The Semantics of K

Formal Systems Laboratory
University of Illinois

August 23, 2017

Please feel free to contribute to this report in all ways. You could add new contents, remove redundant ones, refactor and organize the texts, and correct typos, but please follow the FSL rules for editing, though; e.g., <80 characters per line, each sentence on a new line, etc.

1 Matching Logic

Let us recall the basic grammar of matching logic from [**?**]. Assume a matching logic *signature* (S, Σ), and let Var_s be a countable set of *variables* of sort s. For simplicity, here we assume that the sets of *sorts* S and of *symbols* Σ are finite. We partition Σ in sets of symbols $\Sigma_{s_1 \ldots s_n, s}$ of *arity* $s_1 \ldots s_n, s$, where $s_1, \ldots, s_n, s \in S$. Then *patterns* of sort $s \in S$ are generated by the following grammar:

Add references.

$$
\begin{array}{rl}
\varphi_s ::= & x{:}s \quad \text{where } x \in Var \\
\mid & \varphi_s \wedge \varphi_s \\
\mid & \neg \varphi_s \\
\mid & \exists x{:}s'.\varphi_s \quad \text{where } x \in N \text{ and } s' \in S \\
\mid & \sigma(\varphi_{s_1}, \ldots, \varphi_{s_n}) \quad \text{where } \sigma \in \Sigma \text{ has } n \text{ arguments, and } \ldots
\end{array}
$$

Figure 1: The grammar of matching logic.

The grammar above only defines the syntax of (well-formed) patterns of sort s. It says nothing about their semantics. For example, patterns $x{:}s \wedge y{:}s$ and $y{:}s \wedge x{:}s$ are distinct elements in the language of the grammar, in spite of them being semantically/provably equal in matching logic.

For notational convenience, we take the liberty to use mix-fix syntax for operators in Σ, parentheses for grouping, and omit variable sorts when understood. For example, if $Nat \in S$ and $\_+\_, \_*\_ \in \Sigma_{Nat \times Nat, Nat}$ then we may write $(x + y) * z$ instead of $\_*\_(\_+\_(x{:}Nat, y{:}Nat), z{:}Nat)$.

I think we also need to talk about: other logical connectives as derived, free variables, capture-free substitution, equality. Add more as we need them.

A matching logic *theory* is a triple (S, Σ, A) where (S, Σ) is a signature and A is a set of patterns called *axioms*. Like in many logics, sets of patterns may be presented as *schemas* making use of meta-variables ranging over patterns, sometimes constrained to subsets of patterns using side conditions. For example:

$\varphi[\varphi_1/x] \wedge (\varphi_1 = \varphi_2) \rightarrow \varphi[\varphi_2/x]$	where φ is any pattern and φ_1, φ_2 are any patterns of same sort as x
$(\lambda x.\varphi)\varphi' = \varphi[\varphi'/x]$	where φ, φ' are *syntactic patterns*, that is, ones formed only with variables and symbols This is not true. Pattern φ contains quantifiers.
$\varphi_1 \texttt{ + } \varphi_2 = \varphi_1 +_{Nat} \varphi_2$	where φ, φ' are *ground* syntactic patterns of sort *Nat*, that is, patterns built only with symbols `zero` and `succ`
$(\varphi_1 \rightarrow \varphi_2) \rightarrow (\varphi[\varphi_1/x] \rightarrow \varphi[\varphi_2/x])$	where φ is a *positive context in* x, that is, a pattern containing only one occurrence of x with no negation ($\neg$) on the path to x, and where φ_1, φ_2 are any patterns having the same sort

One of the major goals of this paper is to propose a formal language and an implementation, that allows us to write such pattern schemas.

2 A Calculus of Matching Logic

In this section, we propose a calculus of matching logic as a matching logic theory.

> Many people have developed calculi for mathematical reasoning. A calculus of logics is often called a *logical framework*. I prefer to speak of a *meta-logic* and its *object-logic*.
> By L. Paulson, *The Foundation of a Generic Theorem Prover*, Journal of Automated Reasoning, 1989.

In this proposal, the *object-logic* refers to matching logic, and we propose to use matching logic itself as the *meta-logic*[1]. The calculus of matching logic, denoted as $K = (S_K, \Sigma_K, A_K)$, is a matching logic theory where S_K, Σ_K, and A_K are sets of sorts, symbols, and axioms respectively., The main goal of this section is to present the calculus K, which mainly consists of built-in theories, abstract syntax trees (ASTs) of patterns, and matching logic proof system. They are introduced in detail in the following separate sections.

[1] Coq and Isabelle use fragments of higher-order logic as their meta-logics.

2.1 Built-ins

Two matching logic theories, BOOL for boolean algebra and STRING for strings, are included in the calculus K. Their definitions have been introduced elsewhere (e.g. in [**?**]) and will not be discussed in this proposal. Sorts *Bool* and *String* are included in S_K, and the following symbols are included in Σ_K with their usual axioms [**?**] added to A_K.

$$
\begin{aligned}
&\mathit{true}, \mathit{false} \ : \ \rightarrow \mathit{Bool} \\
&\mathit{notBool} \ : \mathit{Bool} \rightarrow \mathit{Bool} \\
&\mathit{andBool}, \mathit{orBool} \ : \mathit{Bool} \times \mathit{Bool} \rightarrow \mathit{Bool} \\
&\mathit{impliesBool} \ : \mathit{Bool} \times \mathit{Bool} \rightarrow \mathit{Bool}.
\end{aligned}
$$

Whenever we introduce a sort, say X, to S_K, we feel free to use *XList* as the sort of lists over X with the following symbols implicitly declared:

$$
\begin{aligned}
&\mathit{nilXList} \ : \ \rightarrow \mathit{XList} \\
&\mathit{appendXList} \ : \mathit{XList} \times \mathit{XList} \rightarrow \mathit{XList} \\
&\mathit{XListAsX} \ : X \rightarrow \mathit{XList}.
\end{aligned}
$$

For simplicity, we often write as a shorthand *nil* for *nilXList* and φ_1, φ_2 for $\mathit{appendXList}(\varphi_1, \varphi_2)$. We also omit the injection function *XListAsX* and adopt an "order-sorted like" syntax as in [**?**].

2.2 Abstract Syntax Trees

The sort $\mathit{Sort} \in S_K$ is the sort of matching logic sorts, whose only constructor symbol[2] is $\mathit{sort}\colon \mathit{String} \rightarrow \mathit{Sort}$. The sort $\mathit{Symbol} \in S_K$ is the sort of matching logic symbols whose only constructor symbol is $\mathit{symbol}\colon \mathit{String} \times \mathit{SortList} \times \mathit{Sort} \rightarrow \mathit{Symbol}$. The sort $\mathit{Pattern} \in S_K$ is the sort for ASTs of patterns, with the following functional constructor symbols:

$$
\begin{aligned}
&\mathit{variable}\colon \mathit{String} \times \mathit{Sort} \rightarrow \mathit{Pattern} \\
&\mathit{and}, \mathit{or}, \mathit{implies}, \mathit{iff} : \mathit{Pattern} \times \mathit{Pattern} \rightarrow \mathit{Pattern} \\
&\mathit{equals}, \mathit{contains}\colon \mathit{Pattern} \times \mathit{Pattern} \times \mathit{Sort} \rightarrow \mathit{Pattern} \\
&\mathit{not}\colon \mathit{Pattern} \rightarrow \mathit{Pattern} \\
&\mathit{exists}, \mathit{forall}\colon \mathit{String} \times \mathit{Sort} \times \mathit{Pattern} \rightarrow \mathit{Pattern} \\
&\mathit{application}\colon \mathit{Symbol} \times \mathit{PatternList} \rightarrow \mathit{Pattern} \\
&\mathit{top}, \mathit{bottom}\colon \mathit{Sort} \rightarrow \mathit{Pattern}.
\end{aligned}
$$

There are also AST-related symbols included in Σ_K. For example, the symbol $\mathit{wellFormed}\colon \mathit{Pattern} \rightarrow \mathit{Bool}$ determines whether a pattern is well-formed

[2] A constructor symbol has its precise definition in matching logic. Please refer to [**?**].

(or more precisely, it determines whether an abstract syntax tree is a well-formed one of a pattern.) The symbol $\mathit{getSort} \in \Sigma_{\mathit{Pattern},\mathit{Sort}}$ takes a pattern and returns its sort. If the pattern is not well-formed, then *getSort* returns $\bot_{\mathit{Sort}}$; otherwise, *getSort* returns $\mathit{sort}(s)$ if the pattern has sort s. The symbol $\mathit{getFvs}\colon \mathit{Pattern} \to \mathit{PatternList}$ collects all free variables in a pattern. The symbol $\mathit{substitute}\colon \mathit{Pattern} \times \mathit{Pattern} \times \mathit{Pattern} \to \mathit{Pattern}$ takes a target pattern φ, a "find"-pattern ψ_1, and a "replace"-pattern ψ_2, and returns φ in which ψ_2 is substituted for ψ_1, denoted as $\varphi[\psi_2/\psi_1]$.

Side conditions can be defined as functional symbols from *Pattern* to *Bool*. For example, the symbol *syntactic* determines whether a pattern contains only variables and symbol applications. The symbol *ground* determines whether a pattern is variable-free, no matter free or bound. The symbol *groundSyntactic* determines whether a pattern is both syntactic and ground. They all can be easily defined in K. We will provide examples in later sections.

2.3 Proof System

It is strongly recommended that readers read L. Paulson's *The Foundation of a Generic Theorem Prover*, especially Section 2, 3, and 4.

A proof system is a theorem generator. In K, the proof system of matching logic is captured by the functional symbol $\mathit{deducible}\colon \mathit{Pattern} \to \mathit{Bool}$, which returns *true* iff the argument pattern is a theorem. Given a matching logic pattern φ, we use $\mathit{lift}(\varphi)$ to denote its abstract syntax tree, where $\mathit{lift}(\_)$ is called the *lifting function*

Xiaohong: I am considering using brackets $\mathit{lift}[\_]$ instead of parenthesis to writing the lifting function, because it is in fact not a function. Just think this: what is the domain of the lifting "function"?

that maps object-patterns to their meta-representations in K. It worths to point out that the lifting function *cannot* be defined in K no matter what. It is purely a mathematical notation and is not part of the calculus. To see that, simply consider $\mathit{lift}(0)$ and $\mathit{lift}(x - x)$, where $0 = x - x$ but their ASTs are different:

$$\begin{aligned}\mathit{lift}(0) &= \mathit{application}(\mathit{symbol}(\text{“0”}, \cdots), \dots)\\ &\neq \mathit{lift}(x - x)\\ &= \mathit{application}(\mathit{symbol}(\text{“}\_-\_\text{”}, \cdots), \dots)\end{aligned}$$

This means that the following equational substitution deduction

$$\frac{\varphi_1 = \varphi_2}{\mathit{lift}(\varphi_1) = \mathit{lift}(\varphi_2)}\ \text{(WRONG)}$$

does not hold. It is a strong evidence that $\mathit{lift}(\_)$ is not part of the logic.

We introduce the double bracket $[\![\_]\!]$, known as the semantics bracket, as follows:

$$[\![\varphi]\!] \equiv (\mathit{deducible}\,(\mathit{lift}(\varphi)) = \mathit{true})\,.$$

Intuitively, $[\![\varphi]\!]$ means that "φ is deducible". Whenever there is an inference rule (axioms are considered as rules with zero premise)

$$\frac{\varphi_1, \ldots, \varphi_n}{\psi}$$

in matching logic, there is a corresponding axiom in K:

$$[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_n]\!] \rightarrow [\![\psi]\!].$$

Inference modulo theories can be considered in the same way. For any (syntactic) matching logic theory T whose axiom set is A, we add

$$[\![\varphi]\!] \quad \text{for all } \varphi \in A$$

as axioms to K. We sometimes denote the extended theory as $\mathit{lift}(T)$ and call it the *meta-theory for* T.

2.4 Faithfulness

It remains a question whether the calculus K faithfully captures matching logic reasoning. The following definition of *faithfulness* is inspired by [**?**].

Definition 1. The calculus K is said to be faithful for matching logic, if for any matching logic syntactic theory T and its meta-theory $\mathit{lift}(T)$,

$$\varphi \text{ is a theorem in } T \text{ iff } [\![\varphi]\!] \text{ is a theorem in } \mathit{lift}(T), \text{ for any pattern } \varphi.$$

Theorem 2. *The calculus K is faithful for matching logic.*

Proof. TBC. □

Having a faithful calculus for matching logic has at least the following two benefits. Firstly, any implementation of the calculus is guaranteed to be able to conduct any reasoning in matching logic. Secondly, it allows us to define a matching logic theory T by defining its meta-theory $\mathit{lift}(T)$ in the calculus K. The second point is of great importance if we want a formal language to define matching logic theories. We notice that there are many theories whose definitions involve notations that do not belong to the logic itself. For example, in the (β) axiom

$$\lambda x.e[e'] = e[e'/x] \quad \text{where } e \text{ and } e' \text{ are } \lambda\text{-terms},$$

we use the notation for substitution $\_[\_/\_]$, meta-variables e and e', and their range "λ-terms". None of those can be given a formal semantics in the object-logic, but can be defined in the calculus K.

3 The Kore Language

We have shown K, a calculus for matching logic in which we can specify everything about matching logic and matching logic theories, such as whether a pattern is well-formed, what sort a patter has, which patterns are deducible, free variables, fresh variables generation, substitution, etc. The calculus K provides a universe of pattern ASTs and the sound and complete proof system of matching logic. On the other hand, it is usually easier to work at object-level rather than meta-level. Even if all reasoning in a matching logic theory T can be faithfully lifted to and conducted in its meta-theory $\mathit{lift}(T)$, it does not mean one should always do so.

The Kore language is proposed to define matching logic theories using the calculus K. At the same time, it also provides a nice surface syntax (syntactic sugar) to write object-level patterns. We will firstly show the formal grammar of Kore in Section 3.1, followed by some examples in Section 3.2. After that, we will introduce a transformation from Kore definitions to meta-theories as the formal semantics of Kore in Section **??**.

3.1 Syntax and Semantics of Kore

```
// Namespaces for sorts, variables, metavariables,
// symbols, and Kore modules.
Sort           = String
VariableId     = String
MetaVariableId = String
Symbol         = String
ModuleId       = String

Variable       = VariableId:Sort
MetaVariable   = MetaVariableId::Sort

Pattern        = Variable | MetaVariable
               | \and(Pattern, Pattern)
               | \not(Pattern)
               | \exists(Variable, Pattern)
               | Symbol(PatternList)

Sentence       = import ModuleId
               | syntax Sort
               | syntax Sort ::= Symbol(SortList)
               | axiom Pattern
Sentences      = Sentence | Sentences Sentences

Module         = module ModuleId
                   Sentences
                 endmodule
```

In Kore syntax, the backslash "\" is reserved for matching logic connectives and the sharp "#" is reserved for the meta-level, i.e., the K sorts and symbols.

Therefore, the sorts *Bool*, *String*, *Symbol*, *Sort*, and *Pattern* in the calculus K are denoted as `#Bool`, `#String`, `#Symbol`, `#Sort`, and `#Pattern` in Kore respectively. Symbols in K are denoted in the similar way, too. For example, the constructor symbol $\mathit{variable}\colon \mathit{String} \times \mathit{Sort} \to \mathit{Pattern}$ is denoted as `#variable` in Kore.

A Kore module definition begins with the keyword `module` followed by the name of the module-being-defined, and ends with the keyword `endmodule`. The body of the definition consists of some *sentences*, whose meaning are introduced in the following.

The keyword `import` takes an argument as the name of the module-being-imported, and looks for that module in previous definitions. If the module is found, the body of that module is copied to the current module. Otherwise, nothing happens. The keyword `syntax` leads a *syntax declaration*, which can be either a *sort declaration* or a *symbol declaration*. Sorts declared by sort declarations are called *object-sorts*, in comparison to the five *meta-sorts*, `#Bool`, `#String`, `#Symbol`, `#Sort`, and `#Pattern`, in K. Symbols whose argument sorts and return sort are all object-sorts (meta-sorts) are called *object-symbols* (*meta-sorts*).

Patterns are written in prefix forms. A pattern is called an *object-pattern* (or *meta-pattern*) if all sorts and symbols in it are object (or meta) ones. Meta-symbols will be added to the calculus K, while object-sorts and object-symbols will not. They only serve for the purpose to parse an object pattern.

The keyword `axiom` takes a pattern and adds an axiom to the calculus K. If the pattern is a meta-pattern, it adds the pattern itself as an axiom. If the pattern φ is an object-pattern, it adds $[\![\varphi]\!]$ as an axiom to the calculus K.

Recall that we have defined the semantics bracket as

$$[\![\varphi]\!] \equiv (\mathit{deducible}\,(\mathit{lift}(\varphi)) = \mathit{true})\,,$$

where φ is a pattern of the grammar in Figure 1. However, here in Kore we allow φ containing *meta-variables*. As a result, we modify the definition of the semantics bracket as

$$[\![\varphi]\!] \equiv \mathit{mvsc}(\varphi) \to (\mathit{deducible}\,(\mathit{lift}(\varphi)) = \mathit{true}),$$

where the lifting function $\mathit{lift}(\_)$ and the meta-variable sort constraint *mvsc* are implemented in Algorithm 1 and 2, respectively. Intuitively, meta-variables in an object-pattern φ are lifted to variables of the sort *Pattern* with the corresponding sort constraints. For example, the meta-variable $x{::}s$ is lifted to a variable $x{:}\mathit{Pattern}$ in K with the constraint that $\mathit{getSort}(x{:}\mathit{Pattern}) = \mathit{sort}(s)$. The function *mvsc* collects all such meta-variable sort constraint in an object-pattern is implemented in Algorithm 2.

3.2 Examples of Kore

Xiaohong: Add more examples and texts here.

Algorithm 1: Lifting Function $lift(\_)$

Input: An object-pattern φ.
Output: The meta-representation (ASTs) of φ in K

if φ *is* $x{:}s$ **then**
  Return $variable(x, sort(s))$
else if φ *is* $x{::}s$ **then**
  $x{:}Pattern \wedge (sort(s) = getSort(x{:}Pattern))$
else if φ *is* $\varphi_1 \wedge \varphi_2$ **then**
  $and(lift(\varphi_1), lift(\varphi_2)$
else if φ *is* $\neg\varphi_1$ **then**
  $not(lift(\varphi_1))$
else if φ *is* $\exists x{:}s.\varphi_1$ **then**
  $exists(x, sort(s), lift(\varphi_1))$
else if φ *is* $\sigma(\varphi_1, \ldots, \varphi_n)$ *and* $\sigma \in \Sigma_{s_1,\ldots,s_n,s}$ **then**
  $application(symbol(\sigma, (sort(s_1), \ldots, sort(s_n)), sort(s)), lift(\varphi_1), \ldots, lift(\varphi_n))$

Algorithm 2: Meta-Variable Sort Constraint Collection *mvsc*

Input: An object-pattern φ
Output: The meta-variable sort constraint of φ

Collect in set W all meta-variables appearing in φ;
Let $C = \emptyset$;
foreach $x{::}s \in W$ **do**
  $C = C \cup (sort(s) = getSort(x{:}Pattern))$
Return $\bigwedge C$;

The BOOL module.

```
module BOOL
  syntax Bool
  syntax Bool ::= true | false | notBool(Bool)
                | andBool(Bool, Bool) | orBool(Bool, Bool)
  axiom \or(true(), false())
  axiom \exists(X:Bool, \equals(X:Bool, true()))
  axiom \equals(andBool(B1::Bool, B2::Bool),
                andBool(B2::Bool, B1::Bool))
  axiom ... ...
endmodule
```

The BOOL module (desugared).

```
module BOOL
  axiom \equals(
    #true,
    #deducible(#or(#application(#symbol("true", #nilSort, #sort("Bool")),
                                #nilPattern),
                   #application(#symbol("false", #nilSort, #sort("Bool")),
                                #nilPattern))))
  axiom \equals(
    #true,
    #deducible(#exists("X", #sort("Bool"),
                #equals(#variable("X", #sort("Bool")),
                        #application(#symbol("true", #nilSort, #sort("Bool")),
                                     #nilPattern)))))
  axiom \implies(
    \and(\equals(#getSort(B1:Pattern), #sort("Bool")),
         \equals(#getSort(B2:Pattern), #sort("Bool"))),
    \equals(
      #true,
      #deducible(#equals(#application(#symbol("andBool",
                                              (#sort("Bool"), #sort("Bool"))
                                              #sort("Bool")),
                                      (B1:Patern, B2:Pattern)),
                         #application(#symbol("andBool",
                                              (#sort("Bool"), #sort("Bool"))
                                              #sort("Bool")),
                                      (B2:Patern, B1:Pattern))))))
  axiom ... ...
endmodule
```

The LAMBDA module

```
module LAMBDA
  syntax Exp
  syntax Exp ::= app(Exp, Exp) | lambda0(Exp, Exp)
  syntax #Bool ::= isLTerm(#Pattern)
```

```
  axiom \equals(
    isLTerm(#variable(X:String, #sort("Exp"))),
    true)
  axiom \equals(
    isLTerm(#application(
              #symbol("app", (#sort("Exp"), #sort("Exp")), #sort("Exp")),
              (E:Pattern, E':Pattern))),
    andBool(isLTerm(E:Pattern), isLTerm(E':Pattern)))
  axiom \equals(
    isLTerm(#exists(X:String, #sort("Exp"),
                    #application(#symbol("lambda0",
                                         (#sort("Exp"), #sort("Exp")),
                                         #sort("Exp")),
                                 (#variable(X:String, #sort("Exp")),
                                  E:Pattern))),
    isLTerm(E:Pattern))
  axiom \implies(\equals(true,
                         andBool(isLTerm(E:Pattern),
                                 isLTerm(E':Pattern))),
                 \equals(true,
                         deducible(#equals(...1,
                                           ...2))))
endmodule
```