Coq 定理证明器

1 整数算数运算与大小比较

鸡兔同笼问题：

```
Fact chickens_and_rabbits: forall C R: Z,
  C + R = 35 ->
  2 * C + 4 * R = 94 ->
  C = 23.
```

字面意思上，这个命题说的是：对于任意整数 `C` 与 `R`，如果 `C + R = 35` 并且 `2 * C + 4 * R = 94`，那么 `C = 23`。其中，`forall` 与 `->` 是 Coq 中描述数学命题的常用符号。

Coq 表达式 1. 全称量词 forall . 在 Coq 中，`forall` 表示“任意”的意思，例如：

```
forall x: Z, x = x
```

就是一个语法上合法的 Coq 命题，在这个例子中 `Z` 表示整数集合，`forall x: Z, ...` 说的就是“对于任意整数 `x`，某性质成立”。在 `forall` 之后，可以跟一个变量也可以跟多个变量，例如：

```
forall x y: Z, x + y = y + x
```

也是合法的 Coq 命题。另外，`forall` 后的类型标注不是必须的，如果 Coq 系统能够推导出这个类型，那么就可以省略它。Coq 还允许一些 `forall` 之后的变量有类型标注，而另一些没有，例如：

```
forall (x: Z) y, x + y = y + x
```
。

Coq 表达式 2. 表示命题推导的箭头符号 `->` . 在 Coq 中，箭头符号 `->` 表示“如果... 那么...”，例如：

```
x >= 0 -> x + 1 > 0
```

就表示如果 `x` 大于等于 0，那么 `x+1` 大于 0。Coq 规定，这个箭头符号是右结合的，换言之，`P1 -> P2 -> P3` 实际是 `P1 -> (P2 -> P3)` 的简写，其表达的意思是：如果 `P1` 成立，那么 `P2` 能推出 `P3`。逻辑上，这等同于：如果 `P1` 并且 `P2`，那么 `P3`。因此，我们一般会将形如 `P1 -> P2 -> ... -> Pn -> Q` 的命题读作：如果 `P1`、`P2`、...、`Pn` 都成立，那么 `Q` 也成立。

上面的 Coq 代码中，除了逻辑符号 `forall`、`->` 与算数符号之外，还使用了保留字 `Fact`，这是一种 Coq 指令。

Coq 指令 1. Fact 指令. 在 Coq 中，Fact 指令可以用于陈述一个命题。例如，`chickens_and_rabbits` 是在上面 Coq 代码中命题的名字，之后从 `forall` 开头的逻辑算数表达式是这个命题的内容，命题的名字与命题的内容之间用冒号分隔。Coq 系统规定，只要这个命题语法上合法，那么整个 Fact 指令就是合法的。换言之，Coq 不会在执行 Fact 指令的时候检查其声明的命题是不是真命题。不过，执行 Fact 指令之后，用户需要进入 Coq 证明环境证明该结论。在 Coq 中，还有一些保留字与 Fact 功能相同，它们是：Proposition、Example、Lemma、Theorem 与 Corollary。

在 Fact 指令之后，我们可以在 Coq 中证明这个数学命题成立。如果要用中学数学知识完成这一证明，恐怕要使用加减消元法、代入消元法等代数技巧。Coq 并不需要我们掌握这些数学技巧，Coq 系统可以自动完成整数线性运算性质（linear integer arithmetic，简称 lia）的证明，`chickens_and_rabbits` 这一命题在 Coq 中的证明只需一行：

```
Proof. lia. Qed.
```

在这一行代码中，Proof 和 Qed 表示一段证明的开头与结尾，在它们之间的 `lia` 指令是证明脚本。

一般而言，编写 Coq 证明的过程是交互式的——“交互式”的意思是：在编写证明代码的同时，我们可以在 Coq 定理证明环境中运行证明脚本，获得反馈，让定理证明系统告诉我们“已经证明了哪些结论”、“还需要证明哪些结论”等等，并以此为依据编写后续的证明代码。安装 VSCoq 插件的 VSCode 编辑器、安装 proof-general 插件的 emacs 编辑器以及 CoqIDE 都是成熟易用的 Coq 定理证明环境。

以上面的证明为例，执行 `lia` 指令前，证明窗口显示了当前还需证明的性质（亦称为证明目标，proof goal）：

```
------------------------------------------
(1/1)
forall C R : Z,
C + R = 35 -> 2 * C + 4 * R = 94 -> C = 23
```

这里横线上方的是目前可以使用的前提，横线下方的是目前要证明的结论，目前，前提集合为空。横线下方的 (1/1) 表示总共还有 1 个证明目标需要证明，当前显示的是其中的第一个证明目标。利用证明脚本证明的过程中，每一条证明指令可以将一个证明目标规约为 0 个，1 个或者更多的证明目标。执行 `lia` 指令之后，证明窗口显示：Proof finished。表示证明已经完成。一般情况下，Coq 证明往往是不能只靠一条证明指令完成证明的。

Coq 指令 2. Proof 指令与 Qed 指令. Proof 与 Qed 是一段证明的首尾标识，在它们之间的 Coq 指令都是证明脚本。在 Coq 中，用户通过证明脚本完成证明。一般情况下，Coq 证明脚本都能保证其进行的逻辑变换与逻辑规约都是合法的，特殊情况下，Coq 定理证明系统还需要在 Qed 指令时进行额外检验。经过 Qed 检验后，一个数学命题的 Coq 证明才算完成。

Coq 证明脚本 1. lia 指令. 证明指令 `lia` 表示自动证明有关整数线性运算与大小关系的性质，lia 这三个字母是 linear integer arithmetic 的缩写。证明指令 `lia` 是完备的，换言之，所有正确的整数线性运算性质都能够通过这一指令设定的算法完成自动证明。当然，在实际使用中，可能由于待证明的命题规模太大（变量个数太多、约束条件中的表达式太长或约束条件数量太多），算法所需运行时间太长，Coq 系统将其提前终止，因而无法完成自动证明。

Coq 证明指令 `lia` 除了能够证明关于线性运算的等式，也可以证明关于线性运算的不等式。下面这个例子选自熊斌教授所著《奥数教程》的小学部分：幼儿园的小朋友去春游，有男孩、女孩、男老师与女老师共 16 人，已知小朋友比老师人数多，但是女老师比女孩人数多，女孩又比男孩人数多，男孩比男老师人数多，请证明幼儿园只有一位男老师。

```
Fact teachers_and_children: forall MT FT MC FC: Z,
  MT > 0 ->
  FT > 0 ->
  MC > 0 ->
  FC > 0 ->
  MT + FT + MC + FC = 16 ->
  MC + FC > MT + FT ->
  FT > FC ->
  FC > MC ->
  MC > MT ->
  MT = 1.
Proof. lia. Qed.
```

习题 1. 请在 Coq 中描述下面结论并证明：如果今年甲的年龄是乙 5 倍，并且 5 年后甲的年龄是乙的 3 倍，那么今年甲的年龄是 25 岁。

除了线性性质之外，Coq 中还可以证明的一些更复杂的性质。例如下面就可以证明，任意两个整数的平方和总是大于它们的乘积。证明中使用的指令`nia`表示的是非线性整数运算（nonlinear integer arithmetic）求解。

```
Fact sum_of_sqr1: forall x y: Z,
  x * x + y * y >= x * y.
Proof. nia. Qed.
```

不过，`nia`与`lia`不同，后者能够保证关于线性运算的真命题总能被自动证明（规模过大的除外），但是有不少非线性的算数运算性质是`nia`无法自动求解的。例如，下面例子说明，一些很简单的命题`nia`都无法完成自动验证。

```
Fact sum_of_sqr2: forall x y: Z,
  x * x + y * y >= 2 * x * y.
Proof. Fail nia. Abort.
```

这是就需要我们通过编写证明脚本，给出中间证明步骤。证明过程中，可以使用 Coq 标准库中的结论，也可以使用我们自己实现证明的结论。例如，Coq 标准库中，`sqr_pos`定理证明了任意一个整数`x`的平方都是非负数，即：

```
sqr_pos: forall x: Z, x * x >= 0
```

我们可以借助这一性质完成上面`sum_of_sqr2`的证明。

```
Fact sum_of_sqr2: forall x y: Z,
  x * x + y * y >= 2 * x * y.
Proof.
  intros.
  pose proof sqr_pos (x - y).
  nia.
Qed.
```

这段证明有三个证明步骤。证明指令`intros`将待证明结论中的逻辑结构“对于任意整数 x 与 y”移除，并在前提中的“x 与 y 是整数”的假设。第二条指令`pose proof`表示对 x-y 使用标准库中的定理`sqr_pos`。从 Coq 定理证明界面中可以看到，使用该定理得到的命题会被添加到证明目标的前提中去，Coq 系统将这个新命题自动命名为 H（表示 Hypothesis）。最后，`nia`可以自动根据当前证明目标中的前提证明结论。

下面证明演示了如何使用我们刚刚证明的性质`sum_of_sqr1`。

```
Example quad_ex1: forall x y: Z,
  x * x + 2 * x * y + y * y + x + y + 1 >= 0.
Proof.
  intros.
  pose proof sum_of_sqr1 (x + y) (-1).
  nia.
Qed.
```

下面这个例子说的是：如果`x < y`，那么`x * x + x * y + y * y`一定大于零。

```
Fact sum_of_sqr_lt: forall x y: Z,
  x < y ->
  x * x + x * y + y * y > 0.
```

我们可以利用下面两式相等证明：

```
4 * (x * x + x * y + y * y)
3 * (x + y) * (x + y) + (x - y) * (x - y)
```

于是，在`x < y`的假设下，等式右边的两个平方式一个恒为非负，一个恒为正。因此，等式的左边也恒为正。

```
Proof.
  intros.
  pose proof sqr_pos (x + y).
  nia.
Qed.
```

可以看到，在`x < y`的前提下，Coq 的`nia`指令可以自动推断得知`(x - y)`的平方恒为正。不过，我们仍然需要手动提示 Coq，`(x + y)`的平方恒为非负。

Coq 证明脚本 2. intros 指令. 证明指令`intros`表示将待证明结论中的假设移动到证明目标的前提中去。例如，在上面`sum_of_sqr_lt`中，`intros`指令移动了三项前提：`x: Z`、`y: Z`与`H: x < y`。其中`H`是 Coq 定理证明系统自动引入的命名，字母 H 表示 Hypothesis 的简写，当`intros`要添加若干个命题作为前提的时候，Coq 会依次选择`H`、`H0`、`H1`等名字。有时，我们在 Coq 中编写证明脚本代码时，希望能够手动控制这些前提的命名，这只需要在`intros`后添加参数就可以了。例如，`sum_of_sqr_lt`中的`intros`指令就等效于`intros x y H`。Coq 允许我们在`intros`的同时对`forall`后的变量重命名，例如，将`sum_of_sqr_lt`中的`intros`指令改为`intros x1 x2 H`后效果如下。Coq 也允许对一部分前提手动命名，而同时对另一部分前提自动命名，只需用问号占位符表示自动命名的前提即可，例如`intros ? ? H`。

Coq 证明脚本 3. pose proof 指令. 证明指令`pose proof`表示在当前证明中使用一条已经证明过定理或者使用当前证明目标中的一条前提。例如，标准库中已有定理`sqr_pos`

```
sqr_pos: forall x: Z, x * x >= 0
```

那么`pose proof sqr_pos (x + 1)`就会得到`(x + 1) * (x + 1) >= 0`。类似的，假设当前证明目标中有下述前提，

```
x: Z
H: x >= 0
H0: x >= 0 -> x + 1 > 0
```

那么，就可以通过`pose proof H H0`得到`x + 1 > 0`。另外，使用`pose proof`指令时未必需要将所有的前提全部填上，如 Coq 标准库中的`Zmult_ge_compat_r`是下面定理：

```
forall n m p : Z, n >= m -> p >= 0 -> n * p >= m * p
```

假设当前证明目标中有下述前提，

```
k1: Z
k2: Z
x: Z
H: k1 >= k2
H0: x * x >= 0
```

那么，就可以通过以下`pose proof`指令

```
pose proof Zmult_ge_compat_r k1 k2 (x * x) H
pose proof Zmult_ge_compat_r k1 k2 (x * x) H H0
pose proof Zmult_ge_compat_r (x * x) 0 5 H0 ltac:(lia)
```

分别得到以下结论：

```
x * x >= 0 -> k1 * (x * x) >= k2 * (x * x)
k1 * (x * x) >= k2 * (x * x)
x * x * 5 >= 0 * 5
```

可以看到，填写前提中的命题部分时，既可以填写已有前提的名称（如`H`、`H0`等），也可以填写一条证明指令，如`ltac:(lia)`。除此之外，如果`pose proof`指令的一些参数可以由另一些参数推导出来，那么可以用下划线省去这些参数。例如，下面这几条证明指令的效果和上面证明指令的效果时相同的。

```
pose proof Zmult_ge_compat_r _ _ (x * x) H
pose proof Zmult_ge_compat_r _ _ _ H H0
pose proof Zmult_ge_compat_r _ _ 5 H0 ltac:(lia)
```

最后，在 Coq 中还可以指明`pose proof`所生成新命题的名称。例如，

```
pose proof Zmult_ge_compat_r _ _ 5 H0 ltac:(lia) as H5xx
```

得到的新命题是：`H5xx: x * x * 5 >= 0 * 5`。

Coq 证明脚本 4. nia 指令. 证明指令`nia`表示自动证明有关整数非线性算数运算的性质，nia 这三个字母是 nonlinear integer arithmetic 的缩写。证明指令`nia`是不完备的，但是它能够自动完成多项式的展开与线性性质的推理。另外，它也能自动推到乘法与正负数之间的关系。

2 函数与谓词

函数是一类重要的数学对象。例如，“加一”这个函数往往可以写作：$f(x) = x + 1$。在 Coq 中，我们可以用以下代码将其定义为`plus_one`函数。

```
Definition plus_one (x: Z): Z := x + 1.
```

在类型方面，`plus_one (x: Z): Z`表示这个函数的自变量和值都是整数，而`:=`符号右侧的表达式`x + 1`也描述了函数值的计算方法。

我们知道，“在 1 的基础上加一”结果是 2，“在 1 的基础上加一再加一”结果是 3。这些简单论断都可以用 Coq 命题表达出来并在 Coq 中证明。

```
Example One_plus_one: plus_one 1 = 2.
Proof. unfold plus_one. lia. Qed.
```

```
Example One_plus_one_plus_one: plus_one (plus_one 1) = 3.
Proof. unfold plus_one. lia. Qed.
```

Coq 表达式 3. 包含函数的表达式. 在 Coq 中，某函数`F`作用于某参数`X`写作`F X`，不需要写括号。这一语法类似于 Ocaml 等函数式编程语言。另外，这一语法是左结合的。换言之，表达式`F X Y`是`(F X) Y`的简写，而表达`F (G X)`时必须添加括号。

Coq 证明脚本 5. unfold 指令. 证明指令`unfold`表示在待证明的结论中展开某项定义。如果要在证明目标的某前提`H`中展开`X`的定义，可以使用证明指令`unfold X in H`。

下面是更多函数的例子，我们可以采用类似的方法定义平方函数。

```
Definition square (x: Z): Z := x * x.
```

```
Example square_5: square 5 = 25.
Proof. unfold square. lia. Qed.
```

Coq 中也可以定义多元函数。

```
Definition smul (x y: Z): Z := x * y + x + y.
```

```
Example smul_ex1: smul 1 1 = 3.
Proof. unfold smul. lia. Qed.
```

```
Example smul_ex2: smul 2 3 = 11.
Proof. unfold smul. lia. Qed.
```

Coq 表达式 4. 包含多元函数的表达式. Coq 中的二元函数实质上是接收一个参数后会计算得到一个一元函数的函数。例如，当`F`是一个二元函数时，我们通常将“`F`作用于`X`与`Y`的结果”写作`F X Y`，即`(F X) Y`的简写。这是因为`F X`实质上是一个一元函数，当他再接收一个参数`Y`之后的计算结果就是`(F X) Y`。类似的，Coq 中的三元函数实质上是接收一个参数后会计算得到一个二元函数的函数；Coq 中的 n+1 元函数实质上是接收一个参数后会计算得到一个 n 元函数的函数。

下面 Coq 代码定义了“非负”这一概念。在 Coq 中，可以通过定义类型为`Prop`的函数来定义谓词。以下面定义为例，对于每个整数`x`，`:=`符号右侧表达式`x >= 0`是真还是假决定了`x`是否满足性质`nonneg`（即，非负）。

```
Definition nonneg (x: Z): Prop := x >= 0.
```

```
Fact nonneg_plus_one: forall x: Z,
  nonneg x -> nonneg (plus_one x).
Proof. unfold nonneg, plus_one. lia. Qed.
```

```
Fact nonneg_square: forall x: Z,
  nonneg (square x).
Proof. unfold nonneg, square. nia. Qed.
```

习题 2. 请在 Coq 中证明下面结论。

```
Fact nonneg_smul: forall x y: Z,
  nonneg x -> nonneg y -> nonneg (smul x y).
(* 请在此处填入你的证明，以_[Qed]_结束。 *)
```