Languages that require full scanning of words to determine membership

Peter M. Higgins & Suhear Alwan

Department of Mathematical Sciences, University of Essex

NBSAN Norwich April 15th 2014
Compare the five languages:

\[ L_1 = aA^* \cup A^* a, \quad L_2 = A^* a, \quad L_3 = aA^* \]

\[ L_4 = aA^* b \cup bA^* a, \quad L_5 = (A^2)^* \]

\( L_1 \) has the **Factor Property (FP):**

\[ \forall u \in A^* \exists u_1, u_2, u'_1, u'_2 \in A^* : u_1 uu_2 \in L \& u'_1 uu'_2 \in L' \]

\( L_2 \) (resp. \( L_3 \)) has the **Prefix Property (PP)\)** (resp. **Suffix Property (SP)):**

\[ \forall x \in A^* \exists v, v' \in A^* : xv \in L, xv' \in L' \] (resp. \( \exists u, u' \in A^* : ux \in L, u'x \in L' \)).

\( L_4 \) has both the **Prefix and Suffix properties (Weak Scan property (WS)).**
Compare the five languages:

\[ L_1 = aA^* \cup A^*a, \quad L_2 = A^*a, \quad L_3 = aA^* \]

\[ L_4 = aA^*b \cup bA^*a, \quad L_5 = (A^2)^* . \]

\( L_1 \) has the **Factor Property (FP)**:

\[ \forall u \in A^* \exists u_1, u_2, u_1', u_2' \in A^* : uu_1u_2 \in L \& u_1'u_2' \in L' \]

\( L_2 \) (resp. \( L_3 \)) has the **Prefix Property (PP)** (resp. **Suffix Property (SP)**):

\[ \forall x \in A^* \exists v, v' \in A^* : xv \in L \& xv' \in L' \quad (\text{resp. } \exists u, u' \in A^* : ux \in L \& u'x \in L') . \]

\( L_4 \) has both the Prefix and Suffix properties (**Weak Scan property (WS)**).
Compare the five languages:

\[ L_1 = aA^* \cup A^* a, \quad L_2 = A^* a, \quad L_3 = aA^* \]
\[ L_4 = aA^* b \cup bA^* a, \quad L_5 = (A^2)^*. \]

\( L_1 \) has the **Factor Property (FP):**

\[ \forall u \in A^* \exists u_1, u_2, u_1', u_2' \in A^* : uu_1u_2 \in L \& uu_1'u_2' \in L' \]

\( L_2 \) (resp. \( L_3 \)) has the **Prefix Property (PP) (resp. Suffix Property (SP)):**

\[ \forall x \in A^* \exists v, v' \in A^* : xv \in L, xv' \in L' \](resp. \( \exists u, u' \in A^* : ux \in L, u'x \in L' \)).

\( L_4 \) has both the Prefix and Suffix properties (**Weak Scan property (WS)).**
Compare the five languages:

\[ L_1 = aA^* \cup A^* a, \quad L_2 = A^* a, \quad L_3 = aA^* \]

\[ L_4 = aA^* b \cup bA^* a, \quad L_5 = (A^2)^* . \]

\( L_1 \) has the \textbf{Factor Property (FP)}:

\[ \forall u \in A^* \exists u_1, u_2, u'_1, u'_2 \in A^* : u_1 uu_2 \in L \& u'_1 uu'_2 \in L' \]

\( L_2 \) (resp. \( L_3 \)) has the \textbf{Prefix Property (PP)} (resp. \textbf{Suffix Property (SP)}):

\[ \forall x \in A^* \exists v, v' \in A^* : xv \in L, xv' \in L' \ (\text{resp.} \ \exists u, u' \in A^* : ux \in L, u'x \in L') . \]

\( L_4 \) has both the Prefix and Suffix properties (\textbf{Weak Scan property (WS)}).

Peter M. Higgins & Suheear Alwan

Languages that require full scanning of words to determine
A stronger condition still is the *Full Scan Property (FS)*:

\[ \forall u, v \in A^* \exists x, x' \in A^* : uxv \in L' \land ux'v \in L'. \]

$L_5$ has the full scan property.

For regular languages, in terms of the minimal automaton $\mathcal{A}(Q, i, T) = \mathcal{A}(L)$ we have:

- $L \in FP \iff \exists \text{ a sink state } q \in Q : Q \cdot z = q \forall z \in A^*$
- $L \in PP \iff |q \cdot A^*| > 1 \forall q \in Q$; $L \in SP \iff L^R \in PP$
- $L \in WS \iff L \in PP \& L \in SP$
- $L \in FS \iff Lv^{-1} \in PP \forall v \in A^*$.
A stronger condition still is the **Full Scan Property (FS)**:

\[ \forall u, v \in A^* \exists x, x' \in A^* : uxv \in L' \land ux'v \in L'. \]

\(L_5\) has the full scan property.

For regular languages, in terms of the minimal automaton \(A(Q, i, T) = A(L)\) we have:

\[ L \in FP \iff \exists a \text{ sink state } q \in Q : Q \cdot z = q \forall z \in A^* \]

\[ L \in PP \iff |q \cdot A^*| > 1 \forall q \in Q; \quad L \in SP \iff L^R \in PP \]

\[ L \in WS \iff L \in PP \land L \in SP \]

\[ L \in FS \iff L v^{-1} \in PP \forall v \in A^*. \]
A stronger condition still is the *Full Scan Property* (FS):

\[ \forall u, v \in A^* \exists x, x' \in A^* : uxv \in L' \land ux'v \in L'. \]

$L_5$ has the full scan property.

For regular languages, in terms of the minimal automaton $A(Q, i, T) = A(L)$ we have:

\[ L \in FP \iff \exists \text{ a sink state } q \in Q : Q \cdot z = q \forall z \in A^* \]

\[ L \in PP \iff |q \cdot A^*| > 1 \forall q \in Q; \ L \in SP \iff L^R \in PP \]

\[ L \in WS \iff L \in PP \land L \in SP \]

\[ L \in FS \iff L v^{-1} \in PP \forall v \in A^*. \]
A stronger condition still is the *Full Scan Property (FS)*:

\[ \forall u, v \in A^* \exists x, x' \in A^* : uxv \in & ux'v \in L'. \]

$L_5$ has the full scan property.

For regular languages, in terms of the minimal automaton $A(Q, i, T) = A(L)$ we have:

$L \in FP \iff \not\exists$ a sink state $q \in Q : Q \cdot z = q \forall z \in A^*$

$L \in PP \iff |q \cdot A^*| > 1 \forall q \in Q; L \in SP \iff L^R \in PP$

$L \in WS \iff L \in PP & L \in SP$

$L \in FS \iff Lv^{-1} \in PP \forall v \in A^*$. 

Languages that require full scanning of words to determine
A stronger condition still is the Full Scan Property (FS):

\[ \forall u, v \in A^* \exists x, x' \in A^*: uxv \in L \land ux'v \in L'. \]

$L_5$ has the full scan property.

For regular languages, in terms of the minimal automaton $A(Q, i, T) = A(L)$ we have:

\[ L \in FP \iff \exists a \text{ sink state } q \in Q : Q \cdot z = q \forall z \in A^* \]

\[ L \in PP \iff |q \cdot A^*| > 1 \forall q \in Q; L \in SP \iff L^R \in PP \]

\[ L \in WS \iff L \in PP \& L \in SP \]

\[ L \in FS \iff Lv^{-1} \in PP \forall v \in A^*. \]
Let \( \overline{\cdot} : A^* \to A^*/\eta = M(L) \) denote the natural morphism of \( A^* \) onto the syntactic monoid \( M(L) \) of language \( L \subseteq A^* \), so that
\[
\overline{u} = \overline{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta \text{ saturates } L.
\]
We say a set \( X \subseteq M(L) \) is a bridge if
\[
X \eta^{-1} \cap L \neq \emptyset \& X \eta^{-1} \cap L' \neq \emptyset.
\]

**Theorem**

Let \( L \) be regular and let \( I \) be the minimum ideal of \( M = M(L) \) Then

(i) \( L \in FP \) if and only if the \( D \)-class \( I \) of \( M \) is a bridge;
(ii) \( L \in PP \) (resp. \( SP \)) if and only if each \( R \)-class (resp. \( L \)-class) of \( I \) is bridge;
(iii) \( L \in WS \) if and only if each \( R \)-class and each \( L \)-class of \( I \) is bridge;
(iv) \( L \in FS \) if and only if each \( H \)-class of \( I \) is bridge.
Conversely, if \( I \) is a non-trivial group, then \( L \in FS \).
Scanning Conditions in terms of the syntactic monoid

Let $\overline{\cdot} : A^* \to A^*/\eta = M(L)$ denote the natural morphism of $A^*$ onto the syntactic monoid $M(L)$ of language $L \subseteq A^*$, so that

$$\overline{u} = \overline{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta$$

saturnates $L$.

We say a set $X \subseteq M(L)$ is a bridge if

$$X \eta^{-1} \cap L \neq \emptyset \& X \eta^{-1} \cap L' \neq \emptyset.$$ 

**Theorem**

Let $L$ be regular and let $I$ be the minimum ideal of $M = M(L)$ Then

(i) $L \in FP$ if and only if the $D$-class $I$ of $M$ is a bridge;
(ii) $L \in PP$ (resp. $SP$) if and only if each $R$-class (resp. $L$-class) of $I$ is bridge;
(iii) $L \in WS$ if and only if each $R$-class and each $L$-class of $I$ is bridge;
(iv) $L \in FS$ if and only if each $H$-class of $I$ is bridge.

Conversely, if $I$ is a non-trivial group, then $L \in FS$. 

Languages that require full scanning of words to determine
Let $\overline{\cdot} : A^* \rightarrow A^*/\eta = M(L)$ denote the natural morphism of $A^*$ onto the syntactic monoid $M(L)$ of language $L \subseteq A^*$, so that
\[
\overline{u} = \overline{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta \text{ saturates } L.
\]
We say a set $X \subseteq M(L)$ is a bridge if
\[
X\eta^{-1} \cap L \neq \emptyset \& X\eta^{-1} \cap L' \neq \emptyset.
\]

**Theorem**

Let $L$ be regular and let $I$ be the minimum ideal of $M = M(L)$. Then

(i) $L \in FP$ if and only if the $D$-class $I$ of $M$ is a bridge;
(ii) $L \in PP$ (resp. $SP$) if and only if each $R$-class (resp. $L$-class) of $I$ is bridge;
(iii) $L \in WS$ if and only if each $R$-class and each $L$-class of $I$ is bridge;
(iv) $L \in FS$ if and only if each $H$-class of $I$ is bridge.

Conversely, if $I$ is a non-trivial group, then $L \in FS$. 

Peter M. Higgins & Suheear Alwan
Languages that require full scanning of words to determine
Let $\bar{\cdot} : A^* \rightarrow A^*/\eta = M(L)$ denote the natural morphism of $A^*$ onto the syntactic monoid $M(L)$ of language $L \subseteq A^*$, so that

$$\bar{u} = \bar{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta$$ saturates $L$.

We say a set $X \subseteq M(L)$ is a bridge if

$$X \eta^{-1} \cap L \neq \emptyset \& X \eta^{-1} \cap L' \neq \emptyset.$$

**Theorem**

Let $L$ be regular and let $I$ be the minimum ideal of $M = M(L)$ Then

(i) $L \in FP$ if and only if the $D$-class $I$ of $M$ is a bridge;
(ii) $L \in PP$ (resp. $SP$) if and only if each $R$-class (resp. $L$-class) of $I$ is bridge;
(iii) $L \in WS$ if and only if each $R$-class and each $L$-class of $I$ is bridge;
(iv) $L \in FS$ if and only if each $H$-class of $I$ is bridge.

Conversely, if $I$ is a non-trivial group, then $L \in FS$. 

Languages that require full scanning of words to determine
Let \( \overline{\cdot} : A^* \to A^*/\eta = M(L) \) denote the natural morphism of \( A^* \) onto the syntactic monoid \( M(L) \) of language \( L \subseteq A^* \), so that

\[
\overline{u} = \overline{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta \text{saturates } L.
\]

We say a set \( X \subseteq M(L) \) is a bridge if

\[
X \eta^{-1} \cap L \neq \emptyset \& X \eta^{-1} \cap L' \neq \emptyset.
\]

**Theorem**

Let \( L \) be regular and let \( I \) be the minimum ideal of \( M = M(L) \). Then

(i) \( L \in FP \) if and only if the \( \mathcal{D} \)-class \( I \) of \( M \) is a bridge;
(ii) \( L \in PP \) (resp. \( SP \)) if and only if each \( \mathcal{R} \)-class (resp. \( \mathcal{L} \)-class) of \( I \) is bridge;
(iii) \( L \in WS \) if and only if each \( \mathcal{R} \)-class and each \( \mathcal{L} \)-class of \( I \) is bridge;
(iv) \( L \in FS \) if and only if each \( \mathcal{H} \)-class of \( I \) is bridge.

Conversely, if \( I \) is a non-trivial group, then \( L \in FS \).
Let $\overline{\cdot} : A^* \to A^*/\eta = M(L)$ denote the natural morphism of $A^*$ onto the syntactic monoid $M(L)$ of language $L \subseteq A^*$, so that

$$\overline{u} = \overline{v} \iff (puq \in L \iff pvq \in L, \forall p, q \in A^*) - \eta$$

saturates $L$.

We say a set $X \subseteq M(L)$ is a bridge if

$$X \eta^{-1} \cap L \neq \emptyset \& X \eta^{-1} \cap L' \neq \emptyset.$$

**Theorem**

Let $L$ be regular and let $I$ be the minimum ideal of $M = M(L)$. Then

(i) $L \in FP$ if and only if the $D$-class $I$ of $M$ is a bridge;
(ii) $L \in PP$ (resp. $SP$) if and only if each $R$-class (resp. $L$-class) of $I$ is bridge;
(iii) $L \in WS$ if and only if each $R$-class and each $L$-class of $I$ is bridge;
(iv) $L \in FS$ if and only if each $H$-class of $I$ is bridge.

Conversely, if $I$ is a non-trivial group, then $L \in FS$. 
Examples

(A) \( L(f, m, K) = \{ w \in A^*: |w|_f(\text{mod } m) \in K \} \) where \( f \in A^*, \ m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L', L^R, u^{-1}L, Lu^{-1} \) for any \( u \in A^* \);

(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:

(i) \( \{ w \in A^*: |w|_a = |w|_b \} \);

(ii) Primitive words;

(iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \bar{u} \in H \), a maximal subgroup of \( I \).

Take \( k \geq 1 \) such that \( \bar{u} = \bar{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \bar{u} \in H \),

whence \( H \) is not a bridge, contradiction!
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^* : |w|_{f(\text{mod } m)} \in K \} \) where \( f \in A^* \), \( m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L' \), \( L^R \), \( u^{-1}L \), \( Lu^{-1} \) for any \( u \in A^* \);

(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:
(i) \( \{ w \in A^* : |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \bar{u} \in H \), a maximal subgroup of \( I \).

Take \( k \geq 1 \) such that \( \bar{u} = \bar{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \bar{u} \in H \), whence \( H \) is not a bridge, contradiction!
Observations & Examples

Examples

(A) $L(f, m, K) = \{ w \in A^* : |w|_f \text{mod } m \in K \}$ where $f \in A^*$, $m \geq 2$ and $K$ a proper subset of $\{0, 1, \cdots, m-1\}$ is a full scan language.

(B) $L \in FS$ then so is $L'$, $L^R$, $u^{-1}L$, $Lu^{-1}$ for any $u \in A^*$;

(C) $L$ is full scan if and only if $u^{-1}Lv^{-1}$ is proper for all $u, v \in A^*$.

The following languages are not regular:

(i) $\{ w \in A^* : |w|_a = |w|_b \}$; (ii) Primitive words; (iii) Words with borders.

Argument for (ii): $L \in FS$ (easy to check) so suppose $L$ were regular and let $\overline{u} \in H$, a maximal subgroup of $I$.

Take $k \geq 1$ such that $\overline{u} = \overline{u}^{k+1}$: $u^{k+1} \in L' \Rightarrow u \in L' \forall \overline{u} \in H$, whence $H$ is not a bridge, contradiction!
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^* : |w|_f(\text{mod } m) \in K \} \) where \( f \in A^* \), \( m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L', L^R, u^{-1}L, Lu^{-1} \) for any \( u \in A^* \);

(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:

(i) \( \{ w \in A^* : |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \bar{u} \in H \), a maximal subgroup of \( I \).

Take \( k \geq 1 \) such that \( \bar{u} = \bar{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \bar{u} \in H \), whence \( H \) is not a bridge, contradiction!
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^* : |w|_f \text{ (mod } m) \in K \} \) where \( f \in A^* \), \( m \geq 2 \) and \( K \) a proper subset of \( \{0,1,\cdots,m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L' \), \( L^R \), \( u^{-1}L \), \( Lu^{-1} \) for any \( u \in A^* \);
(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u,v \in A^* \).

The following languages are not regular:
(i) \( \{ w \in A^* : |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \bar{u} \in H \), a maximal subgroup of \( I \).
Take \( k \geq 1 \) such that \( \bar{u} = \bar{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \bar{u} \in H \), whence \( H \) is not a bridge, contradiction!
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^*: |w|_f \text{mod } m \in K \} \) where \( f \in A^*, \ m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L', L^R, u^{-1}L, Lu^{-1} \) for any \( u \in A^* \);

(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:

(i) \( \{ w \in A^*: |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \bar{u} \in H \), a maximal subgroup of \( I \).

Take \( k \geq 1 \) such that \( \bar{u} = \bar{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \bar{u} \in H \), whence \( H \) is not a bridge, contradiction!

Peter M. Higgins & Suheer Alwan

Languages that require full scanning of words to determine
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^* : |w|_f(\text{mod } m) \in K \} \) where \( f \in A^* \), \( m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in \text{FS} \) then so is \( L' \), \( L^R \), \( u^{-1}L \), \( Lu^{-1} \) for any \( u \in A^* \);
(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:
(i) \( \{ w \in A^* : |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in \text{FS} \) (easy to check) so suppose \( L \) were regular and let \( \overline{u} \in H \), a maximal subgroup of \( I \).
Take \( k \geq 1 \) such that \( \overline{u} = \overline{u}^{k+1} \). \( u^{k+1} \in L' \Rightarrow u \in L' \forall \overline{u} \in H \), whence \( H \) is not a bridge, contradiction!
Observations & Examples

Examples

(A) \( L(f, m, K) = \{ w \in A^* : |w|_f \text{mod} m \in K \} \) where \( f \in A^* \), \( m \geq 2 \) and \( K \) a proper subset of \( \{0, 1, \cdots, m-1\} \) is a full scan language.

(B) \( L \in FS \) then so is \( L' \), \( L^R \), \( u^{-1}L \), \( Lu^{-1} \) for any \( u \in A^* \);

(C) \( L \) is full scan if and only if \( u^{-1}Lv^{-1} \) is proper for all \( u, v \in A^* \).

The following languages are not regular:

(i) \( \{ w \in A^* : |w|_a = |w|_b \} \); (ii) Primitive words; (iii) Words with borders.

Argument for (ii): \( L \in FS \) (easy to check) so suppose \( L \) were regular and let \( \overline{u} \in H \), a maximal subgroup of \( I \).

Take \( k \geq 1 \) such that \( \overline{u} = \overline{u}^{k+1} \): \( u^{k+1} \in L' \Rightarrow u \in L' \forall \overline{u} \in H \), whence \( H \) is not a bridge, contradiction!
**Chameleon sets**

**Definition**

A set $C \subseteq A^*$ is called a *chameleon set* if $\forall u, v \in A^* \exists u', v' \in A^*$ such that $uu' A^* v' v \cap C = \emptyset$. Equivalently, each two-sided quotient $u^{-1} C v^{-1}$ has an empty two-sided quotient $u'^{-1} (u^{-1} C v^{-1}) v'^{-1}$.

**Examples**

finite sets, complements of ideals.

CP closed under sublanguages, finite unions, reversals, left quotients and right quotients, and so forms a topology on $A^*$. 

Languages that require full scanning of words to determine
Chameleons sets

Definition

A set $C \subseteq A^*$ is called a chameleon set if $\forall u, v \in A^* \exists u', v' \in A^*$ such that $uu' A^* v' v \cap C = \emptyset$. Equivalently, each two-sided quotient $u^{-1} C v^{-1}$ has an empty two-sided quotient $u'^{-1} (u^{-1} C v^{-1}) v'^{-1}$.

Examples

finite sets, complements of ideals.

CP closed under sublanguages, finite unions, reversals, left quotients and right quotients, and so forms a topology on $A^*$.
Chameleon sets

**Definition**

A set $C \subseteq A^*$ is called a *chameleon set* if $\forall u, v \in A^* \exists u', v' \in A^*$ such that $uu' A^* v' v \cap C = \emptyset$. Equivalently, each two-sided quotient $u^{-1} C v^{-1}$ has an empty two-sided quotient $u'^{-1} (u^{-1} C v^{-1}) v'^{-1}$.

**Examples**

finite sets, complements of ideals.

CP closed under sublanguages, finite unions, reversals, left quotients and right quotients, and so forms a topology on $A^*$. 
Chameleon sets

**Definition**

A set $C \subseteq A^*$ is called a *chameleon set* if $\forall u, v \in A^* \exists u', v' \in A^*$ such that $uu' A^* v' v \cap C = \emptyset$. Equivalently, each two-sided quotient $u^{-1} C v^{-1}$ has an empty two-sided quotient $u'^{-1} (u^{-1} C v^{-1}) v'^{-1}$.

**Examples**

finite sets, complements of ideals.

CP closed under sublanguages, finite unions, reversals, left quotients and right quotients, and so forms a topology on $A^*$.
Why chameleon?

Theorem

Let $L$ be full scan and $C$ chameleon. Then $L \cup C$ and $L \setminus C$ are full scan.

Proof Let $u, v \in A^*$. Since $C \in CP \exists u'v' \in A^*$ such that $uu'A^*v'v \cap C = \emptyset$. Since $L \in FS \exists x, x' \in A^*$ such that $(uu')x(v'v) \in L$ and $(uu')x'(v'v) \in L'$. But then:

$$u(u'xv')v \in L \cup C \text{ and } u(u'x'v')v \in L' \cap C' = (L \cup C)', \text{ thus } L \cup C \in FS.$$

$L \in FS \Rightarrow L' \cup C \in FS \Rightarrow (L' \cup C)' \in FS \Rightarrow L \cap C' = L \setminus C \in FS.$
Why chameleon?

**Theorem**

Let \( L \) be full scan and \( C \) chameleon. Then \( L \cup C \) and \( L \setminus C \) are full scan.

**Proof** Let \( u, v \in A^* \). Since \( C \in CP \) \( \exists u'v' \in A^* \) such that \( uu'A^*v'v \cap C = \emptyset \). Since \( L \in FS \) \( \exists x, x' \in A^* \) such that \( (uu')x(v'v) \in L \) and \( (uu')x'(v'v) \in L' \). But then:

\[
u(u'xv')v \in L \cup C \quad \text{and} \quad u(u'x'v')v \in L' \cap C' = (L \cup C)', \text{ thus } L \cup C \in FS.
\]

\[
L \in FS \Rightarrow L' \cup C \in FS \Rightarrow (L' \cup C)' \in FS \Rightarrow L \cap C' = L \setminus C \in FS. \quad \square
\]
Why chameleon?

**Theorem**

Let \( L \) be full scan and \( C \) chameleon. Then \( L \cup C \) and \( L \setminus C \) are full scan.

**Proof** Let \( u, v \in A^* \). Since \( C \in CP \), \( \exists u'v' \in A^* \) such that \( uu'A^*v'v \cap C = \emptyset \). Since \( L \in FS \), \( \exists x, x' \in A^* \) such that \( (uu')x(v'v) \in L \) and \( (uu')x'(v'v) \in L' \). But then:

\[
u(u'xv')v \in L \cup C \text{ and } u(u'x'v')v \in L' \cap C' = (L \cup C)', \text{ thus } L \cup C \in FS.\]

\[
L \in FS \Rightarrow L' \cup C \in FS \Rightarrow (L' \cup C)' \in FS \Rightarrow L \cap C' = L \setminus C \in FS. \qed
\]
Why chameleon?

Theorem

Let $L$ be full scan and $C$ chameleon. Then $L \cup C$ and $L \setminus C$ are full scan.

Proof Let $u, v \in A^*$. Since $C \in CP \exists u'v' \in A^*$ such that $uu'A^*v'v \cap C = \emptyset$. Since $L \in FS \exists x, x' \in A^*$ such that $(uu')x(v'v) \in L$ and $(uu')x'(v'v) \in L'$. But then:

$$u(u'xv')v \in L \cup C \text{ and } u(u'x'v')v \in L' \cap C' = (L \cup C)'$$

Thus $L \cup C \in FS$.

$L \in FS \Rightarrow L' \cup C \in FS \Rightarrow (L' \cup C)' \in FS \Rightarrow L \cap C' = L \setminus C \in FS$.  □
Theorem

A regular chameleon set has none of the five scanning properties.

In consequence, none of the following languages are regular.

Examples

(A) Language of all palindromes is weak scan and chameleon;

(B) The language of all Lyndon words is chameleon and has the factor property;

(C) The Dyck language (of all meaningful parentheses) is chameleon and has the factor property.
Theorem

A regular chameleon set has none of the five scanning properties.

In consequence, none of the following languages are regular.

Examples

(A) Language of all palindromes is weak scan and chameleon;

(B) The language of all Lyndon words is chameleon and has the factor property;

(C) The Dyck language (of all meaningful parentheses) is chameleon and has the factor property.
Theorem

A regular chameleon set has none of the five scanning properties.

In consequence, none of the following languages are regular.

Examples

(A) Language of all palindromes is weak scan and chameleon;

(B) The language of all Lyndon words is chameleon and has the factor property;
(C) The Dyck language (of all meaningful parentheses) is chameleon and has the factor property.
Theorem

A regular chameleon set has none of the five scanning properties.

In consequence, none of the following languages are regular.

Examples

(A) Language of all palindromes is weak scan and chameleon;

(B) The language of all Lyndon words is chameleon and has the factor property;

(C) The Dyck language (of all meaningful parentheses) is chameleon and has the factor property.
Theorem

A regular chameleon set has none of the five scanning properties.

In consequence, none of the following languages are regular.

Examples

(A) Language of all palindromes is weak scan and chameleon;

(B) The language of all Lyndon words is chameleon and has the factor property;

(C) The Dyck language (of all meaningful parentheses) is chameleon and has the factor property.
Letter scan languages

Take the definition of full scan language and strengthen the condition \( uxv \in L, ux'v \in L' \) by insisting that \( x \in A \). If we take \( A = \{a, b\} \) the FSL languages are as follows.

**Definitions**
Let \( E = \{w \in A^* : \lfloor w \rfloor_a \equiv 0 \pmod{2}\} \) and \( O = A^* \setminus E \). Let \( E_n = E \cap A^n \), \( O_n = O \cap A^n \). For any \( L \subseteq A^* \) let \( L_n = L \cap A^n \).

**Theorem**
\( L \) is FSL if and only if \( L_n \in \{E_n, O_n\} \forall n \geq 0 \).

There is then a one-to-one correspondence between FSL languages \( L \) and real numbers \( s_L \) in the interval \([0, 2]\): the initial digit determines the presence or absence of \( \varepsilon \), the \( n \)th digit is 0 if and only if \( L_n = E_n \).

**Theorem**
Let \( L \in \text{FSL} \). Then \( L \) is regular if and only if \( s_L \in \mathbb{Q} \).
Letter scan languages

Take the definition of full scan language and strengthen the condition $uxv \in L, ux'v \in L'$ by insisting that $x \in A$. If we take $A = \{a, b\}$ the FSL languages are as follows.

**Definitions**

Let $E = \{w \in A^* : |w|_a \equiv 0 \pmod{2}\}$ and $O = A^* \setminus E$. Let $E_n = E \cap A^n$, $O_n = O \cap A^n$. For any $L \subseteq A^*$ let $L_n = L \cap A^n$.

**Theorem**

$L$ is FSL if and only if $L_n \in \{E_n, O_n\} \forall n \geq 0$.

There is then a one-to-one correspondence between FSL languages $L$ and real numbers $s_L$ in the interval $[0, 2]$: the initial digit determines the presence or absence of $\varepsilon$, the $n$th digit is 0 if and only if $L_n = E_n$.

**Theorem**

Let $L \in \text{FSL}$. Then $L$ is regular if and only if $s_L \in \mathbb{Q}$.
Take the definition of full scan language and strengthen the condition \(uxv \in L, ux'v \in L'\) by insisting that \(x \in A\). If we take \(A = \{a, b\}\) the FSL languages are as follows.

**Definitions**

Let \(E = \{w \in A^* : |w|_a \equiv 0 \pmod{2}\}\) and \(O = A^* \setminus E\). Let \(E_n = E \cap A^n, O_n = O \cap A^n\). For any \(L \subseteq A^*\) let \(L_n = L \cap A^n\).

**Theorem**

\(L\) is FSL if and only if \(L_n \in \{E_n, O_n\} \forall n \geq 0\).

There is then a one-to-one correspondence between FSL languages \(L\) and real numbers \(s_L\) in the interval \([0, 2]:\) the initial digit determines the presence or absence of \(\varepsilon\), the \(n\)th digit is 0 if and only if \(L_n = E_n\).

**Theorem**

Let \(L \in \text{FSL}\). Then \(L\) is regular if and only if \(s_L \in \mathbb{Q}\).
Letter scan languages

Take the definition of full scan language and strengthen the condition $uxv \in L, ux'v \in L'$ by insisting that $x \in A$. If we take $A = \{a, b\}$ the FSL languages are as follows.

**Definitions**

Let $E = \{w \in A^* : |w|_a \equiv 0 \pmod{2}\}$ and $O = A^* \setminus E$. Let $E_n = E \cap A^n$, $O_n = O \cap A^n$. For any $L \subseteq A^*$ let $L_n = L \cap A^n$.

**Theorem**

$L$ is FSL if and only if $L_n \in \{E_n, O_n\} \forall n \geq 0$.

There is then a one-to-one correspondence between FSL languages $L$ and real numbers $s_L$ in the interval $[0, 2]$: the initial digit determines the presence or absence of $\varepsilon$, the $n$th digit is 0 if and only if $L_n = E_n$.

**Theorem**

Let $L \in \text{FSL}$. Then $L$ is regular if and only if $s_L \in \mathbb{Q}$.
Take the definition of full scan language and strengthen the condition \( uxv \in L, ux'v \in L' \) by insisting that \( x \in A \). If we take \( A = \{a, b\} \) the FSL languages are as follows.

**Definitions**

Let \( E = \{ w \in A^* : |w|_a \equiv 0 \, (\text{mod} \, 2) \} \) and \( O = A^* \setminus E \). Let \( E_n = E \cap A^n \), \( O_n = O \cap A^n \). For any \( L \subseteq A^* \) let \( L_n = L \cap A^n \).

**Theorem**

\( L \) is FSL if and only if \( L_n \in \{E_n, O_n\} \forall n \geq 0 \).

There is then a one-to-one correspondence between FSL languages \( L \) and real numbers \( s_L \) in the interval \([0,2]\): the initial digit determines the presence or absence of \( \varepsilon \), the \( n \)th digit is 0 if and only if \( L_n = E_n \).

**Theorem**

Let \( L \in \text{FSL} \). Then \( L \) is regular if and only if \( s_L \in \mathbb{Q} \).
Languages that require full scanning of words to determine
The universal automaton \( \mathcal{U} \) will recognize a given FSL language \( L \) by putting \( n \) or \( n' \in T \) according as \( L_n = E_n \) or \( L_n = O_n \). In effect we just read \( s_L \) into \( \mathcal{U} \).

Let \( s_L = e_0 \cdot e_1 e_2 \cdots \). If \( s_L \in \mathbb{Q} \) with \( e_k = e_{k+n} \) for some minimum \( k \) and \( n \), then we may identify the pairs of states \( (k, k+n) \) and \( (k', k+n') \). The resulting finite automaton \( \mathcal{A}(L) \) has the form of a cylinder with a trailing tape that leads to a point (0): and \( \mathcal{A}(L) \) is the minimal automaton of \( L \) EXCEPT if \( s_L \) has the form:

\[
s_L = \frac{1}{2^k} \left( n + \frac{t}{1+2^r} \right), 0 \leq k, 0 \leq n \leq 2^k - 1, 1 \leq r, 1 \leq t \leq 2^r.
\]
The universal automaton $\mathcal{U}$ will recognize a given FSL language $L$ by putting $n$ or $n' \in T$ according as $L_n = E_n$ or $L_n = O_n$. In effect we just read $s_L$ into $\mathcal{U}$.

Let $s_L = e_0 \cdot e_1 e_2 \cdots$. If $s_L \in \mathbb{Q}$ with $e_k = e_{k+n}$ for some minimum $k$ and $n$, then we may identify the pairs of states $(k, k+n)$ and $(k', (k+n)')$. The resulting finite automaton $\mathcal{A}(L)$ has the form of a cylinder with a trailing tape that leads to a point (0):

and $\mathcal{A}(L)$ is the minimal automaton of $L$ EXCEPT if $s_L$ has the form:

$$s_L = \frac{1}{2^k} \left( n + \frac{t}{1+2^r} \right), \quad 0 \leq k, 0 \leq n \leq 2^k - 1, 1 \leq r, 1 \leq t \leq 2^r.$$
The universal automaton $\mathcal{U}$ will recognize a given FSL language $L$ by putting $n$ or $n' \in T$ according as $L_n = E_n$ or $L_n = O_n$. In effect we just read $s_L$ into $\mathcal{U}$. Let $s_L = e_0 \cdot e_1 e_2 \cdots$. If $s_L \in \mathbb{Q}$ with $e_k = e_{k+n}$ for some minimum $k$ and $n$, then we may identify the pairs of states $(k, k+n)$ and $(k', (k+n)')$. The resulting finite automaton $\mathcal{A}(L)$ has the form of a cylinder with a trailing tape that leads to a point $(0)$: and $\mathcal{A}(L)$ is the minimal automaton of $L$ EXCEPT if $s_L$ has the form:

$$s_L = \frac{1}{2^k} \left( n + \frac{t}{1+2^r} \right), 0 \leq k, 0 \leq n \leq 2^k - 1, 1 \leq r, 1 \leq t \leq 2^r.$$
The universal automaton $U$ will recognize a given FSL language $L$ by putting $n$ or $n' \in T$ according as $L_n = E_n$ or $L_n = O_n$. In effect we just read $s_L$ into $U$.

Let $s_L = e_0 \cdot e_1 e_2 \cdots$. If $s_L \in \mathbb{Q}$ with $e_k = e_{k+n}$ for some minimum $k$ and $n$, then we may identify the pairs of states $(k, k+n)$ and $(k', (k+n)')$. The resulting finite automaton $A(L)$ has the form of a cylinder with a trailing tape that leads to a point (0):

and $A(L)$ is the minimal automaton of $L$ EXCEPT if $s_L$ has the form:

$$s_L = \frac{1}{2^k} \left( n + \frac{t}{1+2^r} \right), 0 \leq k, 0 \leq n \leq 2^k - 1, 1 \leq r, 1 \leq t \leq 2^r.$$
This special case is where the recurring part of $s_L$ has the form $z\overline{z}$ where $\overline{z}$ is defined by $z + \overline{z} = 11 \cdots 1$ (with $2r$ 1’s), so that $\overline{z}$ is the obtained from $z$ by interchanging the symbols 0 and 1 throughout. In this case the cylinder of circumference $2r$ may be replaced by a Mobius strip of edge length $2r$:

We may identify the pairs of states $(k + r, k')$ and $((k + r)', k)$, the resulting half-twist giving the form of a Mobius strip.
This special case is where the recurring part of $s_L$ has the form $z\bar{z}$ where $\bar{z}$ is defined by $z + \bar{z} = 11 \cdots 1$ (with $2r$ 1’s), so that $\bar{z}$ is the obtained from $z$ by interchanging the symbols 0 and 1 throughout. In this case the cylinder of circumference $2r$ may be replaced by a Mobius strip of edge length $2r$:

We may identify the pairs of states $(k + r, k')$ and $((k + r)', k)$, the resulting half-twist giving the form of a Mobius strip.
This special case is where the recurring part of $s_L$ has the form $z\bar{z}$ where $\bar{z}$ is defined by $z + \bar{z} = 11 \cdots 1$ (with $2r$ 1’s), so that $\bar{z}$ is the obtained from $z$ by interchanging the symbols 0 and 1 throughout. In this case the cylinder of circumference $2r$ may be replaced by a Mobius strip of edge length $2r$:

We may identify the pairs of states $(k + r, k')$ and $((k + r)', k)$, the resulting half-twist giving the form of a Mobius strip.