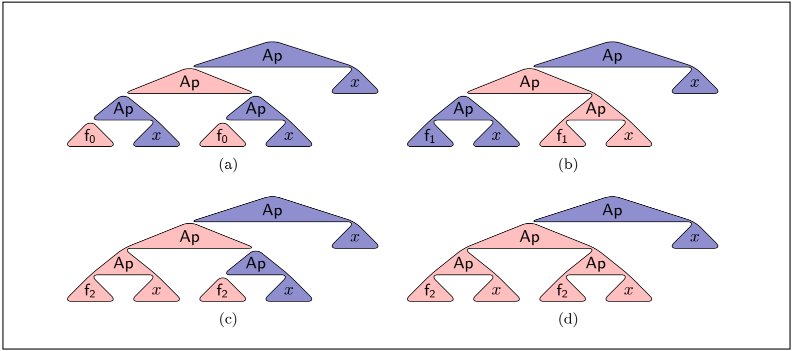
Proof by Picture?!

Vincent van Oostrom

University of Innsbruck

Master Seminar, Wednesday April 11, 2018

Some pictures from rewriting literature



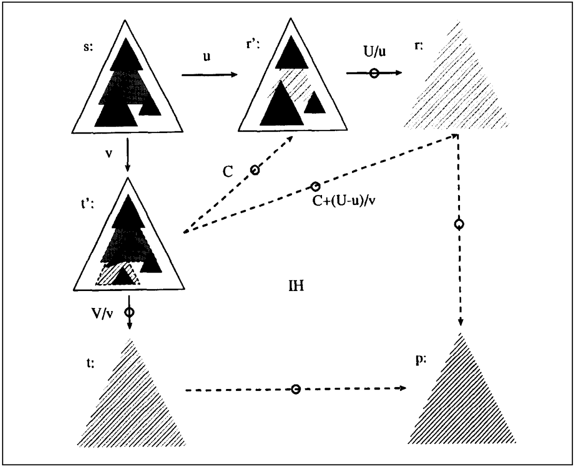
(a)

(b)

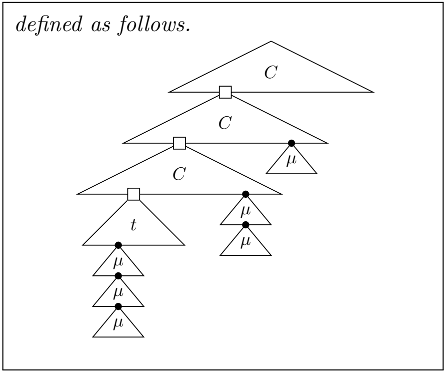
(c)

(d)

Some pictures from rewriting literature

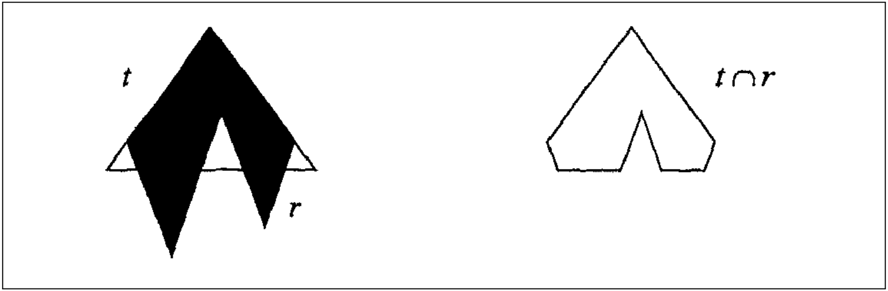


Some pictures from rewriting literature

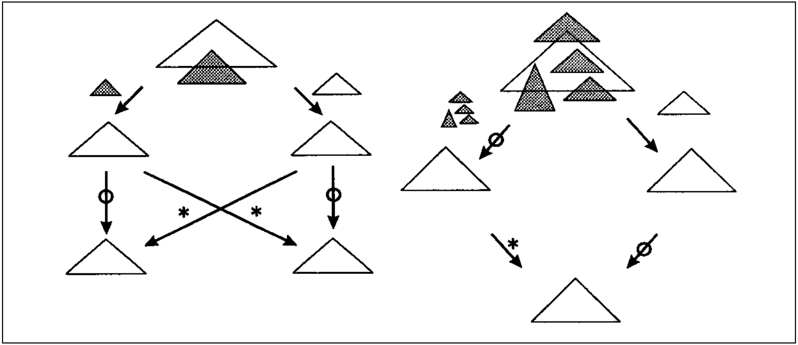


defined as follows.

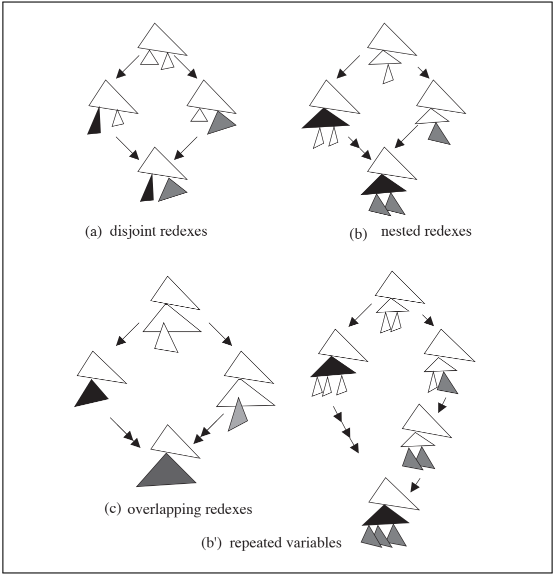
Some pictures from rewriting literature



Some pictures from rewriting literature



Some pictures from rewriting literature



(a) disjoint redexes

(b) nested redexes

(c) overlapping redexes

(b') repeated variables

Some pictures from rewriting literature

some common features of pictures

- identifying parts of terms
- overlap between parts of terms (geometric)
- combining steps on parts of terms (inductive)

note: parts not subterms

why pictures?

Some pictures from rewriting literature

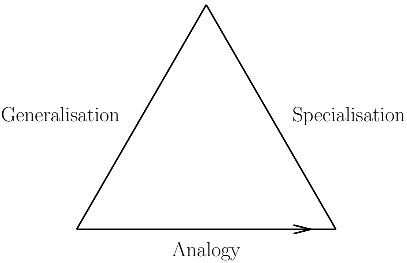
some common features of pictures

- identifying parts of terms
- overlap between parts of terms (geometric)
- combining steps on parts of terms (inductive)

note: parts not subterms

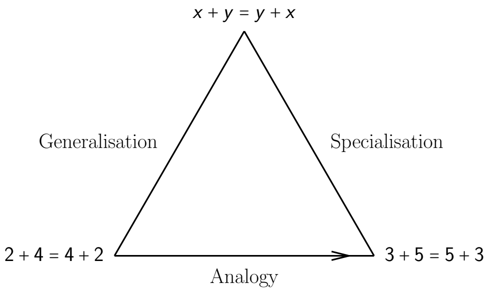
why pictures?
this talk: expressive language for parts of terms

Pólya's triangle



Pólya, Induction and Analogy in Mathematics, 1954, Fig. 2.3

Pólya’s triangle



$x + y = y + x$

Generalisation

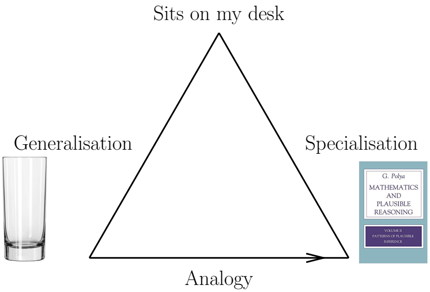
Specialisation

$2 + 4 = 4 + 2$

$3 + 5 = 5 + 3$

Analogy

Pólya's triangle



Sits on my desk

Generalisation

Specialisation

Analogy

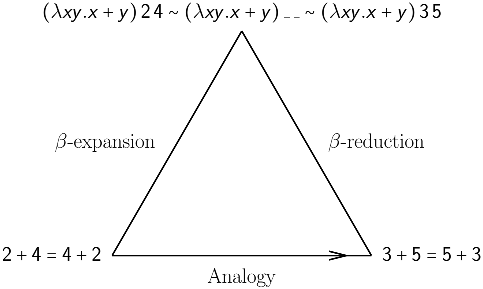
G. Polya

MATHEMATICS AND PLAUSIBLE REASONING

VOLUME II PATTERNS OF PLAUSIBLE INFERENCE

dependent on representation (cf. string, term, graph)

Pólya's triangle



$(\lambda xy.x + y)\,2\,4 \sim (\lambda xy.x + y)\,\_\,\_ \sim (\lambda xy.x + y)\,3\,5$

β-expansion

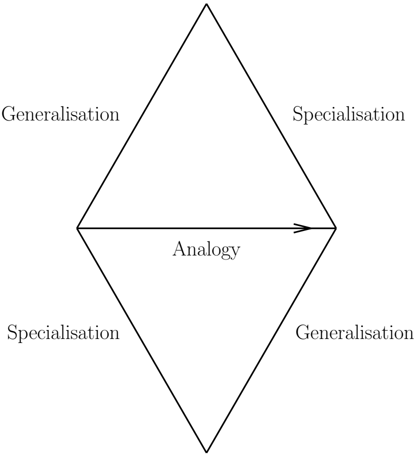
β-reduction

$2 + 4 = 4 + 2$

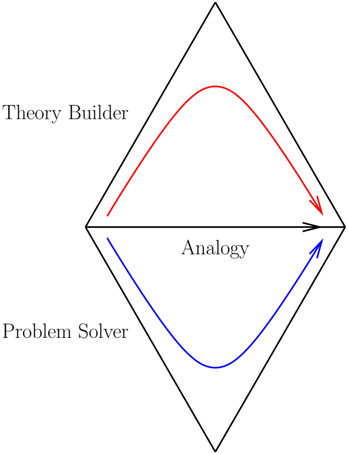
$3 + 5 = 5 + 3$

Analogy

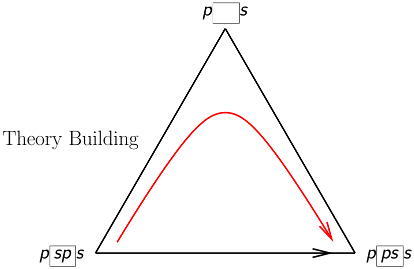
Pólya's triangle



Pólya's triangle

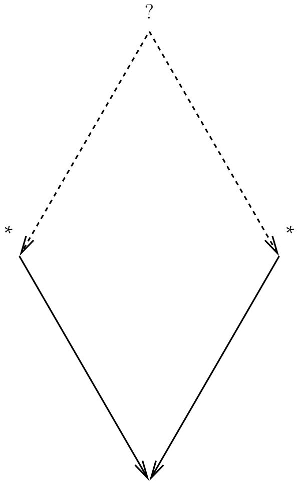


Pólya's triangle



this talk: Theory building, parts of terms theory

Pólya’s triangle

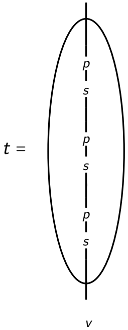


For problem solvers: is (linear) $\lambda\beta$-calculus upward confluent?

Representation?

Example (Running)

$t = p(s(p(s(p(s(v))))))$

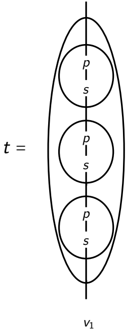


where we want to identify some parts of t

Representation?

Example (Running)

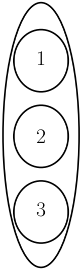
$t = p(s(p(s(p(s(v))))))$



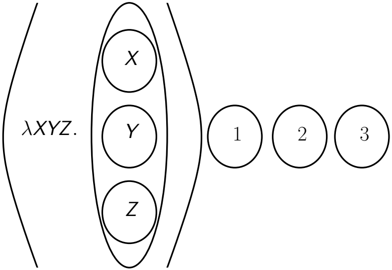
say all three consecutive $p - s$ pairs

Example (Running)

$t = pspsps$



Representation: λ-abstraction/β-expansion

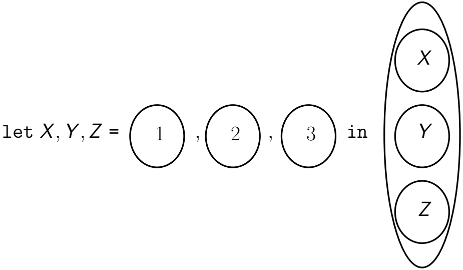


$$(\lambda XYZ.X(Y(Z(v)))(\lambda x.p(s(x)))(\lambda x.p(s(x)))(\lambda x.p(s(x)))$$

represents faithfully, but

- too powerful (repeat β-expansion)?
- universal algebra? (repeated expansion gives higher-order)?
- geometric operations? (union, intersection)
- λ's are scary

Representation: let-expressions (inductive clusters)

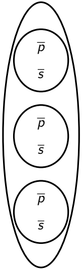


$$\texttt{let}\ X, Y, Z = p(s(v_1)), p(s(v_1)), p(s(v_1))\ \texttt{in}\ X(Y(Z(v)))$$

represents faithfully:

- universal algebra
  (two algebras, for body and for let-block)
- no λ's
  (X,Y,Z of arity 1 so only 1 parameter v_1)
- but geometric operations? (union, intersection)

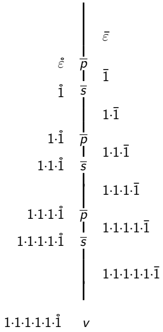
Representation: labelling/overlining



$$\overline{p}(\overline{s}(\overline{p}(\overline{s}(\overline{p}(\overline{s}(\mathsf{v}))))))$$

- loses distinguishing power (1 part vs. 3 parts)?
  (terms not trees; vertices explicit but edges implicit, no label)
- different labels for different parts?
  (but then how about union, intersection?)
- labelling changes signature?
- labels encode coherence of part locally?

Representation: sets of positions (geometric clusters)



$\bar{\varepsilon}$
$\mathring{\varepsilon}$
$\bar{p}$
$\bar{1}$
$\mathring{1}$
$\bar{s}$
$1{\cdot}\bar{1}$
$1{\cdot}\mathring{1}$
$\bar{p}$
$1{\cdot}1{\cdot}\bar{1}$
$1{\cdot}1{\cdot}\mathring{1}$
$\bar{s}$
$1{\cdot}1{\cdot}1{\cdot}\bar{1}$
$1{\cdot}1{\cdot}1{\cdot}\mathring{1}$
$\bar{p}$
$1{\cdot}1{\cdot}1{\cdot}1{\cdot}\bar{1}$
$1{\cdot}1{\cdot}1{\cdot}1{\cdot}\mathring{1}$
$\bar{s}$
$1{\cdot}1{\cdot}1{\cdot}1{\cdot}1{\cdot}\bar{1}$
$1{\cdot}1{\cdot}1{\cdot}1{\cdot}1{\cdot}\mathring{1}$
v

$$\{\mathring{\varepsilon}, \bar{1}, \mathring{1}, 1{\cdot}\mathring{1}, 1{\cdot}1{\cdot}\bar{1}, 1{\cdot}1{\cdot}\mathring{1}, 1{\cdot}1{\cdot}1{\cdot}\mathring{1}, 1{\cdot}1{\cdot}1{\cdot}1{\cdot}\bar{1}, 1{\cdot}1{\cdot}1{\cdot}1{\cdot}\mathring{1}\}$$

represents faithfully

- parts as connected components with vertex borders (both vertex $\mathring{p}$ and edge $\bar{p}$ positions)
- union, intersection simply as sets
- but behaviour under substitution? (tracing positons)

Representation: geometric and inductive

isomorphic: transfer from one to the other when appropriate:

- for union, intersection: geometric cluster
- for substitution: inductive cluster

Theorem

clusters form distributive lattice

Birkhoff's Fundamental Theorem for Distributive Lattices
On geometric clusters, set theoretic (empty set,union,intersection,all)

Representation: geometric and inductive

isomorphic: transfer from one to the other when appropriate:

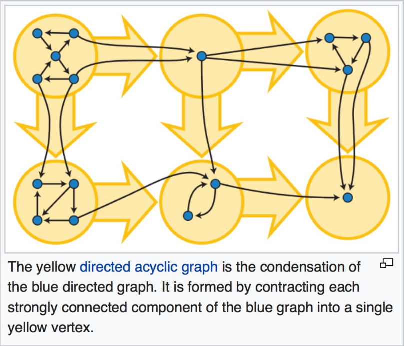
- for union, intersection: geometric cluster
- for substitution: inductive cluster

Theorem

clusters form distributive lattice

Geometric $\rightsquigarrow$ inductive:

collapse connected components to single vertices (let-binding):



The yellow directed acyclic graph is the condensation of the blue directed graph. It is formed by contracting each strongly connected component of the blue graph into a single yellow vertex.

Representation: geometric and inductive

isomorphic: transfer from one to the other when appropriate:

- for union, intersection: geometric cluster
- for substitution: inductive cluster

Theorem

clusters form distributive lattice

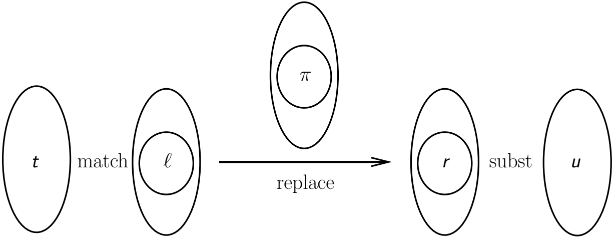
Inductive $\rightsquigarrow$ geometric:
position algebra mapping body to Shift, let-bindings to Tree:

$$\begin{aligned}
\mathcal{Shift}(f)(\vec{P}) &= \begin{pmatrix} \varnothing \\ \varnothing \end{pmatrix} + \begin{pmatrix} \varnothing & \varnothing \\ \varnothing & \{1\} \end{pmatrix} \cdot P_1 + \ldots + \begin{pmatrix} \varnothing & \varnothing \\ \varnothing & \{n\} \end{pmatrix} \cdot P_n; \\
\mathcal{Tree}(f)(\vec{P}) &= \begin{pmatrix} \{\bar{\varepsilon}\} \\ \{\mathring{\varepsilon}\} \end{pmatrix} + \begin{pmatrix} \varnothing & \varnothing \\ \{1\} & \{1\} \end{pmatrix} \cdot P_1 + \ldots + \begin{pmatrix} \varnothing & \varnothing \\ \{n\} & \{n\} \end{pmatrix} \cdot P_n;
\end{aligned}$$

matrix interpretation in Kleene algebra (concatenation, union)

Rewriting as 3-phase process

matching, replacement, substitution



rule $\pi : \ell \to r$

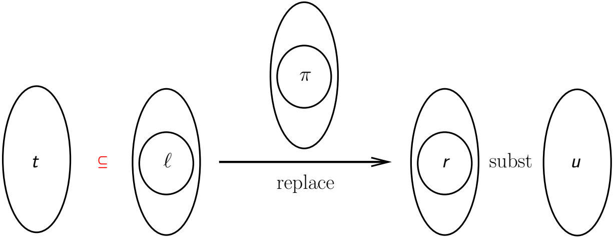
Example (rule $\pi : p(s(v_1)) \to v_1$)

1. matching $t = p(p(s(v)))$ gives `let` $X = p(s(v_1))$ `in` $p(X(v))$
2. replacement step `let` $X = \pi(v_1)$ `in` $p(X(v))$
3. substituting `let` $X = v_1$ `in` $p(X(v))$ yields $p(v) = u$

note: in example π is unary since ℓ is

Rewriting as 3-phase process

matching, replacement, substitution



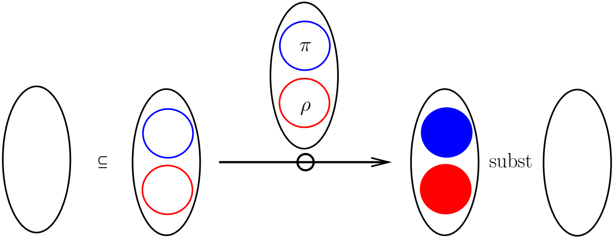
rule $\pi : \ell \to r$

Example (rule $\pi : p(s(v_1)) \to v_1$)

1. matching $t = p(p(s(v)))$ gives `let` $X = p(s(v_1))$ `in` $p(X(v))$
2. replacement step `let` $X = \pi(v_1)$ `in` $p(X(v))$
3. substituting `let` $X = v_1$ `in` $p(X(v))$ yields $p(v) = u$

note: in example π is unary since ℓ is

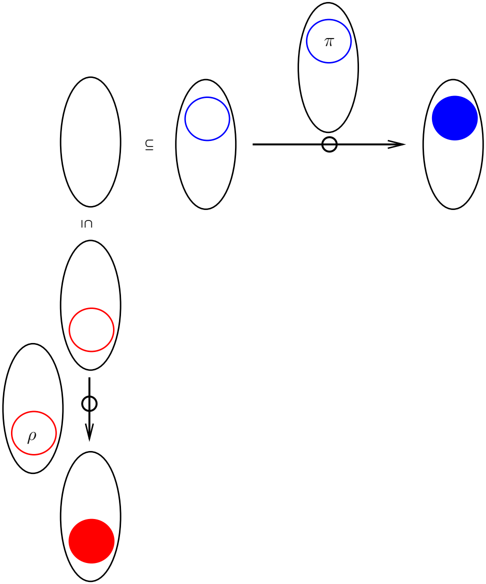
Rewriting on disjoint parts



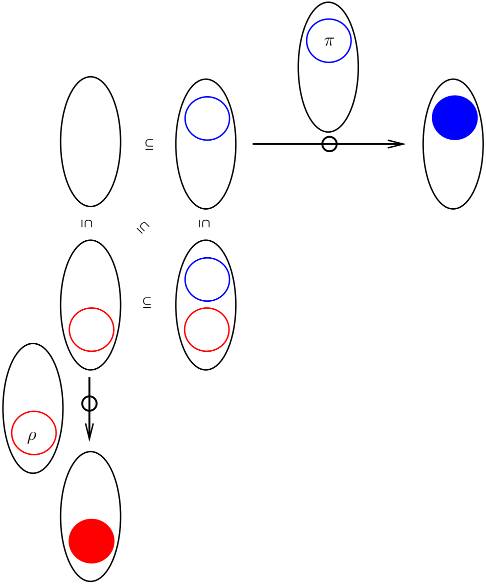
Example (rules $\pi : p(s(v_1)) \to v_1$, $\rho : s(p(v_1)) \to v_1$)

1. matching $t = p(s(p(s(p(v)))))$ gives
   `let` $X, Y = p(s(v_1)), s(p(v_1))$ `in` $X(p(Y(v)))$
2. replacement `let` $X, Y = \pi(v_1), \rho(v_1)$ `in` $X(p(Y(v)))$
3. substituting `let` $X, Y = v_1, v_1$ `in` $X(p(Y(v)))$ gives $p(v) = u$

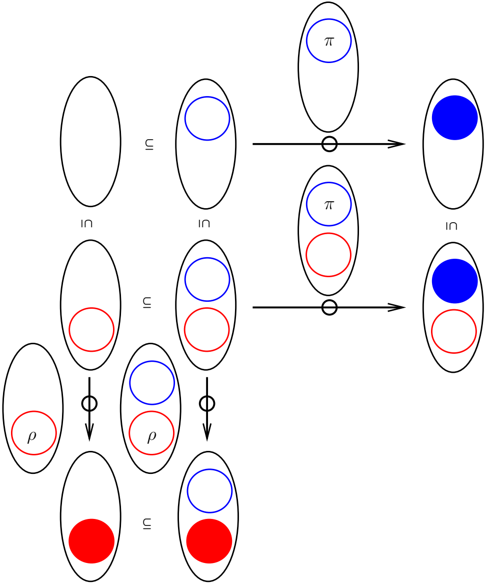
Confluence by Orthogonality (diamond of $\mathrel{-\!\!\circ\!\!\rightarrow}$)



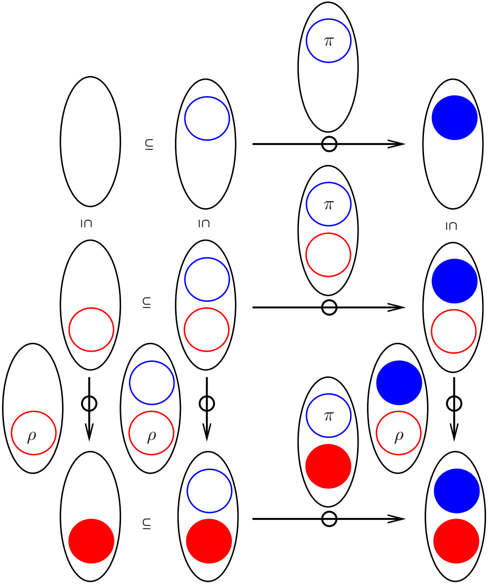
Confluence by Orthogonality (diamond of $\multimap\!\!\to$)



Confluence by Orthogonality (diamond of $\multimap\!\!\rightarrow$)



Confluence by Orthogonality (diamond of $\multimap\!\!\!\rightarrow$)



Critical Peaks

Definition

a peak of steps is critical if its source is the union of their lhss

Example

Peak $\texttt{let}\ X = \pi(v_1)\ \texttt{in}\ X(p(v_1))$ and $\texttt{let}\ Y = \rho(v_1)\ \texttt{in}\ p(Y(v_1))$ is critical

it source $p(s(p(v_1)))$ has positions $\{\mathring{\varepsilon}, \bar{1}, \mathring{1}, 1{\cdot}\bar{1}, 11\mathring{1}\}$

lhs $\texttt{let}\ X = p(s(v_1))\ \texttt{in}\ X(p(v_1))$ has $\{\mathring{\varepsilon}, \bar{1}, \mathring{1}\}$, and

lhs $\texttt{let}\ Y = s(p(v_1))\ \texttt{in}\ p(Y(v_1))$ has $\{\mathring{1}, 1{\cdot}\bar{1}, 11\mathring{1}\}$

must overlap to be critical; here at $\mathring{1}$

Peak $\texttt{let}\ X = \pi(v_1)\ \texttt{in}\ p(X(p(v_1)))$ and $\texttt{let}\ Y = \rho(v_1)\ \texttt{in}\ p(p(Y(v_1)))$ is not critical: p ‘sticks out’

formally: vertex $\mathring{\varepsilon}$ is not in union of lhss

Critical Peaks

Definition

a peak of steps is critical if its source is the union of their lhss

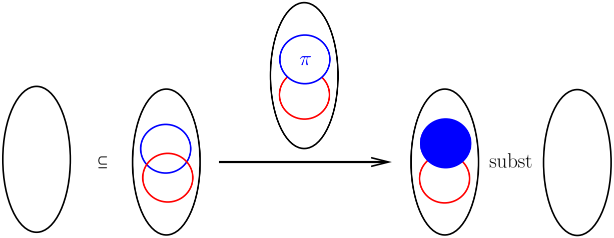
Remark

Without overlap all could be done with so-called multisteps as well

crucial added expressive power lhs:
if $\pi : \ell \to r$, then lhs of $\texttt{let}\ X = \pi\ \texttt{in}\ s$ is $\texttt{let}\ X = \ell\ \texttt{in}\ s$
not a term as for multisteps, but a term with identified pattern ℓ

multisteps/proofterms were designed for representing permutation equivalence, causally independent steps. overlap = dependence. already breaks down for $b \leftarrow a \rightarrow c$.

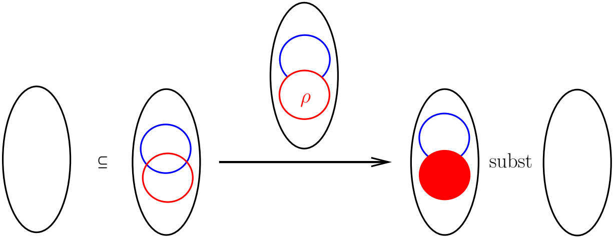
Rewriting on overlapping parts



Example (rules $\pi : p(s(v_1) \rightarrow v_1$, $\rho : s(p(v_1) \rightarrow v_1$)

1. matching on $t = p(s(p(v)))$ gives
   `let` $X = p(s(v_1))$ `in` $X(p(v))$

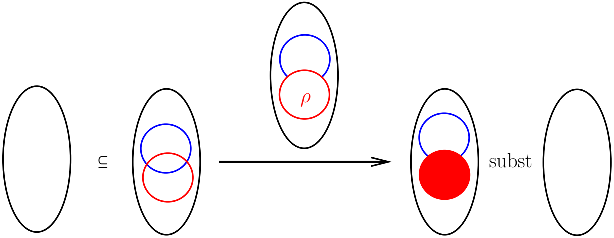
Rewriting on overlapping parts



Example (rules $\pi : p(s(v_1) \to v_1$, $\rho : s(p(v_1) \to v_1)$

1. matching on $t = p(s(p(v)))$ gives
   `let` $X = p(s(v_1))$ `in` $X(p(v))$ or
   `let` $Y = s(p(v_1))$ `in` $p(Y(v))$

Rewriting on overlapping parts



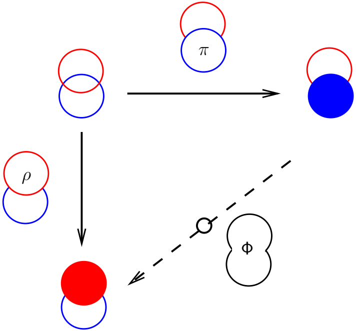
Example (rules $\pi : p(s(v_1) \rightarrow v_1$, $\rho : s(p(v_1) \rightarrow v_1$)

1. matching on $t = p(s(p(v)))$ gives
   `let` $X = p(s(v_1))$ `in` $X(p(v))$ or
   `let` $Y = s(p(v_1))$ `in` $p(Y(v))$
2. both steps can be performed on the union of lhss:
   `let` $Z = p(s(p(v_1)))$ `in` $X(v)$

Confluence by Development Closedness

Theorem

$\multimap\!\!\rightarrow$ has diamond property if all outer–inner critical peaks are development closed



Example

critical peaks between π, ρ are development closed (trivial)

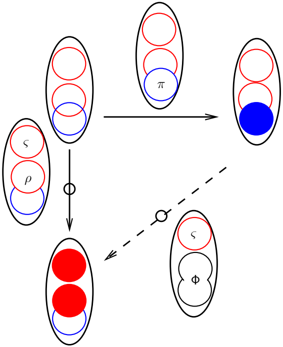
Confluence by Development Closedness

Theorem

$\mathrel{-\!\!\circ\!\!\rightarrow}$ *has diamond property if all outer–inner critical peaks are development closed*

Proof.

induction on amount of overlap with pièce de résistance:



blue step is inner so only overlapped by one red redex □

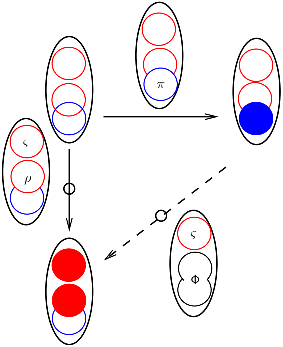
Confluence by Development Closedness

Theorem

$\mathrel{-\!\!\circ\!\!\!\rightarrow}$ has diamond property if all outer–inner critical peaks are development closed

Proof.

induction on amount of overlap with pièce de résistance:

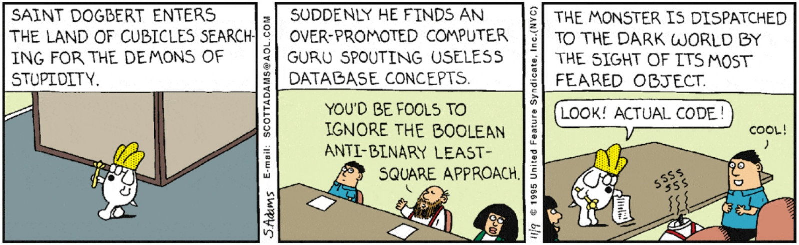


Φ is $\mathrel{-\!\!\circ\!\!\!\rightarrow}$ step on identified part □

Works in theory, not in practice?

Works in theory, not in practice?

Thursday November 09, 1995


SAINT DOGBERT ENTERS THE LAND OF CUBICLES SEARCHING FOR THE DEMONS OF STUPIDITY.
S.Adams E-mail: SCOTTADAMS@AOL.COM
SUDDENLY HE FINDS AN OVER-PROMOTED COMPUTER GURU SPOUTING USELESS DATABASE CONCEPTS.
YOU'D BE FOOLS TO IGNORE THE BOOLEAN ANTI-BINARY LEAST-SQUARE APPROACH.
11/9 © 1995 United Feature Syndicate, Inc.(NYC)
THE MONSTER IS DISPATCHED TO THE DARK WORLD BY THE SIGHT OF ITS MOST FEARED OBJECT.
LOOK! ACTUAL CODE!
COOL!
SSSS SSS

Extension?

What about not left-linear first-order finite term rewriting?

Extension: non-left-linear

- parts $f(x, g(y))$ and $f(z, z)$ in $h(f(g(a), g(a)))$?
- geometric patterns $\{\mathring{1}, 2{\cdot}\bar{1}, 2{\cdot}\mathring{1}\}$ and $\{\mathring{1}\}$ with
- homogeneity relation induced by repeated z: $1{\cdot}\bar{1} \sim 2{\cdot}\bar{1}$
  if related then also corresponding arguments
- union of parts is $\{\mathring{1}, 1{\cdot}\bar{1}, 1{\cdot}\mathring{1}, 2{\cdot}\bar{1}, 2{\cdot}\mathring{1}\}$ with homogeneity
- terms modulo $\sim$ constitute a lattice (Smetsers):
  $t/{\sim_1} \sqcup t/{\sim_2} = t/({\sim_1} \cup {\sim_2})^*$ and $t/{\sim_1} \sqcap t/{\sim_2} = t/({\sim_1} \cap {\sim_2})$
- union is Paterson and Wegman
- not a distributive lattice

Extension: higher-order

- pattern (Miller): restriction on higher-order terms such that unification behaves as in first-order case
- idea: union of terms of patterns but intersection on variables $f(\lambda xy.F(x))$ union $f(\lambda xy.G(y))$ is $f(\lambda xy.H)$, neither x nor y
- to be done ...

Extension: graphs

- pattern: connected component
- idea: as in first-order term case (is embedded)
- warning (already in term graphs): patterns need not be convex
  create redex below by contracting above
- to be done . . .

Extension: infinite terms

- same as for finite terms
- caveat: infinitely many subterms $\rightsquigarrow$ infinite arity
  unify Dershowitz, Klop, . . . (deep) and Rodenburg (wide)?
- to be done . . .

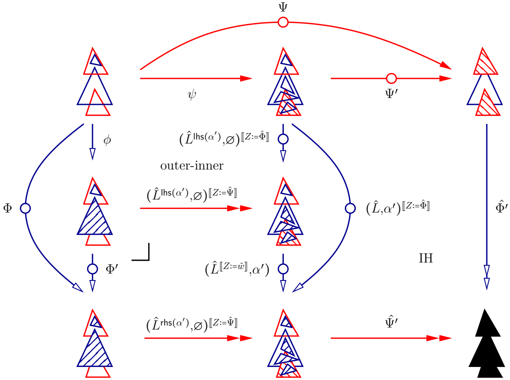
Further Ideas: Modularity

- Confluence Constructor (Peeters)
- clusters allow to represent every layer as a single symbol
  rank $\rightsquigarrow$ height
- factor modularity through non-height increasing TRS
- example: if rules are flat $f(\vec{x}) \to g(\vec{x}')$ and confluent
  then all terms are confluent

Conclusion

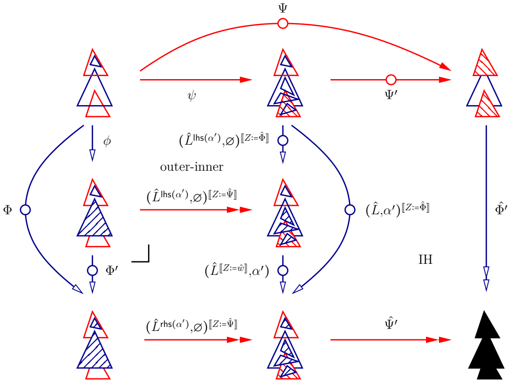
it works

Reviews?


Ψ
ψ
Ψ'
ϕ
$(\hat{L}^{\mathsf{lhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Phi}]\!]}$
outer-inner
$(\hat{L}^{\mathsf{lhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Psi}]\!]}$
Φ
$(\hat{L},\alpha')^{[\![Z:=\hat{\Phi}]\!]}$
$\hat{\Phi}'$
IH
Φ'
$(\hat{L}^{[\![Z:=\hat{w}]\!]},\alpha')$
$(\hat{L}^{\mathsf{rhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Psi}]\!]}$
$\hat{\Psi}'$

The proof of Theorem 44 heavily relies on the figure. Can you provide a proof which can be followed without a figure? (Of course, this is not saying having a figure is bad; it is always nice to have a figure like the presented one to understand the lined proof.)

Reviews?

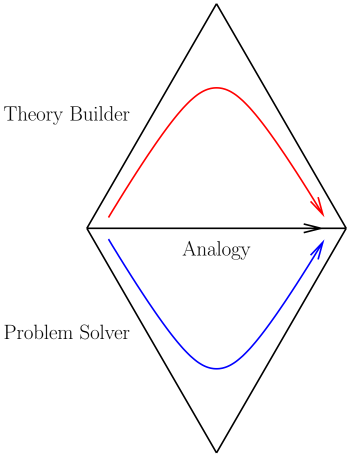

Ψ
ψ
Ψ'
ϕ
$(\hat{L}^{\mathsf{lhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Phi}]\!]}$
outer-inner
Φ
$(\hat{L}^{\mathsf{lhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Psi}]\!]}$
$(\hat{L},\alpha')^{[\![Z:=\hat{\Phi}]\!]}$
$\hat{\Phi}'$
Φ'
$(\hat{L}^{[\![Z:=\hat{w}]\!]},\alpha')$
IH
$(\hat{L}^{\mathsf{rhs}(\alpha')},\varnothing)^{[\![Z:=\hat{\Psi}]\!]}$
$\hat{\Psi}'$

fundamental misunderstanding:

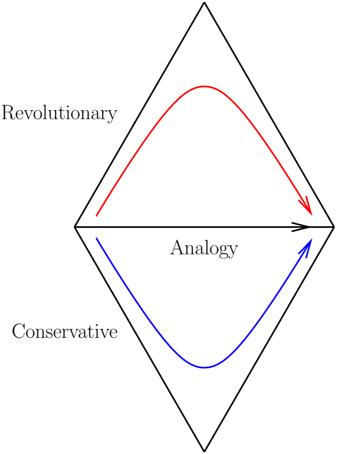
the figure is the proof, formal expressions for steps

sources and targets match up by construction

Epilogue



Epilogue


Revolutionary
Analogy
Conservative