

ANTI MIN KE

北京大学2015年秋 《编译实习》课程项目

史舒扬 张浩千 倪盛恺



Table of Contents

AMK Project	0
Project Management	1
Schedule and Work Allocation	1.1
User Guide with AMK Language	2
AMK Language Design	2.1
AMK Language Design: Propositional logics	2.2
Programmer's Guide of Code Documentation	3
Lexer Design Documentation	3.1
Parser Design Documentation	3.2

AMK: Anti-Min-Ke

AMK (Anti-Min-Ke) 是为书写和检查证明的人们而设计、实现的程序设计语言。设计的目标是：

- 标准化证明书写；
- 自动检查证明的正误；
- 提供在线的编辑器提供Web服务做到上面两点；
- 做一定程度上的（半）自动证明。

AMK (Anti-Min-Ke) is a home-made language for people writing and checking mathematical proofs. It aims to

- standardize proof writings,
- automatically check the correctness of proofs, and
- (future) do semi-automatically proof.

具体实现领域 Implemented Field

作为一学期的课程项目，我们打算做数理逻辑方面的内容，不失一般性地实现该领域下的功能，之后再考虑拓展领域。

As a course project, we decided to implement AMK in the area of Mathematical Logics (propositional logics). Later it can be extended to other areas of mathematics.

项目网站链接 Project Website

访问 <http://amk.website/> 访问我们的在线版本，提供友好的交互界面。

Visit the online version with friendly interaction at <http://amk.website/>.

项目源码链接 Source Code

从 [Github](#) 访问该项目。

Visit the project at [Github](#)

作者信息 Authors

- 史舒扬
- 张浩千
- 倪盛恺

Project Management

- Schedule and Work Allocation

时间和工作安排 Schedule and Work Allocation

开题报告

- 开题报告: 倪盛恺

Open Report

- Project open report: Shengkai Ni
-

10.18~10.25: 语言设计

- 语言设计文档: 史舒扬
- 实现工具和阶段划分调查: 倪盛恺
- 可能有参考性的工具 coq 的调研: 张浩千

Oct 18 -- Oct 25 : language design

- Language design: Shuyang shi
 - Tool and stage investigation: Shengkai Ni
 - Coq investigation: Haoqian Zhang
-

10.25~11.1: 词法分析

- 词法分析实现: 倪盛恺
- 语法分析调研与草稿: 张浩千
- 工具调研(flex+bison vs ANTLR): 史舒扬

Oct 25 -- Nov 1 : lexical analysis

- Lexical analysis using flex: Shengkai Ni
- Syntactical analysis draft: Haoqian Zhang
- Investigation (Bison vs ANTLR, feature and usage): Shuyang shi

11.1~11.12: 语法分析

- 实现语法分析（构件AST）：史舒扬
- 修改词法分析器Lexer代码，并完成它的文档：倪盛恺
- SDT过程设计与草稿：张浩千

Nov 1 -- Nov 12 : syntactical analysis

- Fix bugs and write doc for lexical part: Shengkai Ni
 - Syntactical analysis and build AST: Shuyang shi
 - Design Syntax-directed translation and write doc for that: Haoqian Zhang
-

11.12~11.18: 语法制导翻译SDT

- 语法制导翻译SDT：张浩千
- 写语法分析器Parser的文档：史舒扬
- 继续修改Lexer的Bug：倪盛恺

Nov 12 -- Nov 18: (First demo) Syntax-directed translation

- Write doc for syntactical part: Shuyang shi
 - Syntax-directed translation: Haoqian Zhang
 - Prepare slides for interim report: Shengkai Ni
-

11.19: 中期报告

- 中期报告：史舒扬

Nov 19 : Mid-Term Report

- Mid-Term Report: Shuyang Shi
-

11.20~12.10: 错误报告与完善测试

- 完善Translator的SDT过程: 张浩千
 - 增加更复杂的例子 (添加完整的公理体系)
 - Debug
- Parser完善: 史舒扬
 - 错误报告
 - 增加对括号 '()' 的支持
 - 存储AST节点的行号
- 实现外壳及模块的方式: 史舒扬
- Web 编辑器调研: 倪盛恺

Nov 20 -- Dec 10 : Error Reporting

- Syntax Refinement: Shuyang shi
 - Error Reporting
 - Support for '()'
 - Storage of Line Number in AST Nodes
 - Shell to implement modules: Shuyang shi
 - Translator Refinement: Haoqian Zhang
 - Debug
 - More Complex Examples
 - Web Editors Investigation: Shengkai Ni
-

12.10~12.24: 进一步的工作

- 增加打印导出功能 (PDF) : 张浩千
- Web编辑器与网站搭建: 倪盛恺
- 尝试实现跳一步功能: 史舒扬

12.10 -- 12.24 : Further Work

- Add feature of printing as LaTeX: Haoqian Zhang
 - Web Editor and website: Shengkai Ni
 - Infer feature within one step: Shuyang Shi
-

12.31: 结题报告

- 结题报告: 张浩千
- 文档整理: 史舒扬

12.31: Final Report

- Presentation: Haoqian Zhang
- Documentation: Shuyang Shi

AMK Language User Guide

- AMK Language Design
- AMK Language Design: Propositional logics

AMK Language Design: Overview

By *Shuyang Shi* @ Oct, 2015

下面是关于AMK语言的总述，会出现一些例子，这些例子默认来自于古典命题逻辑领域。具体各领域的內容在各自的说明文件中查阅。

模块 Module

AMK的设计初衷是为了在数学各个领域的证明中发挥相当的作用，由于这样的扩展性，决定采用模块(module)的形式，每个模块对应一个数学领域，包含该领域一些必须的内容等。不导入模块的解释器基本上可以认为只包含了简单的推理逻辑。

标准模块

预置的标准模块中包含的內容有：

- 运算符
- 变量类型（以及默认类型）
- 公理、基本定理及其证明

基本上一个模块对应一个领域，对于有嵌套的领域，用“.”来表示从属关系（见模块导入的例子，logcs模块下有classic 和 intuitionism 两个分支）。

用户定义模块

在导入模块的时候，解释器会根据模块中的证明对其中的定理证明进行验证（公理不验证，仅提示使用者），这样的一方面能每次检查模块是否被修改，保证证明基础的正确性，更重要的是，允许了用户自己撰写模块并使用（用户定义模块）。这一定程度上保证了AMK的可扩展性。

模块导入

模块一旦导入，那么模块中提供的运算符、变量类型、公理定理等都将可以使用。如果导入了不同模块导致使用有冲突，需要在使用的时候用"."标明来自于哪个模块。（当然，目前我们还是暂时考虑一次运行只导入一个模块的情况）

下面是一个在源文件中导入命题逻辑模块的例子（用**import**关键字在文件头导入模块）：

```
import logics.proposition
```

(导入了文件 `modules/logics/proposition.mamk` 中的内容。)

源文件语法

模块源文件 (`.mamk`, module for **amk**)分为四部分，分别是

- 类型
- 运算符
- 公理
- 基本定理或引理。

这里不仔细展开了，有兴趣的读者可以看完后文的语法介绍以后根据下面的一个例子自行琢磨。有几点注意。

- 分为四大部分，分别是TYPE, OPERATOR, AXIOM, THEOREM
- OPERATOR 中的前缀表示结合性(left, right, nonasc, func)，其中nonasc表示没有结合性，func表示不采用中缀的方式而是类似函数调用传参数的方式使用(e.g. fun(a, b))。
- AXIOM部分的公理不需要证明，THEOREM部分的定理需要证明。

`classic.mamk`

```

TYPE
statement
set[statement] # [statement] here means there is a list of several (finite) statements
numeric # numeric values like int, float, etc

OPERATOR # operators are arranged in decreasing order of their priorities
nonasc not (statement) --> statement

right wedge (statement, statement) --> statement

right vee (statement, statement) --> statement

right -> (statement, statement) --> statement

right <-> (statement, statement) --> statement

right |- (set[statement], statement) --> statement

right |-| (set[statement], statement) --> statement

AXIOM
axiom1:
require:
define a of statement
define b of list[statement]
conclude: a, b |- a

axiom2:
...

THEOREM
naive_theorem:
require:
define a of statement
define b of statement
conclude:
a, b |- a
proof:
a, b |- a -[axiom1]

```

用户代码作为模块 (*暂未实现)

有时候证明较长，用户书写的时候需要用到另一个文件里的定理等。这区别于标准模块和用户定义模块，因为这并不是领域的拓新，也不会产生新的运算符、类型、公理等，因此，这一类模块尽管也称为模块，但是解释器实现的时候仅仅是将两个文件顺序拼接当成一个新文件而已。

导入的示例（用引号以与之前的模块区分）：

```
import "logic_theorems" # import theorems in file 'logic_theorems.amk'
```

元素 Elements

类型

AMK本身并不支持类型（theorem, lemma, axiom除外），所有类型都来自于模块定义。

AMK提供元类型，即**list**, **set** 分别表示有序、无序的列表（有限数量的元素）。使用方式：

1. 在模块的源代码声明类型，如

```
list[typeA] set[typeB]
```

2. 运算符默认使用逗号，用

```
a, b
```

表示a, b两个元素组成的set，用

```
[a, b]
```

表示a, b两个元素组成的list。

运算符

AMK本身并没有运算符，所有的相关运算符都来自模块定义（除了numeric，它在模块中声明，一旦声明不需在模块中再定义即可使用）。

运算符的表示理论上可以使用任意符号、单词，甚至Unicode符号如 。暂时使用ASCII码的符号、字母等，后期可以加上运算符别名以适应不同输入环境下的用户需求。

公理、引理、定理

在AMK中，**theorem**, **lemma**, **axiom** 等关键字分别用作定理、引理和公理的声明。作这样的区分，主要是为了用户的书写习惯考虑，辅助用户理清楚要证明的是什么，什么是辅助。由于这些“x理”都有用户命名，因此在使用的时候可以不标明前缀是三种类型中的哪一种。

不失一般性，下面以**theorem**为例说明语法。定义方式类似于Python里的函数，首行为关键字**theorem**和定理名，后加冒号。之后两行分别用"**require:**"和"**conclude:**"标明已知条件和证明的结论。已知条件可以用

- 表达式；
- 变量的定义；
- 定理、引理等的名字或标号。

结论在表达清楚的情况下建议直接写表达式。当结论非常复杂的时候可以用缩进的**where**来声明变量代替长而重复出现的表达式（**where**特性暂未实现）。

再之后"**proof:**"后跟证明过程。具体的证明格式后面再详述。当声明了定理不跟**proof**引导的证明过程的时候，类似函数的声明，表示在后面会找到关于这个定理的证明。这一点是出于证明习惯的考虑，有时候我们会喜欢用分析法，从结论往回分阶段推倒，那么就需要留下一个假设等待后面证明。

下面是一个简单的例子（不要在意例子的具体内容）。

```
theorem A:
    require:
        P, R, Q
    conclude:
        a, b |- c
    where
        state a = R -> P Vee Q
        state b = not (R wedge Q)
        state c = R -> P
    proof:
        a, b |- c [lemma B]
```

注释

采用#作为注释的符号，注释内容从#开始一直到行末。可以看下面的例子：

```
I am not a comment # I am a comment
```

证明过程

下面是一个简单的证明示例：

```

theorem T:
  require:
    define a of statement
    define b of statement
  conclude:
    a -> b |- a -> (a wedge b)
  proof:
    a -> b, a |- a -[axiom belong] <1>
    a -> b, a |- a -> b <2> # using of axiom can be omitted
    a -> b, a |- b -[axiom ->- : <1><2>] <t3>
    a -> b, a |- a wedge b -[:<t3>] <4>
    a -> b |- a -> (a wedge b) [4]

```

首先来看写在定理 / 引理的 `proof:` 后面的内容，每行一句推理过程。

- 每行是一个已定义类型的表达式（或者变量），标示该表达式已经由 `require` 的内容和前面的推理步骤，在该定理内得到证明。
- （可选）尖括号 `<>` 内标示的是推理过程的标号，可以用数字、字母、下划线进行组合。为了方便书写，推理过程的标号仅在当前函数内有效。标号在一行的最右边（注释除外）。
- （可选）符号对 `-[和]` 内标示的是使用的定理、引理。
 - 如果是公理，可以不写清楚具体是哪条公理。但是定理、引理的使用必须明确指明其名字和应用的变量（这点和手写证明的要求是一样的）。
 - 括号内的内容可以用 `:` 分隔作为两部分
 - 前面是使用的定理 / 引理 / 公理的名字，前缀 `axiom` / `lemma` / `theorem` 可以省略。
 - 后面是对那几行推理得到的结果使用该定理 / 引理 / 公理，用标号按定理条件的顺序(不涉及推理的可以打乱顺序) 指明。
 - 如果只有前一半，可以省略 `:`，但是只有后一半不能省略 `:`

编写 Coding

缩进与括号

类似Python，采用缩进作为层次结构的标明，不使用大括号。定理的 `require`，`conclude`，`proof` 等都要相对 `theorem` 缩进，具体的 `where`，证明内容等也要相应缩进。

注意，这里缩进我们仅使用**TAB**。

大小写与命名规则

AMK大小写敏感，内置关键字等统一用小写。推荐用户使用小写字母和符号、数字等命名。变量、定理等的命名需要

高亮与编辑器支持

已有`.amk`和`.mamk`的最基本Vim语法高亮插件。安装方式：

```
$ cp vim-plugin/mamk.vim ~/.vim/syntax/
$ cp vim-plugin/amk.vim ~/.vim/syntax/
$ cp vim-plugin/filetype.vim ~/.vim/
```

支持各模块类型的高亮插件、自动补全插件后期可以补上。

使用

使用的时候只需要运行`amksh.py`即可，例如

```
$ ./amksh.py examples/infer1.amk
bison -d amk.y
flex amk.lex
gcc -o amki amk.tab.c lex.yy.c -O2
Compiling ... done.
=====
checking file `examples/infer1.amk` ... done

[correct 10:10] match conclusion part and requirement part
[correct 11:11] match conclusion part and requirement part
[warning 12:12] infer exactly 0 step(s) to get <x>: Your missing expression is the expression
[warning 12:12] infer exactly 0 step(s) to get <x>: Your missing expression is the expression
[correct 12:12] match conclusion part and requirement part
[correct 13:13] match conclusion part and requirement part
[warning 14:14] infer exactly 0 step(s) to get <x>: Your missing expression is the expression
[correct 14:14] match conclusion part and requirement part
[error 10:14] last line does not match its conclusion

=====

$
```

AMK Language Design: "Mathematical Logics" Module

By *Shuyang Shi* @ Oct, 2015

前言 Preface

我们希望，能从数理逻辑(Mathematical Logics)这一领域作为例子，实现AMK的第一个模块，从而证明该语言的可行性和实用性。相比分析等领域，逻辑较为简单，但是具备了基本的证明要素，十分适宜作为试水的模块。

We use Mathematical Logics (propositional logics) as an example, and as the first module of our AMK project, to prove the AMK project is practicable and utility. Unlike the analysis field, logic is simpler in expression but equipped with most elements of mathematical proofs, which make it suitable as the first module.

继承 Inheritance

继承AMK本身的所有语言特征。

All the features of AMK language.

具体的定义与说明 Specific Definition and Demonstration

命题逻辑部分，参见 `modules/logics/proposition.mamk`。

To view the specific definition of the propositional part, refer to
`modules/logics/proposition.mamk`.

Source Code Example

更多的例子可以见网站。

Visit our website to see more examples.

```
import logics.proposition

theorem belong2:
    require:
        define a of statement
        define b of statement
        define c of statement
        define d of statement
    conclude:
        a,b,c|-c
    proof:
        a,b,c|-a -[belong] <1>
        a,b,(a wedge b) wedge c|- (a wedge b) wedge c -[belong] <2>
        a,b,c|-c -[belong] <2>
```

```
import logics.proposition

theorem belong2:
    require:
        define a of statement
        define b of statement
        define c of statement
        define d of statement
    conclude:
        a,b,c|-c
    proof:
        a,b,c|-a -[belong] <1>
        a,b,a wedge b|- a wedge b -[belong] <2>
        a,b,c|-c -[belong] <2>

theorem belong3:
    require:
        define x of statement
        define y of statement
        define z of statement
        define p of statement
    conclude:
        x,y,z|-y
    proof:
        x,y,z|-y -[belong2] <2>
```

```
import logics.proposition

theorem zero_step_and_one_step_at_one_time:
    require:
        define a of statement
        define b of statement
    conclude:
        not a -> not b |- b -> a
proof:
    not a, not a -> not b, b |- not a -[belong]
    not a, not a -> not b, b |- not a -> not b -[belong]
#     not a, not a -> not b, b |- not b -[theorem contain_cancellation: ]
    not a, not a -> not b, b |- b -[belong]
    not a -> not b, b |- a -[theorem not_cancellation: ] # s are here
    not a -> not b |- b -> a -[contain_introduction]
```

Source Code Documentation

- [Lexer Design Documentation](#)
- [Parser Design Documentation](#)

Documentation:Lex

By *Shengkai Ni* @ 2015

Keyword

```
define import
therom axiom lemma
require conclude proof
where of
not wedge vee
```

Operator

Spaces in operators are not allowed.

```
not vee wedge
|- | -> <->
```

Identifier

Identifiers are strings that

1. starts with letters;
2. may contain letters, numbers, and '_'.

Label

Labels contain letters, numbers and '_'.

Labels are wrapped with '<' and '>'.

Spaces in labels are not allowed.

Here are some examples below.

```
<1>
<belong>
<_____yoooooo_____>
<3q_>
```

Other Symbol

The symbol to show which lemma or conclusion appeared before and now is used in this proof line.

```
-[]

Example:
a |- b -[theorem t: <1> <2>] <1>
```

Parrens.

```
(  
)
```

Comments

```
#  
Example:  
#This is a comment.
```

Square brackets

```
[  
]
```

Colons

```
:
```

Documentation: Syntax

By *Shuyang Shi* @ Nov 2015

Data Union

The union for data passing during **yylex()** is defined as follows.

```
/* union */
%union {
    char *str;
    struct ast_node *ptr;
};
```

AST Node

Struct Definition

struct ast_node is defined as follows.

```
/* node in AST */
struct ast_node {
    /* info of this node */
    enum node_types node_type;
    void * data;

    /* links */
    int num_links;
    struct ast_node ** links;
};
```

Explanation for Node Types

For different node types, we define the member variables as the following table indicates.

node_type	data	num_links	links
nd_program	NULL	2	import_part, proof_part
nd_import_part	NULL	n	import_expr
nd_proof_part	NULL	n	proof_block
nd_import_expr	(char *) str	0	NULL
nd_rich_expressions	NULL	n	rich_expr
nd_expressions	NULL	n	expr
nd_proof_block	(char *) name	3	proof_require, proof_conclude, proof_body / NULL
nd_rich_expr	(struct ast_node *) expr	2	theorem_ref, label
nd_ref_body	(char *) theorem_name	2	ref_pref, ref_labels
nd_ref_labels	NULL	n	identifier
nd_expr	(enum operators) op (NULL if none)	k = 1 or 2	(ast_node) expr / (char) var
nd_of_expr	var	1	type
nd_of_expressions	NULL	n	of_expr
nd_proof_req	NULL	2	of_expressions, expressions
nd_type	(char *) identifier / "set" / "list"	0 or 1	(sub-)type

Operator Precedence and Associativity

```
/* operator precedence and associativity */
%right dget
%right get
%right dcontain
%right contain
%right vee
%right wedge
%nonassoc not
```