Regularity and context-freeness over word rewriting systems

Didier Caucal and Trong Hieu Dinh

LIGM / CNRS

University Paris - Est

France

Rewriting system R over an alphabet A

R | **u** $\longrightarrow$ **v** **with** **u , v** **in** **A***

Rewriting $xuy \underset{R}{\longrightarrow} xvy$ with $u\ R\ v$

Derivation $\overset{*}{\underset{R}{\longrightarrow}}$ preserves regularity (cf.) if

$$\overset{*}{\underset{R}{\longrightarrow}}(L) = \{\, v \mid \exists\, u \in L \quad u \overset{*}{\underset{R}{\longrightarrow}} v \,\}$$

is regular (cf.) for **L** regular (cf.)

Example

$$R \;\Big|\; ba \longrightarrow ab$$

$\xrightarrow[R]{*}$ does not preserve regularity and cf.

$$\xrightarrow[R]{*} (ab)^* \cap a^* b^* = \{ a^n b^n \mid n \geq 0 \}$$

$$\xrightarrow[R]{*} \{ (ab)^n c^n \mid n \geq 0 \} \cap a^* b^* c^*$$

$$= \{ a^n b^n c^n \mid n \geq 0 \}$$

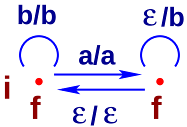
$R \mid a \longrightarrow a\,b$

$\xrightarrow[R]{*}$ preserves regularity and cf.

$\xrightarrow[R]{*}$ is a regular substitution

$\xrightarrow[R]{*}(a) = a\,b^*$ and $\xrightarrow[R]{*}(b) = b$

recognized by the transducer



Fact

Any rational relation preserves reg. and cf.

Recognizable relation R

$$R \;\Big|\; U_i \longrightarrow V_i \quad \text{with} \quad U_i, V_i \ \text{regular}$$

$$R = \bigcup_i U_i \times V_i$$

Fact

Any recognizable relation R is rational, hence

the image $R(L)$ is regular (cf.) if L is regular (cf.)

Monadic system

$$R \left| \begin{array}{lll} u \longrightarrow a & \text{with } a \text{ letter} \\ u \longrightarrow \varepsilon & u \text{ context-free} \end{array} \right.$$

Proposition

Any monadic system R is reg / cf - preserving :

$\xrightarrow[R]{*}$ preserves regularity

$\xrightarrow[R^{-1}]{*}$ preserves context-freeness

Dyck system

$$D \left| \begin{array}{lcl} \overrightarrow{a}\, a & \longrightarrow & \varepsilon \\ a\, \overleftarrow{a} & \longrightarrow & \varepsilon \end{array} \right. \qquad \text{with } a \text{ letter}$$

The Dyck system **D** is reg / cf - preserving :

$\xrightarrow[D]{*}$ preserves regularity (but not cf.)

$\xrightarrow[D^{-1}]{*}$ preserves context-freeness (but not reg.)

$$\overleftarrow{a_1 \ldots a_n} = \overleftarrow{a_n} \ldots \overleftarrow{a_1} \quad \text{and} \quad \overrightarrow{a_1 \ldots a_n} = \overrightarrow{a_n} \ldots \overrightarrow{a_1}$$

Decomposition of $\xrightarrow[R]{*}$ for any R

$$\overleftrightarrow{R} \;\Big|\; \varepsilon \longrightarrow \overleftarrow{u}\, w\, \overrightarrow{v} \quad \text{for} \quad u\,v\ R\ w$$

$$x\,u\,v\,y \xrightarrow{R} x\,w\,y$$

$$x\,u\,v\,y \xrightarrow{\overleftrightarrow{R}} x\,u\,\overleftarrow{u}\,w\,\overrightarrow{v}\,v\,y \xrightarrow[D]{*} x\,w\,y$$

Endrullis, Haufbauer, Waldmann 06

$$\xrightarrow[R]{*} \;=\; \Big(\xrightarrow[\overleftrightarrow{R}]{*} \circ \xrightarrow[D]{*}\Big) \cap A^* \times A^*$$

D , $\overleftrightarrow{R}^{-1}$ are monadic

Restriction of $\xrightarrow[\overleftrightarrow{R}]{*}$

Insertion of $L \subseteq A^*$ in a word $a_1 \ldots a_n$

$$(a_1 \ldots a_n)\,[L] \;=\; L^*\, a_1 \ldots L^*\, a_n\, L^*$$

$[L]$ is a substitution : $a[L] \;=\; L^*\, a\, L^*$

We say that R is decomposable into $S \subseteq R$ if

$$\xrightarrow[R]{*} \;=\; (\,[L] \circ \xrightarrow[S \cup D]{*}\,) \cap A^* \times A^*$$

with $L \;=\; \{\ \overleftarrow{u}\ w\ \overrightarrow{v} \mid u\, v\ R - S\ w\ \}$

R is iteratively decomposable into $S \subseteq R$ if

there exists $n > 0$ and $R \supseteq S_1 \supseteq \ldots \supseteq S_n = S$

such that R decomposable into S_1

$S_i \cup D$ decomposable into $S_{i+1} \cup D$

Theorem

If R is iteratively decomposable into S

with R-S recognizable and S monadic

then R is reg / cf - preserving.

$$\xrightarrow[R]{*} \;=\; (h \circ \xrightarrow[S \cup D]{*}) \cap A^* \times A^*$$

for some regular substitution h

Deleting system

$$\mathbf{R} \;\Big|\; \mathbf{u} \longrightarrow \mathbf{v}$$

$\forall$ **b** in **Alph(v)** $\exists$ **a** in **Alph(u)** **a** $>$ **b**

Any deleting system is iteratively decomp. into $\emptyset$

Proposition Haufbauer, Waldmann 03

Any recognizable deleting system is

reg / cf - preserving

$\mathbf{B} \subseteq \mathbf{A}$ is a prefix sub-alphabet of $\mathbf{R}$ if

$$\mathbf{R} \subseteq \mathbf{A^* \times (A - B)^*} \ \mathbf{U} \ \mathbf{B\,A^* \times B(A - B)^*}$$

i.e. for any rule $\mathbf{u \; R \; a\,v}$ with $\mathbf{a}$ in $\mathbf{A}$:

$$\mathbf{v} \in \mathbf{(A - B)^*} \quad \text{and} \quad \mathbf{a} \in \mathbf{B} \implies \mathbf{u(1)} \in \mathbf{B}$$

Proposition

For any prefix sub-alphabet $\mathbf{B}$ of $\mathbf{R}$

$\mathbf{R}$ is decomposable into

$$\{ \ \mathbf{(u,v)} \in \mathbf{R} \mid \mathbf{Alph(uv)} \cap \ \mathbf{B} = \emptyset \ \}$$

Same result for a suffix sub-alphabet $\mathbf{B}$ of $\mathbf{R}$:

$$\mathbf{R} \subseteq \mathbf{A^* \times (A - B)^*} \ \mathbf{U} \ \mathbf{A^*\,B \times (A - B)^*\,B}$$

Application 1 : prefix derivation

$$\mathbf{u\,y} \underset{\mathbf{R}}{\longmapsto} \mathbf{v\,y} \quad \text{for} \quad \mathbf{u\,R\,v} \quad \text{and} \quad \mathbf{y} \in \mathbf{A^*}$$

Proposition

The prefix derivation $\overset{*}{\underset{\mathbf{R}}{\longmapsto}}$ of any recognizable $\mathbf{R}$

preserves regularity and context-freeness

$\mathbf{\#R} = \{\, (\mathbf{\#u}, \mathbf{\#v}) \mid \mathbf{u\,R\,v} \,\}$ for $\mathbf{\#}$ not in $\mathbf{A}$

$\overset{*}{\underset{\mathbf{R}}{\longmapsto}}(\mathbf{L}) = \mathbf{\#}^{-1}(\overset{*}{\underset{\mathbf{\#R}}{\longrightarrow}}(\mathbf{\#L}))$ for any $\mathbf{L} \subseteq \mathbf{A^*}$

$\{\#\}$ is a prefix sub-alphabet of $\mathbf{\#R}$ and $(\mathbf{\#R})^{-1}$

$\mathbf{\#R}$ and $(\mathbf{\#R})^{-1}$ are decomposable into $\emptyset$

Application 2 : left-to-right derivation

$$x\,u\,y \;\underset{R}{\longrightarrow}_{|x|}\; x\,v\,y \quad \text{for} \quad u\,R\,v \quad \text{and} \quad x\,,y \in A^*$$

$$u \;\overset{*}{\underset{R}{\hookrightarrow}}\; v \;\;\text{if}\;\; u = u_0 \underset{R}{\longrightarrow}_{p_1} \ldots \underset{R}{\longrightarrow}_{p_n} u_n = v$$

$$\text{with} \;\; p_1 \leq \;\ldots\; \leq p_n$$

Proposition

The left-to-right derivation $\overset{*}{\underset{R}{\hookrightarrow}}$ for R recognizable

preserves context-freeness, and its inverse regularity

$$\mathbf{R}_{\#} = \{\, (\mathbf{u}, \#\mathbf{v}) \mid \mathbf{u}\ \mathbf{R}\ \mathbf{v} \,\}$$ for $\#$ not in $\mathbf{A}$

For any $\mathbf{L} \subseteq \mathbf{A}^*$

$$\underset{\mathbf{R}}{\overset{*}{\hookrightarrow}}(\mathbf{L}) = \pi(\underset{\mathbf{R}_{\#}}{\overset{*}{\hookrightarrow}}(\mathbf{L})) = \pi(\underset{\mathbf{R}_{\#}}{\overset{*}{\longrightarrow}}(\mathbf{L}))$$

$\{\#\}$ is a prefix sub-alphabet of $(\mathbf{R}_{\#})^{-1}$

$(\mathbf{R}_{\#})^{-1}$ is decomposable into $\emptyset$

Application 3 : tagged system

$$R \subseteq (A \cup B)^* \times (A \cup B)^*$$

with **B** a new alphabet of tags

Tagged bifix system $R_p \cup R_s \cup R_b$

R_p	$\#u$	$\longrightarrow$	$\#v$	prefix rules
R_s	$u\&$	$\longrightarrow$	$v\&$	suffix rules
R_b	$\#u\&$	$\longrightarrow$	$\#v\&$	bifix rules

with $\#,\& \in B$ and $u,v \in A^*$

Proposition Altenbernd 08

$\xrightarrow[R]{*}$ rational for **R** recognizable tagged bifix system

Tag-adding bifix system

$$R = R_p \cup R_s \cup R_b \cup R_d \cup R_i$$

R_p	|	$\#u$	$\longrightarrow$	$\#v$	tag-adding prefix rules
R_s	|	$u\&$	$\longrightarrow$	$v\&$	tag-adding suffix rules
R_b	|	$\#u\&$	$\longrightarrow$	$\#v\&$	tag-adding bifix rules
R_d	|	$\#u\#$	$\longrightarrow$	$\#$	tag-deleting rules
R_i	|	u	$\longrightarrow$	v	tag-adding infix rules

with $\#,\& \in B$, $u \in A^*$, $v \in (A \cup B)^*$

A tagged system $\mathbf{R}$ is context-free if

$$\mathbf{R}^{-1} \cap \mathbf{A}^* \times \mathbf{A}^* \quad \text{is monadic}$$

$$\mathbf{R} - \mathbf{A}^* \times \mathbf{A}^* \quad \text{is recognizable}$$

Proposition

Any context-free tag-adding bifix system is

cf / reg - preserving

True for $(\mathbf{R} \cap \mathbf{A}^* \times \mathbf{A}^*) \cup \mathbf{D}^{-1}$ cf / reg - preserving

Tag-adding prefix/suffix system

$$R = R_p \cup R_s \cup R_b \cup R_i$$

R_p	|	#u	$\longrightarrow$	#'v	tag-adding prefix rules
R_s	|	u&	$\longrightarrow$	v&'	tag-adding suffix rules
R_b	|	#u&	$\longrightarrow$	#'v&'	tag-adding bifix rules
R_i	|	u	$\longrightarrow$	v	tag-adding infix rules

with $\#,\#' \in B$ and $\&,\&' \in C$

$u \in A^*$ and $v \in (A \cup B \cup C)^*$

For any tag-adding prefix/suffix system **R**

B is a prefix sub-alphabet of $\mathbf{R}^{-1}$

C is a suffix sub-alphabet of $\mathbf{R}^{-1}$

$\mathbf{R}^{-1}$ is 2-decomposable into $\mathbf{R}^{-1} \cap \mathbf{A}^* \times \mathbf{A}^*$

For **R** context-free, **R** is cf / reg - preserving

True for $(\mathbf{R} \cap \mathbf{A}^* \times \mathbf{A}^*) \cup \mathbf{D}^{-1}$ cf / reg - preserving

Conclusion

Decomposition by removing letters :

deleting systems

prefix and suffix sub-alphabets

Another decompositions ?