

# A Case for Completion Modulo Equivalence

Kristoffer H. Rose

IBM Thomas J. Watson Research Center

*1st International Workshop on Confluence*  
May 29, 2012, Nagoya, Japan

We present simple use cases from compiler writing for the use of completion modulo an equational theory, corresponding to the semantics of the target language.

## Disclaimer

The information contained in this presentation is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided "as is", without warranty of any kind, express or implied. In addition, this information is based on IBM's current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other documentation. Nothing contained in this presentation is intended to, or shall have the effect of:

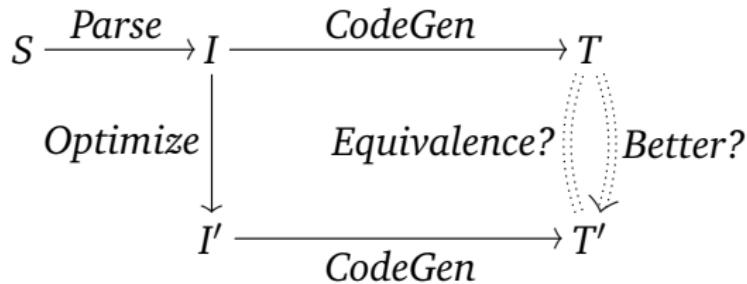
- creating any warranty or representation from IBM (or its affiliates or its or their suppliers and/or licensors); or
- altering the terms and conditions of the applicable license agreement governing the use of IBM software.

CRSX is Higher Order Rewriting engine with special support for writing compilers.

- RTA 1996 free variable matching rules (w/Haskell interpreter)
- XML 2005 starts as internal optimizer in Java impl. of XPath
- HOR 2007 as executable HOR tool
- IBM 2008 Compiler specification language started
- HOR 2010 “Compromised Rewriting Systems” described as compiler source language
- RTA 2011 Environment extension and polymorphic sorts “tricks”
- HOR 2012 Issues for Real Programmers “+ Blessings of Completion”

See <http://crsx.sf.net>.

# Compiling by Rewriting



## 1 Optimization by Completion

- Peano
- Compilation

## 2 Conclusion

## 1 Optimization by Completion

- Peano
- Compilation

## 2 Conclusion

$N ::= Z \mid S[N] ;$   
 $OP ::= + \mid \times ;$

$E ::= Op[OP, E, E] \mid Nat[N] ;$

$C ::= list[INS] ;$   
 $INS ::= OP[OP] \mid PUSH[N] ;$

Note:  $list[INS]$  has members  $(INS; \dots; INS;)$

$N ::= Z \mid S[N] ;$

$OP ::= + \mid \times ;$

$E ::= Op[OP, E, E] \mid Nat[N] ;$

$C ::= list[INS] ;$

$INS ::= OP[OP] \mid PUSH[N] ;$

Note:  $list[INS]$  has members  $(INS; \dots; INS;)$

$N ::= Z \mid S[N] ;$

$OP ::= + \mid \times ;$

$E ::= Op[OP, E, E] \mid Nat[N] ;$

$C ::= \mathbf{list}[INS] ;$

$INS ::= OP[OP] \mid PUSH[N] ;$

Note:  $\mathbf{list}[INS]$  has members  $(INS; \dots; INS;)$

$T[E, C] :: C ;$

$T[\text{Nat}[N], C] \rightarrow (\text{PUSH}[N]; C) ;$

$T[\text{Op}[\text{OP}, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (\text{OP}[\text{OP}]; C)]] ;$

## Sample

$T[\text{Op}[\times, \text{Nat}[S[Z]], \text{Op}[+, \text{Nat}[Z], \text{Nat}[S[S[Z]]]]], 0]$

$\rightarrow T[\text{Nat}[S[Z]], T[\text{Op}[+, \text{Nat}[Z], \text{Nat}[S[S[Z]]]], (\text{OP}[\times];)]]$

$\rightarrow (\text{PUSH}[S[Z]]; T[\text{Nat}[Z], T[\text{Nat}[S[S[Z]]]], (\text{OP}[+]; \text{OP}[\times];)])$

$\rightarrow \rightarrow (\text{PUSH}[S[Z]]; \text{PUSH}[Z]; \text{PUSH}[S[S[Z]]]; \text{OP}[+]; \text{OP}[\times];)$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times]; )]]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]]], (OP[+]; OP[\times]; )])$

$\rightarrow \rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times]; )$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times];)])]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]]], (OP[+]; OP[\times];)])$

$\rightarrow \rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times];)$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times]; )]]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]], (OP[+]; OP[\times]; )]])$

$\rightarrow \rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times]; )$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times]; )]]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]], (OP[+]; OP[\times]; )]])$

$\rightarrow \rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times]; )$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times]; )]]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]]], (OP[+]; OP[\times]; )]])$

$\rightarrow \rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times]; )$

$T[E, C] :: C ;$

$T[Nat[N], C] \rightarrow (PUSH[N]; C) ;$

$T[Op[OP, E_1, E_2], C] \rightarrow T[E_1, T[E_2, (OP[OP]; C)]] ;$

## Sample

$T[Op[\times, Nat[S[Z]], Op[+, Nat[Z], Nat[S[S[Z]]]]], 0]$

$\rightarrow T[Nat[S[Z]], T[Op[+, Nat[Z], Nat[S[S[Z]]]], (OP[\times]; )]]$

$\rightarrow (PUSH[S[Z]]; T[Nat[Z], T[Nat[S[S[Z]]]], (OP[+]; OP[\times]; )])$

$\rightarrow (PUSH[S[Z]]; PUSH[Z]; PUSH[S[S[Z]]]; OP[+]; OP[\times]; )$

# Peano Optimization

$\text{Op}[+, \text{Nat}[Z], N] \rightarrow N ;$

$\text{Op}[\times, \text{Nat}[S[Z]], N] \rightarrow N ;$

- [Discard[N]]:

$\text{Op}[\times, \text{Nat}[Z], N] \rightarrow \text{Nat}[Z] ;$

## Sample

$T[\text{Op}[\times, \text{Nat}[S[Z]], \text{Op}[+, \text{Nat}[Z], \text{Nat}[S[S[Z]]]]], 0]$

$\rightarrow T[\text{Op}[+, \text{Nat}[S[S[Z]]], \text{Nat}[Z]], 0]$

$\rightarrow T[\text{Nat}[S[S[Z]]], 0]$

$\rightarrow (\text{PUSH}[S[S[Z]]];)$

# Peano Optimization

$\text{Op}[+, \text{Nat}[Z], N] \rightarrow N ;$

$\text{Op}[\times, \text{Nat}[S[Z]], N] \rightarrow N ;$

-[Discard[N]]:

$\text{Op}[\times, \text{Nat}[Z], N] \rightarrow \text{Nat}[Z] ;$

## Sample

$T[\text{Op}[\times, \text{Nat}[S[Z]], \text{Op}[+, \text{Nat}[Z], \text{Nat}[S[S[Z]]]]], 0]$

$\rightarrow T[\text{Op}[+, \text{Nat}[S[S[Z]]], \text{Nat}[Z]], 0]$

$\rightarrow T[\text{Nat}[S[S[Z]]], 0]$

$\rightarrow (\text{PUSH}[S[S[Z]]];)$

# Peano Optimization

$\text{Op}[+, \text{Nat}[Z], N] \rightarrow N ;$

$\text{Op}[\times, \text{Nat}[S[Z]], N] \rightarrow N ;$

-[Discard[N]]:

$\text{Op}[\times, \text{Nat}[Z], N] \rightarrow \text{Nat}[Z] ;$

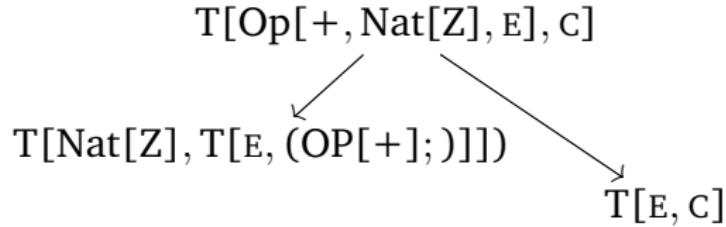
## Sample

$T[\text{Op}[\times, \text{Nat}[S[Z]], \text{Op}[+, \text{Nat}[Z], \text{Nat}[S[S[Z]]]]], 0]$

$\rightarrow T[\text{Op}[+, \text{Nat}[S[S[Z]]], \text{Nat}[Z]], 0]$

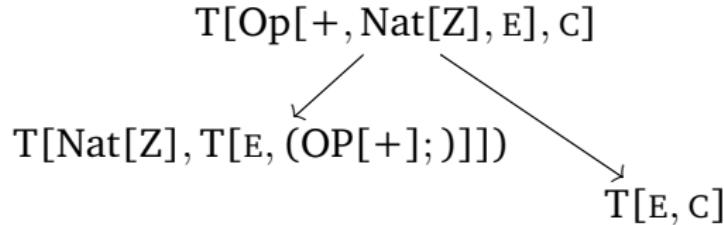
$\rightarrow T[\text{Nat}[S[S[Z]]], 0]$

$\rightarrow (\text{PUSH}[S[S[Z]]];)$



### Observation

One branch is a “data reduction.”



### Observation

One branch is a “data reduction.”

$T[Op[+, Nat[Z], E_2], c] \rightarrow T[E_2, c] ;$

$T[Op[+, Nat[S[N_1]], E_2], c] \rightarrow T[Nat[S[N_1]], T[E_2, (OP[+]; c)]] ;$

-[Discard[ $E_2$ ]]:

$T[Op[\times, Nat[Z], E_2], c] \rightarrow (PUSH[Z]; c) ;$

$T[Op[\times, Nat[S[Z]], E_2], c] \rightarrow T[E_2, c] ;$

$T[Op[\times, Nat[S[S[N_1]]], E_2], c] \rightarrow T[Nat[S[S[N_1]]], T[E_2, (OP[\times]; c)]] ;$

$T[Op[+, Nat[Z], E_2], c] \rightarrow T[E_2, c] ;$

$T[Op[+, Nat[S[N_1]], E_2], c] \rightarrow T[Nat[S[N_1]], T[E_2, (OP[+]; c)]] ;$

$- [Discard[E_2]] :$

$T[Op[\times, Nat[Z], E_2], c] \rightarrow (PUSH[Z]; c) ;$

$T[Op[\times, Nat[S[Z]], E_2], c] \rightarrow T[E_2, c] ;$

$T[Op[\times, Nat[S[S[N_1]]], E_2], c] \rightarrow T[Nat[S[S[N_1]]], T[E_2, (OP[\times]; c)]] ;$

$E ::= v \mid \text{Int}[\mathbf{integer}] \mid \text{Str}[\mathbf{String}] \mid \text{Nil} \mid \text{Seq}[E, E]$   
 $\mid \text{IfNil}[E, E, E] \mid \text{Let}[E, x: E.E] ;$

$A ::= \mathbf{list}[AI] ;$

$AI ::= L[AL] \mid J[AL] \mid PU \mid POJN[AL] \mid POR[AR]$   
 $\mid SR[AR] \mid SI[\mathbf{integer}] \mid SS[\mathbf{String}] ;$

$AL ::= label ;$

$AR ::= register ;$

# Compilation Rules

Idea:  $(PU; \{in: r_i; out: r_o\}T[E, (POR[r_o]);])$

$\{\rho\} = \{E: AR\}$  ;

$\{\rho\} T[E, A] :: A$  ;

$\{v: r\} T[v, A] \rightarrow (SR[r]; A)$  ;

$\{\neg v\} T[v, A] \rightarrow (SS["Error"]); A$  ;

$\{\rho\} T[Int[I], A] \rightarrow (SI[I]; A)$  ;

$\{\rho\} T[Str[s], A] \rightarrow (SS[s]; A)$  ;

$\{\rho\} T[Nil, A] \rightarrow A$  ;

$\{\rho\} T[Seq[E_1, E_2], A] \rightarrow \{\rho\} T[E_1, \{\rho\} T[E_2, A]]$  ;

$\neg [Fresh[f, fi]]$  :

$\{\rho\} T[If[E, E_1, E_2], A]$

$\rightarrow (PU; \{\rho\} T[E, (POJN[f]; \{\rho\} T[E_1, (J[fi]; L[f]); \{\rho\} T[E_2, (L[fi]; A)])]))$  ;

$\neg [Fresh[x, r]]$  :

$\{\rho\} T[Let[E_1, v_1.E_2[v_1]], A] \rightarrow (PU; \{\rho\} T[E_1, (POR[r]; \{\rho; x: r\} T[E_2[x], A]))$  ;

# Compilation Rules

Idea:  $(PU; \{in: r_i; out: r_o\}T[E, (POR[r_o]);])$

$\{\rho\} = \{E: AR\}$  ;

$\{\rho\} T[E, A] :: A$  ;

$\{v: r\} T[v, A] \rightarrow (SR[r]; A)$  ;

$\{\neg v\} T[v, A] \rightarrow (SS["Error"]; A)$  ;

$\{\rho\} T[Int[I], A] \rightarrow (SI[I]; A)$  ;

$\{\rho\} T[Str[s], A] \rightarrow (SS[s]; A)$  ;

$\{\rho\} T[Nil, A] \rightarrow A$  ;

$\{\rho\} T[Seq[E_1, E_2], A] \rightarrow \{\rho\} T[E_1, \{\rho\} T[E_2, A]]$  ;

$\neg [Fresh[f, fi]]$  :

$\{\rho\} T[If[E, E_1, E_2], A]$

$\rightarrow (PU; \{\rho\} T[E, (POJN[f]; \{\rho\} T[E_1, (J[fi]; L[f]; \{\rho\} T[E_2, (L[fi]; A)]))])$  ;

$\neg [Fresh[x, r]]$  :

$\{\rho\} T[Let[E_1, v_1.E_2[v_1]], A] \rightarrow (PU; \{\rho\} T[E_1, (POR[r]; \{\rho; x: r\} T[E_2[x], A]))$  ;

# Compilation Optimization



If[Nil, E<sub>1</sub>, E<sub>2</sub>] → E<sub>2</sub> ;

Let[E, x.x] → E ;

Let[E<sub>1</sub>, x<sub>1</sub>.E<sub>2</sub>[]] → E<sub>2</sub> ;

(PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[Let[If[Nil, Nil, in], x.x], (POR[r<sub>o</sub>];)])  
→ (PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[If[Nil, Nil, in], (POR[r<sub>o</sub>];)])  
→ (PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[in, (POR[r<sub>o</sub>];)])  
→ (PU; SR[r<sub>in</sub>]; POR[r<sub>o</sub>];)

# Compilation Optimization



If[Nil, E<sub>1</sub>, E<sub>2</sub>] → E<sub>2</sub> ;

Let[E, x.x] → E ;

Let[E<sub>1</sub>, x<sub>1</sub>.E<sub>2</sub>[]] → E<sub>2</sub> ;

(PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[Let[If[Nil, Nil, in], x.x], (POR[r<sub>o</sub>];)])  
→ (PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[If[Nil, Nil, in], (POR[r<sub>o</sub>];)])  
→ (PU; {in: r<sub>i</sub>; out: r<sub>o</sub>}T[in, (POR[r<sub>o</sub>];)])  
→ (PU; SR[r<sub>in</sub>]; POR[r<sub>o</sub>];)

# Compilation Optimization



$\{s\}T[\text{If}[\text{Nil}, E_1, E_2], A] \rightarrow \{s\}T[E_2, A]$

-[Fresh[f, fi]] :

$\{s\}T[\text{If}[E \not\equiv \text{Nil}, E_1, E_2], A]$

$\rightarrow (\text{PU}; \{s\}T[E, (\text{POJN}[f]); \{s\}T[E_1, (\text{J}[fi]; \text{L}[f]); \{s\}T[E_2, (\text{L}[fi]; A))])]$

$\{s\}T[\text{Let}[E, x.x], A] \rightarrow \{s\}T[E, A] ;$

$\{s\}T[\text{Let}[E_1, x_1.E_2[]], A] \rightarrow E_2, A] ;$

-[Fresh[x, r]] :

$\{s\}T[\text{Let}[E_1, v_1.E_2[!v_1] \not\equiv v_1], A] \rightarrow (\text{PU}; \{s\}T[E_1, (\text{POR}[r]; \{s!x:r\}T[E_2]))]$

# Compilation Optimization



$\{s\}T[\text{If}[\text{Nil}, E_1, E_2], A] \rightarrow \{s\}T[E_2, A]$

-[Fresh[f, fi]] :

$\{s\}T[\text{If}[E \not\equiv \text{Nil}, E_1, E_2], A]$

$\rightarrow (\text{PU}; \{s\}T[E, (\text{POJN}[f]; \{s\}T[E_1, (\text{J}[fi]; \text{L}[f]; \{s\}T[E_2, (\text{L}[fi]; A)]))])$

$\{s\}T[\text{Let}[E, x.x], A] \rightarrow \{s\}T[E, A] ;$

$\{s\}T[\text{Let}[E_1, x_1.E_2[]], A] \rightarrow E_2, A] ;$

-[Fresh[x, r]] :

$\{s\}T[\text{Let}[E_1, v_1.E_2[!v_1] \not\equiv v_1], A] \rightarrow (\text{PU}; \{s\}T[E_1, (\text{POR}[r]; \{s !x: r\}T[E_2]))])$

## 1 Optimization by Completion

- Peano
- Compilation

## 2 Conclusion

- Compilers are complete modulo the target language semantic equations.
- Optimizations are intermediate language (data) rewrites.
- Completion with priority to data rewrites integrates translation schemes and optimization rules.
- Peano integers never get boring!

- Compilers are complete modulo the target language semantic equations.
- Optimizations are intermediate language (data) rewrites.
- Completion with priority to data rewrites integrates translation schemes and optimization rules.
- Peano integers never get boring!

- Compilers are complete modulo the target language semantic equations.
- Optimizations are intermediate language (data) rewrites.
- Completion with priority to data rewrites integrates translation schemes and optimization rules.
- Peano integers never get boring!

- Compilers are complete modulo the target language semantic equations.
- Optimizations are intermediate language (data) rewrites.
- Completion with priority to data rewrites integrates translation schemes and optimization rules.
- Peano integers never get boring!

Thank You!