meanings, e.g., the initial active node, or a decision between flows.

Initial Node *InitialNodes* represent the entry-point to the activity.

Final Node A special node representing the final node for the activity.

Data Node *DataNodes* encapsulate the meaning of *data* inside activities.

Fork Node *ForkNodes* are used to model parallelism, by creating one token on each of its output flows when executed. Fork nodes shall only have one input flow.

Join Node *JoinNodes* are the pair of fork nodes; the additional created tokens are swallowed by this node, by only sending out one token, regardless of the number of input flows.

Decision Node *DecisionNodes* create branches across multiple output flows. An input flow's token is removed, and sent out to a single output flow - depending on which of the output flows are *enabled* (see Section 3.3.1.2).

Merge Node *MergeNodes* forward all incoming tokens as soon as they arrive, one by one. They are used to *merge* different flow paths (created using *decisions*).

3.3.1.4 Root Structure

Every model has to have a root element structure; Activity Language is not any different. The meta-model described in this section can be seen in Figure 3.7.

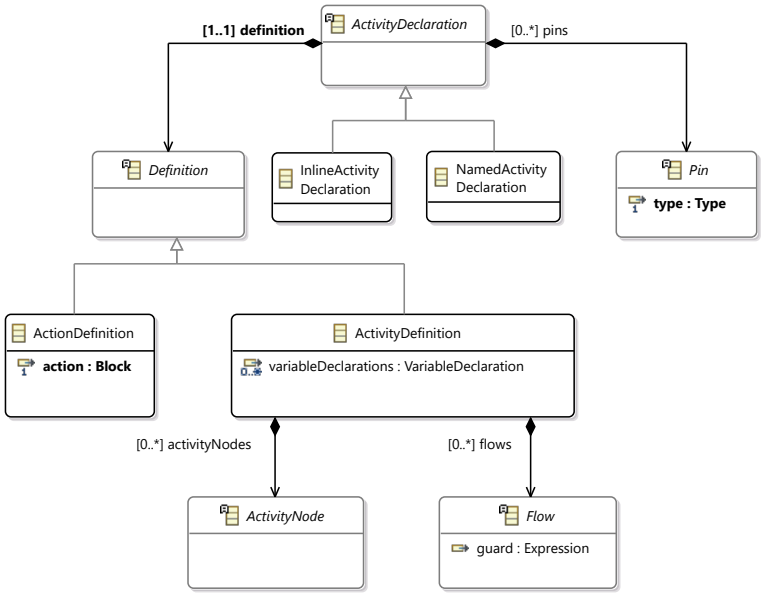


Figure 3.7: The root structure of the language

Activity Declaration All elements inside an activity are contained in a root *Activity-Declaration* element. It contains *Pins* needed for value passing (see Section 3.3.1.1) and a *Definition*. A declaration can be *InlineActivityDeclaration*, which means they are declared in an other declaration, or *NamedActivityDeclaration*, which is a standalone activity declaration. The difference will be clarified in **??**.

Definition The definition *defines* how the activity is described; using activity nodes, or by the Gamma Activity Language[1]. *ActionDefinition* contains a single *Block*[2], which is

[1] Gamma Activity Language is a lightweight programming language-like construct for writing simple algorithms

[2] A *Block* contains multiple *Actions* which are executed one after the other

executed as-is when the activity is executed[3]. *ActivityDefinition* contains *ActivityNodes* (Section 3.3.1.3) and *Flows* (Section 3.3.1.2).

3.3.1.5 Composing Activities

When one defines an *ActionNode*, it may or may not add an *ActivityDeclarationReference*. If the node does not have any, it is considered a *simple* node, without any implementation. If it does contain an *ActivityDeclarationReference*, then when this node is executed, the underlying *ActivityDeclaration* is executed as well. This gives us the power to pre-declare specific activities, or inline declare them in the language; and use any definition defined above (Section 3.3.1.4). Figure 3.8 shows this part of the meta-model.

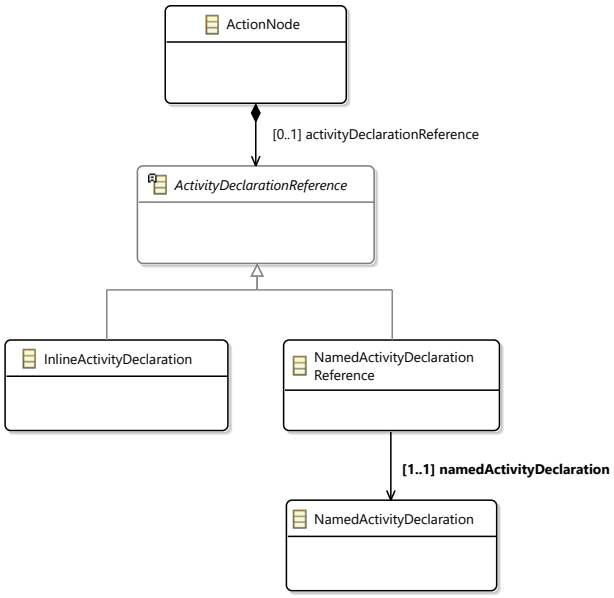


Figure 3.8: The action node containment hierarchy.

3.3.1.6 Data Source-Target Reference

In order to correctly set a data flow's data source and data target, we have to keep in mind where we are referencing the pins. A given *InputPin* can be considered a *DataTarget* from outside of the associated *Activity*, however, it is a *DataSource* from inside the *Defintion*. *DataNodes* can be considered both data sources and data targets. See Figure 3.9 for the meta-model.

Example of inside-input/outside-output pin Figure 3.10 shows an example of the aforementioned effect. From the perspective of the flow $P_1 \rightarrow P_2$ the pin P_1 is a *DataSource* and the pin P_2 is a *DataTarget*. However, from the perspective of the flow $P_2 \rightarrow P_3$, the pin P_2 is a *DataSource* - because the latter flow is inside the composite activity.

[3]This fact means, that if one used Action to define an activity, that activity is executed atomically; it will not be interlaced with other XSTS transitions. See Chapter 4.

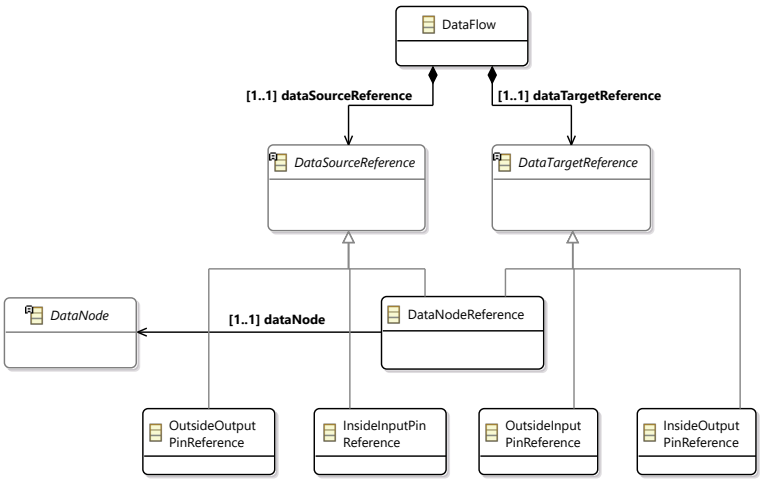


Figure 3.9: The Data Source-Target reference structure

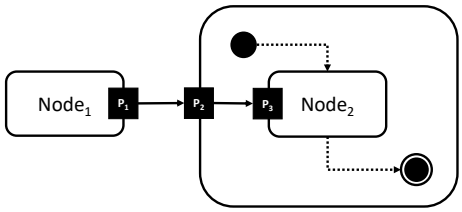


Figure 3.10: The Data node reference structure

3.3.1.7 Pin Reference

Pin references are used to define a rigid pin-reference structure in the model. *InputPinReferences* have a reference to a specific *InputPin*, *OutputPinReferences* have a reference to a specific *OutputPin*. *InsidePinReferences* and *OutsidePinReferences* are used to tell the direction the reference sees the given pin; inside reference sees it from the inside, outside reference sees it from the outside. The *OutsidePinReference* must also have a reference to the specific *ActionNode* the pin is associated with[4]. See Figure 3.11 for the meta-model.

3.3.2 Concrete Syntax

In order to make it easier to test and visualise activities, I created an Xtext domain-specific language. The SysMLv2 language served as the main inspiration for the language, however, much of the *syntactic sugars* have been omitted for simplicities' sake. As a result, many constructs are inherently repetitive to write. You can see an example in Listing 3.1. For the exact language definition, please see Section A.2

[4]Note, that the inside pin reference does not have a node reference, because it is implicitly the containing activity

```
activity CompilationProcess {
    var errors : boolean := false

    initial Initial
    merge Merge
    fork Fork

    action Read1 : activity(
        out p : integer
    )
    action Compile1 : activity(
        in p : integer
    )
    action Read2 : activity(
        out p : integer
    )
    action Compile2 : activity(
        in p : integer
    )

    join Join
    decision Decision
    action Edit
    final Final

    control flow from Initial to Merge
    // ...
    data flow from Read1.p to Compile1.p
    control flow from Fork to Read2
    // ...
    control flow from Decision to Edit [errors]
    control flow from Edit to Merge [!errors]
    control flow from Decision to Final
}
```

Listing 3.1: Gamma Activity Language representation of the compilation activity.

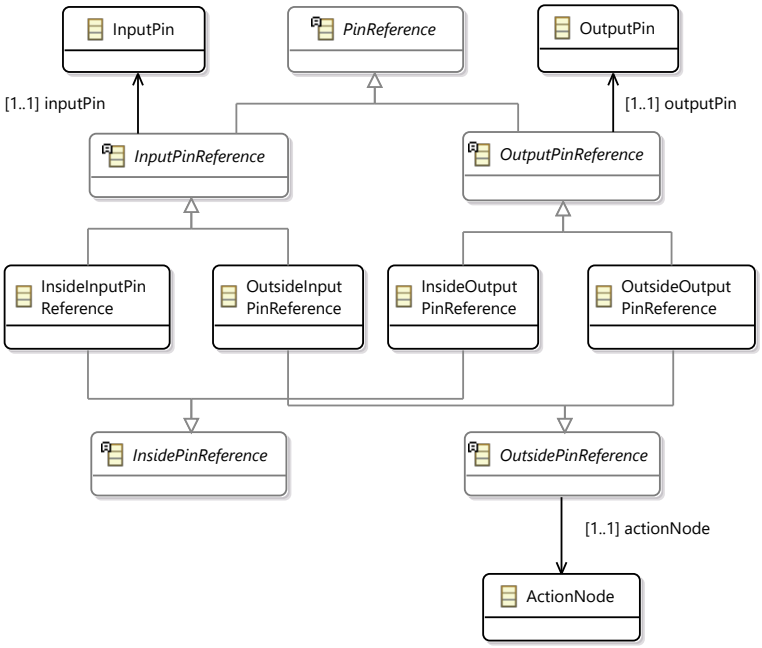


Figure 3.11: The Pin reference structure

Chapter 4

Integrating Activity Language Into Gamma

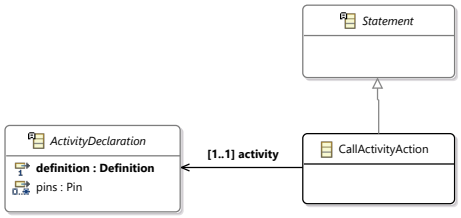
In the previous chapter, I have introduced the Gamma Activity Language - it's formalism and textual representation. However, in order to use it for model verification, it has to be integrated into the Gamma transformation pipeline (see Figure 2.7). For this integration, I start by defining *how* it integrates with Gamma in Section 4.1, after which I show transformation algorithm in Section 4.2. Finally, I talk a bit about the implementation details in Section 4.3.

4.1 Activities Alongside Statecharts

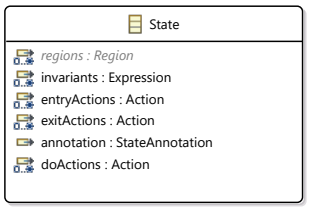
In the real world, SysML models often contain state machines that call activity diagrams using a state's *do behaviour*, as often times this makes the design process easier. However, Gamma does not support statecharts containing *do activity* actions. Thus, first I extend the Gamma meta-model to support calling activities from states.

4.1.0.1 Gamma Extensions

First, I extended Gamma *States* to include *do* actions as well among *entry* and *exit* actions (see Figure 4.1b). In order to enable calling activities from statecharts, I added the *CallActivityAction*, which has a reference to an activity declaration (see Figure 4.1b). These changes can be seen visually in Figure 4.2. Added parts are outlined with green, while edited parts are outlined with orange.



(a) Call activity action



(b) State modification

Figure 4.1: Extensions to the Gamma meta-model

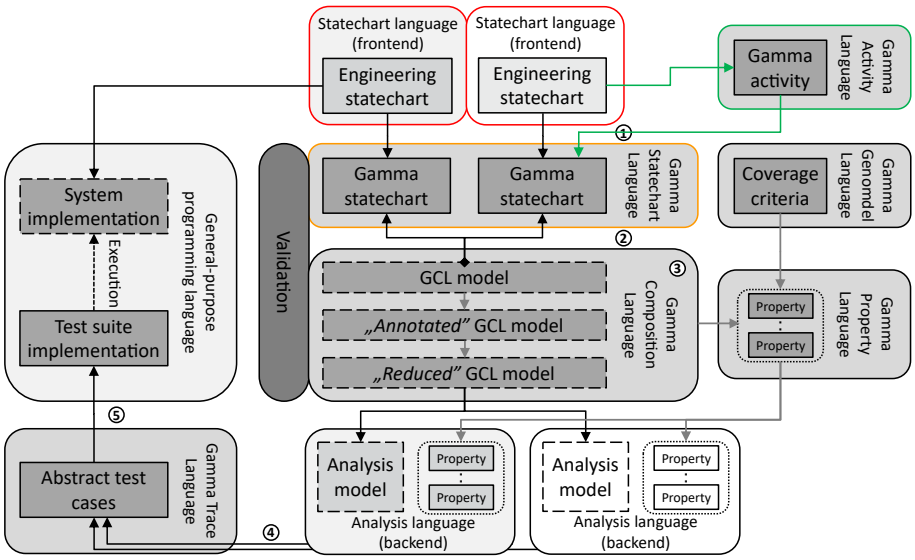


Figure 4.2: Overview of my modification to the Gamma framework.

4.1.1 Challenges of Calling Activities

how to integrate into gamma, action, do action, etc, transition gamma használja az xsts-t -> xsts-re kell trafózni

ez a rész csak jegyzet

Initial/final node, action node, composite activity node, decision/merge node, fork/join node, pins, control/data flows.

Do behaviour:

call activity action on transitions is not supported, would have to be inlined and flattened, atomic functions, interlaced with line execution semantics (future work)

- call activity context függő - hosszú (nem atomikus) dolgot nem lehet akárhol meghívni (transition, sima action, stb) - validáció: inline környezetben tilos várakozás és ciklus (vagy csak x mélységig) - sok köztes nem stabil állapot - nagy változás a trafóban

ezért csak do activity van

fentebb leírtam, hogyan lehet leképezni proces modelleket, de így csak magukban futhatnak. Itt most leírom, hogy milyen nehézségeket okozhat, amikor egymás mellett futtatjuk őket:

do activity:

amikor a state aktív, futhat az aktivity, de ez pár problémával jön: párhuzamossági és szinkornizációs problémák

action:

egy activity értékét ki kell lapítani, hogy a tranzíció tüzelése után azonnal értelmezhető legyen (nem felezhetjük el az adott tranzíciót, mert akkor érvénytelen állapotaink lesznek)

unblock:

még nagy kérdés, hogy milyen szinten bontsuk fel a lépéseket, mennyire legyenek atomikusak, mik futhatnak egymás mellett, stb

alap elképzelésben akár egy activity akár egy állapotgép definiálhat egy komponsens viselkedését, de jelenleg ezt is egyszerűsített módon implementáltuk; az activity az állapotgép része lehet

megoldás a kérdésekre:

megoldásként a legegyszerűbb módszert választottuk, mert első sorban az a kérdés, hogy ez a módszer egyáltalán alkalmazható-e; a másik mindenképpen bonyolítja az implementációt és a létrejövő modellt is

4.2 Activities as State-based Models

Because Gamma transforms its models to XSTS in order to run formal verification on them, to integrate activity language, it must also be transformed to XSTS. Luckily, the structure and behaviour of the activity formalism makes this transformation easy.

4.2.1 Structure Transformation

The state of the activities (State Transition Functions 3.2) are defined by the nodes' *states* and *values*, it's pins' *values* and the flows' *states* and *values*. All of these can be represented using *variables* in the XSTS formalism. Because XSTS variables and types must have unique names, the name of the variables must be derived from their containers. In the following I will show the transformation algorithm using the XSTS's textual representation for easier understanding.

The transformations steps are:

Step One In order to create variables for the *state* of the nodes and flows, we must declare a *NodeState* and *FlowState* type.

```
type NodeState : { __Idle__, __Running__, __Done__ }
type FlowState : { __Empty__, __Full__ }
```

Step Two Now that we have the types needed, we can create the mappings of the nodes and flows. Flows are mapped to a variable of type *FlowState*. If it is a *DataFlow*, an additional *value* variable has to be added, which will have a type of the connected pin[1].

```
var activity_flow : FlowState = __Empty__
var activity_flow_value : integer = 0
```

Nodes are mapped to a variable of type *NodeState*. After which, according to its exact type, additional variables are added; if it is a *DataNode*, a single *value* variable will be added with the type of the connected *DataNode*. Otherwise, if it is an *ActionNode*, then all of its pins will be mapped to a variable - each with their own types.

```
var activity_node : NodeState = __Idle__
var activity_node_pin1 : integer = 0
var activity_node_pin2 : integer = 0
var activity_decision_node : NodeState = __Idle__
var activity_decision_value : boolean = false
```

[1]The connected pin can be determined by traversing the chain of flows in a direction, until we reach a pin

Step Three Now that we have mapped the state of the activity, we must map the state-transition functions. This

```
1 assume ((activity_flow == __Empty__) && (activity_node == __Done__));
2 activity_flow := __Full__;
3 activity_node := __Idle__;
```

4.3 Implementation Regards

implementációs nehézségek, kérdések, stb (commitok, line change, projektek magas szinten)

Chapter 5

Evaluation

spacecraft model bevezetése, leírása statechart elemek referálása, activity-k megírása kézzel (modell mehet apendix-be) mérések végzése és leírása

Chapter 6

Conclusion

ami nem ment evaluationbe - lassú - nem skálázódik - de egy első lépés, hogy unified verifikációt lehessen csinálni komplikált rendszereken

Future Work - sysml activity nagyobb mappelése - signalok bevezetése - activity külön komponensként - activity inline-olása transition action-re

Acknowledgements

Ez nem kötelező, akár törölhető is. Ha a szerző szükségét érzi, itt lehet köszönetet nyilvánítani azoknak, akik hozzájárultak munkájukkal ahhoz, hogy a hallgató a szakdolgozatban vagy diplomamunkában leírt feladatokat sikeresen elvégezze. A konzulensnek való köszönetnyilvánítás sem kötelező, a konzulensnek hivatalosan is dolga, hogy a hallgatót konzultálja.

List of Figures

Listings

Todo list

nyugi, ez a rész nem lesz benne a véglegesben

Bibliography

[1] Technical operations international council on systems engineering incose. incose systems engineering vision 2020. technical report. URL https://sebokwiki.org/wiki/INCOSE_Systems_Engineering_Vision_2020.

[2] Mbse wiki. URL https://www.omgwiki.org/MBSE/doku.php?id=start.

[3] David A. Larsen K. G. Håkansson J. Pettersson P. Yi W. Hendriks M. Behrmann, G. Uppaal 4.0. 2006.

[4] G Bellinger. Modeling & simulation: An introduction. 2004. URL http://www.systems-thinking.org/modsim/modsim.htm.

[5] DoD. Dod modeling and simulation (m&s) glossary. *DoD Manual 5000.59-M. Arlington, VA, USA: US Department of Defense*, 1998. URL https://apps.dtic.mil/sti/pdfs/ADA349800.pdf.

[6] D. Dori. Object-process methodology: A holistic system paradigm. *New York, NY, USA: Springer.*, 2002.

[7] A. Moore R. Steiner Friedenthal, S. and M. Kaufman. A practical guide to sysml: The systems modeling language, 3rd edition. *MK/OMG Press.,*, 2014.

[8] Bence Graics, Vince Molnár, András Vörös, István Majzik, and Dániel Varró. Mixed-semantics composition of statecharts for the component-based design of reactive systems. *Software and Systems Modeling*, 19(6):1483–1517, Nov 2020. ISSN 1619-1374. DOI: 10.1007/s10270-020-00806-5. URL https://doi.org/10.1007/s10270-020-00806-5.

[9] Object Management Group. Omg system modeling language. URL https://www.omg.org/spec/SysML/.

[10] Object Management Group. Mda foundation model. omg document number ormsc/2010-09-06. 2010.

[11] Object Management Group. Semantics of a foundational subset for executable uml models. 2018. URL https://www.omg.org/spec/FUML/1.4/.

[12] Object Management Group. Systems modeling language version 2 (sysmlv2). 2020. URL https://www.omgsysml.org/SysML-2.htm.

[13] D. Harel and A. Pnueli. On the development of reactive systems. In Krzysztof R. Apt, editor, *Logics and Models of Concurrent Systems*, pages 477–498, Berlin, Heidelberg, 1985. Springer Berlin Heidelberg. ISBN 978-3-642-82453-1.

[14] David Harel. Statecharts: a visual formalism for complex systems. *Science of Computer Programming*, 8(3):231–274, 1987. ISSN 0167-6423. DOI: https://doi.org/10.1016/0167-6423(87)90035-9. URL https://www.sciencedirect.com/science/article/pii/0167642387900359.

[15] Osman Hasan and Sofiène Tahar. Encyclopedia of information science and technology, third edition. pages 7162–7170., 2015. DOI: https://doi:10.4018/978-1-4666-5888-2.ch705.

[16] Benedek Horváth, Bence Graics, Ákos Hajdu, Zoltán Micskei, Vince Molnár, István Ráth, Luigi Andolfato, Ivan Gomes, and Robert Karban. Model checking as a service: Towards pragmatic hidden formal methods. In *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, MODELS '20, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450381352. DOI: 10.1145/3417990.3421407. URL https://doi.org/10.1145/3417990.3421407.

[17] Edward Huang, Leon F. McGinnis, and Steven W. Mitchell. Verifying sysml activity diagrams using formal transformation to petri nets. *Systems Engineering*, 23(1):118–135, 2020. DOI: https://doi.org/10.1002/sys.21524. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sys.21524.

[18] Balcer MJ. Mellor SJ. Executable uml: A foundation for model- drivenarchitecture. *The Addison-Wesley Object TechnologySeries: Addison-Wesley Professional*, 2002.

[19] Vince Molnár, Bence Graics, András Vörös, István Majzik, and Dániel Varró. The Gamma statechart composition framework: design, verification and code generation for component-based reactive systems. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceeedings*, pages 113–116. ACM, 2018. DOI: 10.1145/3183440.3183489.

[20] Milán Mondok. Extended symbolic transition systems: an intermediate language for the formal verification of engineering models. *Scientific Students' Association Report*, 2020.

[21] Zoltán Micskei Márton Elekes. Towards testing the uml pssm test suite. 2021.

[22] T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, 1989. DOI: 10.1109/5.24143.

[23] OMG. Precise semantics of uml state machines (pssm). *formal/19-05-01.*, 2019.

[24] Gianna Reggio, Maurizio Leotta, and Filippo Ricca. Who knows/uses what of the uml: A personal opinion survey. In Juergen Dingel, Wolfram Schulte, Isidro Ramos, Silvia Abrahão, and Emilio Insfran, editors, *Model-Driven Engineering Languages and Systems*, pages 149–165, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11653-2.

[25] Hajdu A. Vörös A. Micskei Z. Majzik I. Tóth, T. Theta: a framework for abstraction refinement-based model checking. *Stewart, D., Weissenbacher, G. (eds.), Proceedings of the 17th Conference on Formal Methods in Computer-Aided Design.*

Appendix

A.1 XSTS Language

In this appendix I introduce the exact language constructs for the XSTS language

A.1.1 Types

XSTS contains two default variable types, logical variables (*boolean*) and mathematical integers (*integer*). XSTS also allows the user to define *custom types*, similarly to enum types in common programming languages.

A custom type can be declared the following way:

```
type <name> : { <literal_1>, . . . , <literal_n> }

type Color : { RED, GREEN, BLUE }
```

A.1.2 Variables

Variables can be declared the following way, where `<value>` denotes the value that will be assigned to the variable in the initialization vector:

```
var <name> : <type> = <value>
```

The variable can only take a value of the specified type.

If the user wishes to declare a variable without an initial value, this is possible as well:

```
var <name> : <type>
```

A variable can be tagged as a control variable with the keyword `ctrl`:

```
ctrl var <name> : <type>
```

In which case the variable v will also be added to V_C (the set of control variables).

Examples:

```
var a : integer
ctrl var b : boolean = false
var c : Color = RED
```

A.1.2.1 Operations

The behaviour of XSTS can be described using basic and composite operations. Basic operations include assignments, assumptions and havocs. In the following, you can see an

example for each, where `<expr>` is an expression returning a value, and `<varname>` is the name of a variable.

```
assume <expr>

<varname> := <expr>

havoc <varname>
```

Composite operations are either non-deterministic choices, sequences or parallels. Nondeterministic choices have the following syntax, where `<operation>` are arbitrary basic or composite operations:

```
choice {
    <operation>
} or {
    <operation>
}
```

Sequences have the following syntax:

```
<operation>
<operation>
<operation>
```

And parallels have the following syntax:

```
parallel {
    <operation>
    <operation>
}
```

A.1.2.2 Transitions

Each transition is a single operation (basic or composite). We distinguish between three sets of transitions, *Tran*, *Init* and *Env*. Transitions are described with the following syntax, where `<transition-set>` is either `tran`, `env` or `init`:

```
<transition-set> {
    <operation>
} or {
    <operation>
} or
//...
or {
    <operation>
}
```

A.2 Gamma Activity Language

A.2.1 Language elements

The following section gives high level introduction into the syntax of the Gamma Activity DSL.

A.2.1.1 Pins

Pins can be declared the following way, where `<direction>` can be either `in` or `out`, and `type` a valid Gamma Expression type.

```
<direction> <name> : <type>
```

For example:

```
in examplePin : integer
```

where the direction is `in`, the name is `examplePin` and the type is `integer`.

Nodes

Nodes can be declared by stating the `type` of the node and then it's name. The type determines the underlying meta element.

The available node types:

```
initial InitialNode
decision DecisionNode
merge MergeNode
fork ForkNode
join JoinNode
final FinalNode
action ActionNode
```

A.2.1.2 Flows

The behaviour of the Activity can be described by stating `data` or `control` flows between two nodes. Flows may have `guard`s on them, which limits when the flow can fire. Activities may only be from the current activity definition's children. Pins can be accessed using the `.` accessor operator, the activity on the left hand side, and the pin's name on the right. The enclosing activity's name is `self`.

Flows can be declared the following way, where `<kind>` is the kind of flow, `<source>` and `<target>` is the source/target node or pin, and `<guard>` is a Gamma Expression returning boolean:

```
<kind> flow from <source> to <target> [<guard>]

control flow from activity1 to activity2
control flow from activity1 to activity2 [x == 10]
data flow from activity1.pin1 to activity2.pin2
data flow from self.pin to activity3.pin2
```

A.2.1.3 Declarations

Activity declarations state the *name* of the activity, as well as its *pins* A.2.1.1.

```
activity Example (
    //..pins..
) {
    //..body..
}
```

You can also declare activities inline by using the `:` operator:

```
activity Example {
    action InlineActivityExample : activity
}
```

A.2.1.4 Definitions

Activities also have definitions, which give them bodies. The body language can be either `activity` or `action` depending on the language metadata set. Using the `action` language let's you use any Gamma Action expression, including timeout resetting, raising events through component ports, or simple arithmetic operations.

An example activity defined by an action body:

```
activity Example (
    in x : integer,
    out y : integer
) [language=action] {
    self.y := self.x * 2;
}
```

Inline activities may also have pins and be defined using action language:

```
activity Example {
    action InlineActivityExample : activity (
        in input : integer,
        out output : integer
    ) [language=action] {
        self.output := self.input;
    }
}
```

A composite activity can be easily created using inline activity definition:

```
activity Example {
    initial Init1
    final Final1

    action InlineActivityExample : activity {
        initial Init2
        final Final2

        control flow from Init2 to Final2
    }

    control flow from Init1 to InlineActivityExample
    control flow from InlineActivityExample to Final1
}
```

A.2.1.5 Example

A simple example activity can be seen in Listing A.2.1. This example shows two read actions, that return a random value, which are added together and logged to the console. The addition is defined using a *NamedActivitDeclaraionReference*.

```
activity Adder(
    in x : integer,
    in y : integer,
    out o : integer
) [language=action]  {
    self.o := self.x + self.y;
}

activity Example {
    initial Initial

    action ReadSelf1 : activity (
        out x : integer
    )
    action ReadSelf2 : activity (
        out x : integer
    )
    action Add : Adder
    action Log : activity (
        in x : integer
    )

    final Final

    control flow from Initial to ReadSelf1
    control flow from Initial to ReadSelf2
    control flow from Initial to Add
    data flow from ReadSelf1.x to Add.x
    data flow from ReadSelf2.x to Add.y
    data flow from Add.o to Log.x
    control flow from Log to Final
}
```

Listing A.2.1: Basic adder example.

A.3 Spacecraft Model