Automata Theory :: Properties of Context-Free Languages

Jörg Endrullis

Vrije Universiteit Amsterdam

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$.

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$:

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \rightarrow S_1 \mid S_2$, and pick S as start variable.

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \to S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$:

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \rightarrow S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$: Add $S \rightarrow S_1 S_2$, and pick S as start variable.

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \to S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$: Add $S \to S_1 S_2$, and pick S as start variable.
- L_1^*:

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \to S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$: Add $S \to S_1 S_2$, and pick S as start variable.
- L_1^*: Add $S \to S_1 S \mid \lambda$, and pick S as start variable.

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \to S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$: Add $S \to S_1 S_2$, and pick S as start variable.
- L_1^*: Add $S \to S_1 S \mid \lambda$, and pick S as start variable.
- L_1^R:

Basic Properties of Context-Free Languages

Theorem

If L_1 and L_2 are context-free, then also

$$L_1 \cup L_2 \qquad L_1 L_2 \qquad L_1^* \qquad L_1^R$$

Proof.

Let G_i be a context-free grammar with start variable S_i s.t.

$$L_i = L(G_i)$$

for $i = 1, 2$. Let G_1 and G_2 have no variables in common.

- $L_1 \cup L_2$: Add rules $S \to S_1 \mid S_2$, and pick S as start variable.
- $L_1 L_2$: Add $S \to S_1 S_2$, and pick S as start variable.
- L_1^*: Add $S \to S_1 S \mid \lambda$, and pick S as start variable.
- L_1^R: Reverse all rules ($x \to y$ becomes $x^R \to y^R$).

□

Basic Properties of Context-Free Languages

The intersection $L_1 \cap L_2$ is **not** always context-free.
(for context free languages L_1 and L_2)

Basic Properties of Context-Free Languages

The intersection $L_1 \cap L_2$ is not always context-free.
(for context free languages L_1 and L_2)

The languages L_1 and L_2 are context-free:

$$L_1 = \{\, a^n b^n c^m \mid n \geq 0 \wedge m \geq 0 \,\}$$
$$L_2 = \{\, a^n b^m c^m \mid n \geq 0 \wedge m \geq 0 \,\}$$

Basic Properties of Context-Free Languages

The intersection $L_1 \cap L_2$ is not always context-free.
(for context free languages L_1 and L_2)

The languages L_1 and L_2 are context-free:

$$L_1 = \{\, a^n b^n c^m \mid n \geq 0 \wedge m \geq 0 \,\}$$
$$L_2 = \{\, a^n b^m c^m \mid n \geq 0 \wedge m \geq 0 \,\}$$

However $L_1 \cap L_2 = \{\, a^n b^n c^n \mid n \geq 0 \,\}$ is **not** context-free.

Basic Properties of Context-Free Languages

The intersection $L_1 \cap L_2$ is not always context-free.
(for context free languages L_1 and L_2)

The languages L_1 and L_2 are context-free:

$$L_1 = \{ a^n b^n c^m \mid n \geq 0 \wedge m \geq 0 \}$$
$$L_2 = \{ a^n b^m c^m \mid n \geq 0 \wedge m \geq 0 \}$$

However $L_1 \cap L_2 = \{ a^n b^n c^n \mid n \geq 0 \}$ is **not** context-free.

Also $\overline{L_1}$ and $L_1 \setminus L_2$ are not always context-free.
(for context free languages L_1 and L_2)

Basic Properties of Context-Free Languages

The intersection $L_1 \cap L_2$ is not always context-free.
(for context free languages L_1 and L_2)

The languages L_1 and L_2 are context-free:

$$L_1 = \{ a^n b^n c^m \mid n \geq 0 \wedge m \geq 0 \}$$
$$L_2 = \{ a^n b^m c^m \mid n \geq 0 \wedge m \geq 0 \}$$

However $L_1 \cap L_2 = \{ a^n b^n c^n \mid n \geq 0 \}$ is **not** context-free.

Also $\overline{L_1}$ and $L_1 \setminus L_2$ are not always context-free.
(for context free languages L_1 and L_2)

Namely, we have:

$$L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}} \qquad \overline{L_1} = \Sigma^* \setminus L_1$$

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

We construct an NPDA $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \widehat{\delta}, \widehat{q}_0, z, \widehat{F})$ where

$$\widehat{Q} = Q \times R \qquad \widehat{q}_0 = (q_0, r_0) \qquad \widehat{F} = F \times G$$

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

We construct an NPDA $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \widehat{\delta}, \widehat{q}_0, z, \widehat{F})$ where

$$\widehat{Q} = Q \times R \qquad \widehat{q}_0 = (q_0, r_0) \qquad \widehat{F} = F \times G$$

The transition function $\widehat{\delta}$ is defined by:

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

We construct an NPDA $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \widehat{\delta}, \widehat{q}_0, z, \widehat{F})$ where

$$\widehat{Q} = Q \times R \qquad \widehat{q}_0 = (q_0, r_0) \qquad \widehat{F} = F \times G$$

The transition function $\widehat{\delta}$ is defined by:

- $\widehat{M}$: $(q, r) \xrightarrow{a[b/v]} (q', r')$ if M: $q \xrightarrow{a[b/v]} q'$ and N: $r \xrightarrow{a} r'$

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

We construct an NPDA $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \widehat{\delta}, \widehat{q}_0, z, \widehat{F})$ where

$$\widehat{Q} = Q \times R \qquad \widehat{q}_0 = (q_0, r_0) \qquad \widehat{F} = F \times G$$

The transition function $\widehat{\delta}$ is defined by:

- $\widehat{M}$: $(q, r) \xrightarrow{a[b/v]} (q', r')$ if M: $q \xrightarrow{a[b/v]} q'$ and N: $r \xrightarrow{a} r'$
- $\widehat{M}$: $(q, r) \xrightarrow{\lambda[b/v]} (q', r)$ if M: $q \xrightarrow{\lambda[b/v]} q'$

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \cap L_2$ is context-free.

Construction

Let

- $M = (Q, \Sigma, \Gamma, \delta, q_0, z, F)$ be an NPDA accepting L_1, and
- $N = (R, \Sigma, \epsilon, r_0, G)$ a DFA accepting L_2.

We construct an NPDA $\widehat{M} = (\widehat{Q}, \Sigma, \Gamma, \widehat{\delta}, \widehat{q}_0, z, \widehat{F})$ where

$$\widehat{Q} = Q \times R \qquad \widehat{q}_0 = (q_0, r_0) \qquad \widehat{F} = F \times G$$

The transition function $\widehat{\delta}$ is defined by:

- $\widehat{M}$: $(q, r) \xrightarrow{a[b/v]} (q', r')$ if M: $q \xrightarrow{a[b/v]} q'$ and N: $r \xrightarrow{a} r'$
- $\widehat{M}$: $(q, r) \xrightarrow{\lambda[b/v]} (q', r)$ if M: $q \xrightarrow{\lambda[b/v]} q'$

Then $L(\widehat{M}) = L(M) \cap L(N)$.

Question

Question

Why does the construction not work for two NPDA's?
(instead of an NPDA and a DFA)

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \setminus L_2$ is context-free.

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \backslash L_2$ is context-free.

Proof.

$\overline{L_2}$ is regular, thus $L_1 \backslash L_2 = L_1 \cap \overline{L_2}$ is context-free. □

Basic Properties of Context-Free Languages

Theorem

If L_1 is context-free and L_2 regular, then $L_1 \backslash L_2$ is context-free.

Proof.

$\overline{L_2}$ is regular, thus $L_1 \backslash L_2 = L_1 \cap \overline{L_2}$ is context-free. □

$L_2 \backslash L_1$ is **not** always context-free. Namely

$$\overline{L_1} = \Sigma^* \backslash L_1$$

Applications

$L \setminus \{ \lambda \}$ is context-free for every context-free language L.

Applications

$L \setminus \{ \lambda \}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Applications

$L \setminus \{\lambda\}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Show that the language

$$L = \{ w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w) \}$$

is **not** context-free.

Applications

$L \setminus \{\lambda\}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Show that the language

$$L = \{ w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w) \}$$

is **not** context-free.

For a contradiction, assume L was context-free.

Applications

$L \setminus \{\lambda\}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Show that the language

$$L = \{ w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w) \}$$

is **not** context-free.

For a contradiction, assume L was context-free.

The language $L(a^* b^* c^*)$ is regular, thus

$$L \cap L(a^* b^* c^*) = \{ a^n b^n c^n \mid n \geq 0 \}$$

would be context-free.

Applications

$L \setminus \{\lambda\}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Show that the language

$$L = \{ w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w) \}$$

is **not** context-free.

For a contradiction, assume L was context-free.

The language $L(a^* b^* c^*)$ is regular, thus

$$L \cap L(a^* b^* c^*) = \{ a^n b^n c^n \mid n \geq 0 \}$$

would be context-free. However, we know that it is not.

Applications

$L \setminus \{ \lambda \}$ is context-free for every context-free language L.

$\{ a^n b^n \mid n \geq 1000 \}$ is context-free.

Show that the language

$$L = \{ w \in \{a, b, c\}^* \mid n_a(w) = n_b(w) = n_c(w) \}$$

is **not** context-free.

For a contradiction, assume L was context-free.

The language $L(a^* b^* c^*)$ is regular, thus

$$L \cap L(a^* b^* c^*) = \{ a^n b^n c^n \mid n \geq 0 \}$$

would be context-free. However, we know that it is not.

Contradiction. Thus L is not context-free.

Decidability

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?
5. Does $L(G) = L(H)$ hold ?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?
5. Does $L(G) = L(H)$ hold ?
6. Is $L(G) \cap L(H)$ empty ?

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?
5. Does $L(G) = L(H)$ hold ?
6. Is $L(G) \cap L(H)$ empty ?

Only the first two questions are decidable.

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?
5. Does $L(G) = L(H)$ hold ?
6. Is $L(G) \cap L(H)$ empty ?

Only the first two questions are decidable.

Remove all λ and unit productions.

1. $\{ v \mid S \Rightarrow^* v,\ |v| \leq |w| \}$ can be computed in finite time.
2. $L(G)$ is empty $\iff$ starting variable is useless.

Basic Questions about Context-Free Grammars

Given context-free grammar G and H.

Which of the following questions are **decidable**?

1. Given $w \in \Sigma^*$, do we have $w \in L(G)$?
2. Is $L(G)$ empty ?
3. Does $L(G) = \Sigma^*$ hold ?
4. Does $L(G)$ contain a palindrome ($w = w^R$) ?
5. Does $L(G) = L(H)$ hold ?
6. Is $L(G) \cap L(H)$ empty ?

Only the first two questions are decidable.

Remove all λ and unit productions.

1. $\{ v \mid S \Rightarrow^* v,\ |v| \leq |w| \}$ can be computed in finite time.
2. $L(G)$ is empty $\iff$ starting variable is useless.

Surprisingly all other questions are undecidable.