Formal Modeling (SS 2020)
Some Proseminar Topics

Wolfgang Schreiner
Research Institute for Symbolic Computation (RISC)
Wolfgang.Schreiner@risc.jku.at

Your task is to study a particular mathematical/computational problem from literature, to develop a formal model of this problem in the RISCAL language, and to validate this model with the RISCAL software. The model may be the specification of a computational problem by a pair of pre/post-conditions; these may be validated by analyzing the inputs/output pairs allowed by the specification and by checking the correctness of an algorithm with respect to this specification. The model may also describe a computational system by an initial state condition and a transition relation; these may be validated by analyzing the sequences of system states arising from this model. Your goal is to find a way to formalize the model and to validate this formalization that is adequate for the chosen problem.

The following topics are only examples, you may also suggest other ones (to be approved by the lecturer). Most topics may also be pursed in the frame of *bachelor theses* (in this case, also algorithms have to be formulated and appropriately annotated such that verification conditions can be generated and checked).

Topic: Formalization of Floating Point Arithmetic

A floating point number is essentially a triple (s, m, e) of sign $s = \pm 1$, mantissa $m \in \mathbb{N}$ with $n < 2^n$, and exponent $e \in \mathbb{N}$, for some $n \in \mathbb{N}$; it describes the rational number $x = s \cdot (1 + m/2^n) \cdot 2^e$. All s, m, e are represented as sequences of bits; the IEEE 754 standard defines the details of the representation including special values such as ± 0, $\pm\infty$ and NaN (not a number).

The goal is to model (an appropriately chosen subset of) this standard and some arithmetic operation (e.g., floating point addition) such that from given inputs in binary form the appropriate output is appropriately determined.

Topic: Formalization of Multi-Precision Integer Arithmetic

A multi-precision integer (s, d) consists of a sign $s = \pm 1$ and a sequence d of n integers $d_0 \ldots d_{n-1}$ representing the value $x = s \cdot \sum_{i=0}^{n-1} d_i \cdot b^i$.

The goal is to develop an appropriate RISCAL model of multi-precision integer arithmetic including algorithms for some basic operations (comparison, addition, multiplication). The challenge is to cope

with the fact that RISCAL only supports arrays of fixed lengths (possibly denoted by model parameters) such that the model has to deal with the problem of overflow/underflow appropriately. One possibility is to extend the representation to include special "overflow" values; another possibility is to appropriately restrict inputs by preconditions that rule out overflow. A third possibility is to represent the results by larger arrays; however, then before continuing with the result an overflow check and a conversion to the smaller arrays is needed. The various alternatives shall be investigated and their advantages respectively disadvantages compared.

Topic: Formalization of the Knuth-Morris-Pratt Algorithm

The Knuth-Morris-Pratt (KMP) algorithm solves the problem of searching in a string s for (the first position respectively all positions) of a string w. While the naive searching algorithm has asymptotic time complexity $|s| \cdot |w|$, the KMP algorithm has complexity $|s| + |w|$. The algorithm proceeds in two phases: in the first phase w is analyzed and translated into a table W; the second phase uses W rather than w for searching in s.

The goal is to model the problem of the precomputation of W and the problem of the search in s with respect to W (both by suitable pre/postconditions).

Topic: Formalization of Directed Graphs

A directed graph (V, E) consists of a set of "vertices" V and an relation $E \subseteq V \times V$ (the set of directed "edges").

The goal is to model the basic theory of directed graphs including basic notions (such as indegree, outdegree, neighborhood, reachability, connectivity, acyclicity, etc.), basic properties (theorems), and basic problems (e.g., shortest path computation). For this suitable literature has to be investigated.

Topic: Formalization of Trees

In graph theory, a tree is an undirected graph in which any two vertices are connected by exactly one path (or, equivalently, a connected acyclic graph).

The goal is to model the theory of trees including derived/equivalent definitions of trees, basic notions of trees (the leaves of a tree, the height/degree of the tree with respect to a given root, etc.), basc properties (theorems), and basic problems (e.g., spanning tree computation). For this suitable literature has to be investigated.

Topic: Formalization of Binary Decision Diagrams

A binary decision diagram is a directed acyclic graph (representing a binary tree with sharing of subtrees) that encodes a multi-ary boolean function. In a nutshell, the leaves of the diagram are constant truth values while every non-leaf node represents a boolean variable; its outgoing edges denote those subtrees that describe the value of the function for the two possible values of the variable.

The goal is to model the basic notion of binary decision diagrams and their evaluation and modeling the computation of binary decision diagrams from propositional formulas.

Topic: Formalization of Games

Choose some (card/tabletop) game and model it as a non-deterministic state transition system with an initial state condition, transition relation, and termination (winning) condition. The transition relation models the possible actions of the players by some guard condition (when is it possible to perform the action?) and the resulting effect on the state of the game. Formulate also some invariants that every state of the game must satisfy and check in RISCAL whether the game indeed preserves these invariants.

Topic: Formalization of the "Steam Boiler Control Specification"

In the accompanying lecture, a simple variant of the "steam boiler control specification" problem is modeled that describes a controller that regulates the ingoing flow of water to a steam boiler such that the boiler operations remains safe. This simple model assumes that most devices operate without failures.

The original problem described in literature, however, deals with the failures of the various components such as the pump controllers or the devices for measuring the water level, steam quantity, or pump flow. The goal is to extend/adapt/revise the presented model to capture also these problems. For this, the available literature (many publications describe formalizations of this problem) provides ample ideas.

Topic: Formalization of a Distributed Resource Allocator

The case study "A Resource Allocator"[1] describes a model of a distributed system where a finite set of resources is managed and distributed to a number of clients. A model of this case study has been developed in the specification language TLA+ and validated with the TLA model checker (a simplified version of this model is available from the lecturer).

The goal is to develop a corresponding RISCAL model and to appropriately validate it in RISCAL.

[1] https://link.springer.com/chapter/10.1007/978-3-540-74107-7_8
https://members.loria.fr/SMerz/projects/tla/allocator.pdf