## The process semantics of regular expressions

## Clemens Grabmayer



Computer Science group

Gran Sasso Science Institute
L'Aquila

## Postdoc Seminar

November 19, 2018


## Regular Expressions

## Definition

The set $\operatorname{Reg}(A)$ of regular expressions over alphabet $A$ is defined by the grammar:

$$
e, f::=0|1| a|(e+f)|(e \cdot f) \mid\left(e^{\star}\right) \quad(\text { for } a \in A) .
$$

## Regular Expressions (Copi-Elgot-Wright, 1958; based on Kleene, 1951)

## Definition

The set $\operatorname{Reg}(A)$ of regular expressions over alphabet $A$ is defined by the grammar:

$$
e, f::=0|1| a|(e+f)|(e \cdot f) \mid\left(e^{*}\right) \quad(\text { for } a \in A) .
$$

Note, here:

- symbol 0 instead of $\varnothing$
- symbol 1 used (often dropped, definable as $0^{*}$ )
- no complementation operation $\bar{e}$


## Language semantics $\llbracket \cdot \rrbracket_{L} \quad$ (Copi-Elgot-Wright, 1958)

$0 \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}$ empty language $\varnothing$
$1 \stackrel{\llbracket \rrbracket_{L}}{\longmapsto}\{\epsilon\} \quad(\epsilon$ the empty word)
$a \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}\{a\}$

## Language semantics $\llbracket \cdot \rrbracket_{L} \quad$ (Copi-Elgot-Wright, 1958)

$0 \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}$ empty language $\varnothing$
$1 \stackrel{\llbracket \rrbracket_{L}}{\longmapsto} \quad\{\epsilon\} \quad(\epsilon$ the empty word)
$a \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}\{a\}$
$e+f \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}$ union of $\llbracket e \rrbracket_{L}$ and $\llbracket f \rrbracket_{L}$
$e \cdot f \stackrel{\llbracket \cdot \rrbracket_{L}}{\longmapsto}$ element-wise concatenation of $\llbracket e \rrbracket_{L}$ and $\llbracket f \rrbracket_{L}$
$e^{*} \xrightarrow{\llbracket \cdot \rrbracket_{L}}$ set of words formed by concatenating words in $\llbracket e \rrbracket_{L}$ plus the empty word $\epsilon$

## Process semantics $\llbracket \cdot \rrbracket_{P} \quad$ (Miner, 1984)

$0 \stackrel{\Vdash \Vdash \|_{P}}{\longleftrightarrow}$ deadlock $\delta$, no termination
$1 \stackrel{\Vdash \cdot \|_{P}}{\longleftrightarrow}$ empty process $\epsilon$, then terminate
$a \xrightarrow{\llbracket \|_{P}}$ atomic action $a$, then terminate

## Process semantics $\llbracket \cdot \rrbracket_{P} \quad$ (Miner, 1984)

$0 \stackrel{\llbracket \cdot \mathbb{P}_{P}}{\longleftrightarrow}$ deadlock $\delta$, no termination
$1 \stackrel{\llbracket \eta_{P}}{\longleftrightarrow}$ empty process $\epsilon$, then terminate
$a \xrightarrow{\llbracket \|_{P}}$ atomic action $a$, then terminate
$e+f \xrightarrow{\llbracket!\|_{P}}$ alternative composition of $\llbracket e \rrbracket_{P}$ and $\llbracket f \rrbracket_{P}$
$e \cdot f \xrightarrow{\|\cdot\|_{P}}$ sequential composition of $\llbracket e \rrbracket_{P}$ and $\llbracket f \rrbracket_{P}$
$e^{*} \xrightarrow{\llbracket \|_{P}}$ unbounded iteration of $\llbracket e \rrbracket_{P}$, option to terminate

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$\left(a a(b a)^{*} b\right)^{*} 0$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$\left(a \cdot a \cdot(b \cdot a)^{*} \cdot b\right)^{*} \cdot 0$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$\left(a \cdot a \cdot(b \cdot a)^{*} \cdot b\right)^{*} \cdot 0$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$a \cdot(a \cdot(b+b \cdot a))^{*} \cdot 0$
$\left(a \cdot a \cdot(b \cdot a)^{*} \cdot b\right)^{*} \cdot 0$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$\llbracket a \cdot(a \cdot(b+b \cdot a))^{*} \cdot 0 \rrbracket_{P}$

$$
\llbracket\left(a \cdot a \cdot(b \cdot a)^{*} \cdot b\right)^{*} \cdot 0 \rrbracket_{P}
$$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)


$\llbracket a(a(b+b a))^{*} 0 \rrbracket_{P}$

$\llbracket\left(a a(b a)^{*} b\right)^{*} 0 \rrbracket_{P}$

## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Process semantics $\llbracket \cdot \rrbracket_{P}$ (examples)



## Expressible process graphs (under bisimulation $\leftrightarrows$ )



## Expressible process graphs (under bisimulation $\leftrightarrows$ )



$$
\notin i m\left(\mathbb{[} \cdot \rrbracket_{P}\right)
$$

## Expressible process graphs (under bisimulation $\leftrightarrows$ )


$\epsilon \operatorname{im}\left(\llbracket \rrbracket_{P}\right)$
$\notin i m\left(\mathbb{\llbracket} \rrbracket_{P}\right)$
$\llbracket \cdot \rrbracket_{P}$-expressible

## Expressible process graphs (under bisimulation $\leftrightarrows$ )


$\epsilon \operatorname{im}\left(\llbracket \rrbracket_{P}\right)$

$$
\notin i m\left(\llbracket \rrbracket_{P}\right)
$$

$\llbracket \cdot \rrbracket_{P}$-expressible

## Expressible process graphs (under bisimulation $\leftrightarrows$ )



$$
\begin{gathered}
\epsilon \operatorname{im}\left(\llbracket \cdot \rrbracket_{P}\right) \\
\llbracket \cdot \rrbracket_{P} \text {-expressible }
\end{gathered}
$$

$$
\notin \operatorname{im}\left(\llbracket \rrbracket_{P}\right)
$$


$\llbracket \cdot \|_{p}$-expressible

## Expressible process graphs (under bisimulation $\leftrightarrows$ )



## Expressible process graphs (under bisimulation $\leftrightarrows$ )



## Properties of $P$

- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible.

not $\llbracket \cdot \rrbracket_{P}$-expressible $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$


## Properties of $P$

- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible.
- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$.

not $\llbracket \cdot \rrbracket_{P}$-expressible
$\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$

not $\llbracket \cdot \rrbracket_{P}$-expressible not $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$


## Properties of $P$

- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible.
- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$.
- Fewer identities hold for $\leftrightarrows_{P}$ than for $=_{L}: \quad \leftrightarrows_{P} \varsubsetneqq={ }_{L}$.


## Properties of $P$

- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible.
- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$.
- Fewer identities hold for $\leftrightarrows_{P}$ than for $=_{L}: \quad \leftrightarrows_{P} \varsubsetneqq=_{L}$.



## Properties of $P$

- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible.
- Not every finite-state process is $\llbracket \cdot \rrbracket_{P}$-expressible modulo $\leftrightarrows$.
- Fewer identities hold for $\leftrightarrows_{P}$ than for $=_{L}: \quad \leftrightarrows_{P} \varsubsetneqq=_{L}$.


$$
a \cdot(b+c)
$$

$4_{P}$
$a \cdot b+a \cdot c$

## Complete axiomatization of $=_{L} \quad$ (Aanderaa/Salomaa, 1965/66)

Axioms:
(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B8) $e \cdot 0=0$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $e^{*}=1+e \cdot e^{*}$
(B5) $e \cdot(f+g)=e \cdot f+e \cdot g$
(B11) $e^{*}=(1+e)^{*}$
(B6) $\quad e+e=e$
Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \text { FIX } \quad \text { if } \underbrace{\{\epsilon\} \notin \llbracket f \rrbracket_{L}}_{\begin{array}{c}
\text { non-empty-word } \\
\text { property }
\end{array}})
$$

## Sound and unsound axioms with respect to $\leftrightarrows_{P}$

## Axioms:

(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B8) $e \cdot 0=0$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $\quad(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $\quad e^{*}=1+e \cdot e^{*}$
(B5) $e \cdot(f+g)=e \cdot f+e \cdot g$
(B11) $e^{*}=(1+e)^{*}$
(B6) $e+e=e$
Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \text { FIX } \quad \text { if } \underbrace{\{\epsilon\} \notin \llbracket f \rrbracket_{L}}_{\begin{array}{c}
\text { non-empty-word } \\
\text { property }
\end{array}})
$$

## Sound and unsound axioms with respect to $\leftrightarrows_{P}$

## Axioms:

(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B8) $e \cdot 0=0$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $\quad(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $\quad e^{*}=1+e \cdot e^{*}$
(B5) $e \cdot(f+g)=e \cdot f+e \cdot g$
(B11) $\quad e^{*}=(1+e)^{*}$
(B6) $e+e=e$
$(\mathrm{B} 8)^{\prime} \quad 0 \cdot e=0$
Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \text { FIX } \quad \text { if } \underbrace{\{\epsilon\} \notin \llbracket f \rrbracket_{L}}_{\begin{array}{c}
\text { non-empty-word } \\
\text { property }
\end{array}})
$$

## Adaptation for $\overleftrightarrow{S}_{P} \quad$ (Milner, 1984) (Mil $=$ Mil $^{-}+$RSP $\left.^{*}\right)$

## Axioms:

(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $\quad(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $\quad e^{*}=1+e \cdot e^{*}$
(B11) $\quad e^{*}=(1+e)^{*}$
(B6) $e+e=e$
$(\mathrm{B} 8)^{\prime} \quad 0 \cdot e=0$

Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \mathrm{RSP}^{*}(\text { if } \underbrace{\{\epsilon\} \notin \llbracket f \rrbracket_{L}}_{\begin{array}{c}
\text { non-empty-word } \\
\text { property }
\end{array}})
$$

## Adaptation for $\overleftrightarrow{S}_{P} \quad$ (Milner, 1984) (Mil $=$ Mil $^{-}+$RSP $\left.^{*}\right)$

Axioms:
(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B8) $\quad 0 \cdot e=0$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $e^{*}=1+e \cdot e^{*}$
(B11) $e^{*}=(1+e)^{*}$
(B6) $e+e=e$
Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \mathrm{RSP}^{*}(\text { if } \underbrace{\text { property }}_{\text {non-empty-word }} \begin{array}{|c|}
\left\{\epsilon \notin \llbracket f \rrbracket_{L}\right.
\end{array})
$$

## Adaptation for $\overleftrightarrow{S}_{P} \quad$ (Milner, 1984) (Mil $=$ Mil $^{-}+$RSP $\left.^{*}\right)$

Axioms:
(B1) $e+(f+g)=(e+f)+g$
(B7) $e \cdot 1=e$
(B2) $\quad(e \cdot f) \cdot g=e \cdot(f \cdot g)$
(B8) $\quad 0 \cdot e=0$
(B3) $\quad e+f=f+e$
(B9) $e+0=e$
(B4) $(e+f) \cdot g=e \cdot g+f \cdot g$
(B10) $e^{*}=1+e \cdot e^{*}$
(B11) $\quad e^{*}=(1+e)^{*}$
(B6) $e+e=e$
Inference rules: equational logic plus

$$
\frac{e=f \cdot e+g}{e=f^{*} \cdot g} \mathrm{RSP}^{*}(\text { if } \underbrace{\{\epsilon\} \notin \llbracket f \rrbracket_{L}}_{\begin{array}{c}
\text { non-empty-word } \\
\text { property }
\end{array}})
$$

## Milner's questions

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

## Milner's questions

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

## Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

## Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

- definability by well-behaved specifications (Baeten/Corradini, 2005)

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

## Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

- definability by well-behaved specifications (Baeten/Corradini, 2005)
- that is decidable (super-exponentially) (Baeten/Corradini/G, 2007)

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

- definability by well-behaved specifications (Baeten/Corradini, 2005)
- that is decidable (super-exponentially) (Baeten/Corradini/G, 2007)

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

- Mil is complete for perpetual-loop expressions (Fokkink, 1996)
- every iteration $e^{*}$ occurs as part of a 'no-exit' subexpression $e^{*} \cdot 0$


## Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

- definability by well-behaved specifications (Baeten/Corradini, 2005)
- that is decidable (super-exponentially) (Baeten/Corradini/G, 2007)

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

- Mil is complete for perpetual-loop expressions (Fokkink, 1996)
- every iteration $e^{*}$ occurs as part of a 'no-exit' subexpression $e^{*} \cdot 0$
- Mil is complete when restricted to 1 -return-less expressions
(Corradini, De Nicola, Labella, 2002)


## Milner's questions, and partial results

Q1. Recognition: Which structural property of finite process graphs characterizes $\llbracket \cdot \rrbracket_{P}$-expressibility modulo $\leftrightarrows$ ?

- definability by well-behaved specifications (Baeten/Corradini, 2005)
- that is decidable (super-exponentially) (Baeten/Corradini/G, 2007)

Q2. Complete axiomatization: Is Mil complete for $\leftrightarrows_{P}$ ?

- Mil is complete for perpetual-loop expressions (Fokkink, 1996)
- every iteration $e^{*}$ occurs as part of a 'no-exit' subexpression $e^{*} \cdot 0$
- Mil is complete when restricted to 1 -return-less expressions (Corradini, De Nicola, Labella, 2002)
- $\mathrm{Mil}^{-}$+ one of two stronger rules (than RSP*) is complete ( $G$, 2006)

New approach: Loop Existence and Elimination (LEE)


New approach: Loop Existence and Elimination (LEE)


New approach: Loop Existence and Elimination (LEE)


New approach: Loop Existence and Elimination (LEE)


New approach: Loop Existence and Elimination (LEE)


New approach: Loop Existence and Elimination (LEE)


## New approach: Loop Existence and Elimination (LEE)



## New approach: Loop Existence and Elimination (LEE)


eliminate loop


## New approach: Loop Existence and Elimination (LEE)


eliminate loop


## New approach: Loop Existence and Elimination (LEE)


eliminate loop


## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop entry


## New approach: Loop Existence and Elimination (LEE)


eliminate loop

garbage collection eliminate loop entry


## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop


## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop

repeated elimination of loops leads to a process graph without infinite trace

## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop

repeated elimination of loops leads to a process graph without infinite trace

## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop

repeated elimination of loops leads to a process graph without infinite trace


## New approach: Loop Existence and Elimination (LEE)


eliminate loop

eliminate loop

repeated elimination of loops leads to a process graph without infinite trace


## LEE-witness


structured
LEE-witness

## LEE-witness: structure constrained process graph



LEE-witness
structured
LEE-witness

## LEE-witness yields star expression



LEE-witness
structured
LEE-witness

$\llbracket\left(a a(b a)^{*} b\right)^{*} 0 \rrbracket_{P}$

## structure constraints (L'Aquila)



## structure constraints (L'Aquila)



## Loop elimination



## Loop elimination



## Loop elimination



## Loop elimination




## Loop elimination




## Loop elimination




## Loop elimination




## Loop elimination





## Loop elimination






## Loop elimination






## Loop elimination






## Loop elimination






## Loop elimination






## Loop elimination






## Loop elimination





$\xrightarrow{\longrightarrow}$ elim

## Loop elimination





$\xrightarrow{\longrightarrow}$ elim

## Loop elimination




$\xrightarrow{\longrightarrow}$ elim


## Loop elimination




$\xrightarrow{ } \mathrm{elim}$


## Loop elimination



$\xrightarrow{ } \mathrm{elim}$


## Loop elimination


$\longrightarrow \mathrm{elim}$
$v_{1}$

$\xrightarrow{l}$ elim

$\longrightarrow{ }^{*}$ elim


## Loop elimination


$\longrightarrow$ elim
$v_{1}$


$$
\longrightarrow \mathrm{elim}
$$



## Loop elimination


$\longrightarrow$ elim
$v_{1}$

$\xrightarrow{ } \mathrm{elim}$


## Loop elimination


$\longrightarrow$ elim
© $v_{1}$


$\longrightarrow \mathrm{elim}$

$\xrightarrow{ } \mathrm{elim}$
$\longrightarrow{ }^{*}{ }^{*}$


## Loop elimination


$\longrightarrow$ elim
$v_{1}$


$\xrightarrow{\nrightarrow}$ elim


## LEE fails



## LEE fails



## LEE fails



## LEE fails




## LEE under bisimulation

## Observation

- LEE is not invariant under bisimulation.
- LEE is not preserved by converse functional bisimulation.


LEE
$\neg L E E$


LEE
ᄀLEE

## LEE is preserved under bisimulation collapse


$\llbracket a(a(b+b a))^{*} 0 \rrbracket_{P}$
$\llbracket\left(a a(b a)^{*} b\right)^{*} 0 \rrbracket_{P}$

## Goals

- Completeness proof
- is a large project: report (now $\sim 250$ pages, $\sim 2$ years)
- writing up a crucial step:
- pseudo-collapse LEE-witnesses with 1-transitions
- Interpretations of equational theories
- Characterize interpretations with respect to what kinds of properties they permit to transfer between equational theories


## Goals

- Completeness proof
- is a large project: report (now $\sim 250$ pages, $\sim 2$ years)
- writing up a crucial step:
- pseudo-collapse LEE-witnesses with 1-transitions
- Interpretations of equational theories
- Characterize interpretations with respect to what kinds of properties they permit to transfer between equational theories
- Recognition problem
- polynomial if no 1-transitions (testing for property LEE)
- Cost models for $\lambda$-calculus

