

# Conjunctive grammars generate non-regular unary languages

Artur Jež

August 21, 2007

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.
- Extension of **Context-free grammars** by an **intersection** in a rule body.

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.
- Extension of **Context-free grammars** by an **intersection** in a rule body.
- Productions of the form

$$A \rightarrow \alpha_1 \& \alpha_2 \& \dots \& \alpha_k, \quad \text{for } \alpha_i \in (\Sigma \cup N)^* .$$

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.
- Extension of **Context-free grammars** by an **intersection** in a rule body.
- Productions of the form

$$A \rightarrow \alpha_1 \& \alpha_2 \& \dots \& \alpha_k, \quad \text{for } \alpha_i \in (\Sigma \cup N)^* .$$

- Intuition of the **semantics**:

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.
- Extension of **Context-free grammars** by an **intersection** in a rule body.
- Productions of the form

$$A \rightarrow \alpha_1 \& \alpha_2 \& \dots \& \alpha_k, \quad \text{for } \alpha_j \in (\Sigma \cup N)^* .$$

- Intuition of the **semantics**:
  - ▶  $w$  is derived such production iff it is derived by **each**  $\alpha_j$

# History

- **Conjunctive grammars** introduced in 2001 by A. Okhotin.
- Extension of **Context-free grammars** by an **intersection** in a rule body.
- Productions of the form

$$A \rightarrow \alpha_1 \& \alpha_2 \& \dots \& \alpha_k, \quad \text{for } \alpha_j \in (\Sigma \cup N)^* .$$

- Intuition of the **semantics**:
  - ▶  $w$  is derived such production iff it is derived by **each**  $\alpha_j$
  - ▶  $w$  is derived from  $\alpha_j = N_1 \cdot N_2 \cdot \dots \cdot N_k$  iff  $w = w_1 w_2 \dots w_k$  and  $w_j$  is derived from  $N_j$  for each  $j$

# Example

## Example

$$\Sigma = \{a, b, c\},$$

$$N = \{S, B, C, E, A\}$$

$$S \rightarrow (AE)\&(BC)$$

$$A \rightarrow aA|\epsilon$$

$$B \rightarrow aBb|\epsilon$$

$$C \rightarrow cC|\epsilon$$

$$E \rightarrow bEc|\epsilon$$



# Example

## Example

$$\Sigma = \{a, b, c\},$$

$$N = \{S, B, C, E, A\}$$

$$S \rightarrow (AE) \& (BC) \qquad \{a^n b^n c^n : n \in \mathbb{N}\}$$

$$A \rightarrow aA | \epsilon \qquad a^*$$

$$B \rightarrow aBb | \epsilon \qquad \{a^n b^n : n \in \mathbb{N}\}$$

$$C \rightarrow cC | \epsilon \qquad c^*$$

$$E \rightarrow bEc | \epsilon \qquad \{b^n c^n : n \in \mathbb{N}\}$$

# Motivation

- **natural extension** of CFG

# Motivation

- **natural extension** of CFG
- very close connection to **language equations**

# Motivation

- **natural extension** of CFG
- very close connection to **language equations**
- from **possible extensions** of CFG this keeps the meaning of **language equations**

# Motivation

- **natural extension** of CFG
- very close connection to **language equations**
- from **possible extensions** of CFG this keeps the meaning of **language equations**
- good **parsing** properties

# Formal syntax

## Definition

A **conjunctive grammar** is a  $\langle \Sigma, N, S, P \rangle$  where

# Formal syntax

## Definition

A **conjunctive grammar** is a  $\langle \Sigma, N, S, P \rangle$  where

- $\Sigma$  is a finite alphabet

# Formal syntax

## Definition

A **conjunctive grammar** is a  $\langle \Sigma, N, S, P \rangle$  where

- $\Sigma$  is a finite alphabet
- $N$ —set of non-terminal symbols



# Formal syntax

## Definition

A **conjunctive grammar** is a  $\langle \Sigma, N, S, P \rangle$  where

- $\Sigma$  is a finite alphabet
- $N$ —set of non-terminal symbols
- $S$ —starting symbol

# Formal syntax

## Definition

A **conjunctive grammar** is a  $\langle \Sigma, N, S, P \rangle$  where

- $\Sigma$  is a finite alphabet
- $N$ —set of non-terminal symbols
- $S$ —starting symbol
- $P$ —set of productions of a form

$$A \rightarrow \alpha_1 \& \alpha_2 \& \dots \& \alpha_k, \quad \alpha_i \in (\Sigma \cup N)^*$$

# Rewriting

## Semantics

By *term rewriting*.

# Rewriting

## Semantics

By *term rewriting*.

Generalizes the **Chomsky rewriting**.

# Rewriting

## Semantics

By *term rewriting*.

Generalizes the **Chomsky rewriting**.

Drawbacks

# Rewriting

## Semantics

By *term rewriting*.

Generalizes the **Chomsky rewriting**.

Drawbacks

- There are more generalizations.

# Rewriting

## Semantics

By *term rewriting*.

Generalizes the **Chomsky rewriting**.

Drawbacks

- There are more generalizations.
- Slightly problematic to handle.

# Language equations

## Semantics

*With each nonterminal  $A$  we associate a language  $L_A$ .*



# Language equations

## Semantics

*With each nonterminal  $A$  we associate a language  $L_A$ . The rule*

$$A \rightarrow B \& CD | a$$

*is replaced by*

$$L_A = (L_B \cap L_A \cdot L_D) \cup \{a\}$$

# Language equations

## Semantics

With each nonterminal  $A$  we associate a language  $L_A$ . The rule

$$A \rightarrow B \& CD | a$$

is replaced by

$$L_A = (L_B \cap L_A \cdot L_D) \cup \{a\}$$

The language corresponding to the component  $L_S$  in the **least solution** of the system is a language of the grammar.

# Language equations

## Semantics

With each nonterminal  $A$  we associate a language  $L_A$ . The rule

$$A \rightarrow B \& CD | a$$

is replaced by

$$L_A = (L_B \cap L_A \cdot L_D) \cup \{a\}$$

The language corresponding to the component  $L_S$  in the **least solution** of the system is a language of the grammar.

## Remark

In the CFG case the only allowed operations are  $\cup$  and  $\cdot$ .

# Example revisited

## Example

$$\Sigma = \{a, b, c\},$$

$$N = \{S, B, C, E, A\}$$

$$S \rightarrow (AE)\&(BC)$$

$$A \rightarrow aA|\epsilon$$

$$B \rightarrow aBb|\epsilon$$

$$C \rightarrow cC|\epsilon$$

$$E \rightarrow bEc|\epsilon$$

# Example revisited

## Example

$$\Sigma = \{a, b, c\},$$

$$N = \{S, B, C, E, A\}$$

$$S \rightarrow (AE) \& (BC)$$

$$L_S = (L_A \cdot L_E) \cap (L_B \cdot L_C)$$

$$A \rightarrow aA | \epsilon$$

$$L_A = \{a\} \cdot L_A \cup \{\epsilon\}$$

$$B \rightarrow aBb | \epsilon$$

$$L_B = \{a\} \cdot L_B \cdot \{b\} \cup \{\epsilon\}$$

$$C \rightarrow cC | \epsilon$$

$$L_C = \{c\} \cdot L_C \cup \{\epsilon\}$$

$$E \rightarrow bEc | \epsilon$$

$$L_E = \{b\} \cdot L_E \cdot \{c\} \cup \{\epsilon\}$$

# Example revisited

## Example

$$\Sigma = \{a, b, c\},$$

$$N = \{S, B, C, E, A\}$$

$S \rightarrow (AE) \& (BC)$	$L_S = (L_A \cdot L_E) \cap (L_B \cdot L_C)$	$\{a^n b^n c^n : n \in \mathbb{N}\}$
$A \rightarrow aA   \epsilon$	$L_A = \{a\} \cdot L_A \cup \{\epsilon\}$	$a^*$
$B \rightarrow aBb   \epsilon$	$L_B = \{a\} \cdot L_B \cdot \{b\} \cup \{\epsilon\}$	$\{a^n b^n : n \in \mathbb{N}\}$
$C \rightarrow cC   \epsilon$	$L_C = \{c\} \cdot L_C \cup \{\epsilon\}$	$c^*$
$E \rightarrow bEc   \epsilon$	$L_E = \{b\} \cdot L_E \cdot \{c\} \cup \{\epsilon\}$	$\{b^n c^n : n \in \mathbb{N}\}$

# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$ .

# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$
- Chomsky's normal form



# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$
- Chomsky's normal form
- Efficient parsing by CYK

# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$
- Chomsky's normal form
- Efficient parsing by CYK
- High expressive power

## Example

$$\{w c w : w \in \{a, b\}^*\}$$

# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$
- Chomsky's normal form
- Efficient parsing by CYK
- High expressive power

## Example

$$\{w c w : w \in \{a, b\}^*\}$$

## Negative results

# Basic results

## Positive results

- Resolved language equations with  $\cup$ ,  $\cap$  and  $\cdot$
- Chomsky's normal form
- Efficient parsing by CYK
- High expressive power

## Example

$$\{w cw : w \in \{a, b\}^*\}$$

## Negative results

- Mainly open questions

# Problem

## Problem

Do all conjunctive grammars over *unary* alphabet generate only *regular* languages?

# Problem

## Problem

*Do all conjunctive grammars over **unary** alphabet generate only **regular** languages? (This is true for CFG.)*

# Problem

## Problem

*Do all conjunctive grammars over **unary** alphabet generate only **regular** languages? (This is true for CFG.)*

## Conjecture

Yes

# Problem

## Problem

Do all conjunctive grammars over *unary* alphabet generate only *regular* languages? (This is true for CFG.)

## Conjecture

Yes

## Intuition

This *should be true* since regular sets are closed under

- *concatenation*



# Problem

## Problem

Do all conjunctive grammars over *unary* alphabet generate only *regular* languages? (This is true for CFG.)

## Conjecture

Yes

## Intuition

This *should be true* since regular sets are closed under

- *concatenation*
- *intersection*

# Problem

## Problem

Do all conjunctive grammars over *unary* alphabet generate only *regular* languages? (This is true for CFG.)

## Conjecture

Yes

## Intuition

This *should be true* since regular sets are closed under

- concatenation
- intersection
- union

# Result

## Theorem (Disproving the conjecture)

*Conjunctive grammars generate **non-regular languages** over unary alphabet.*

# Result

## Theorem (Disproving the conjecture)

Conjunctive grammars generate *non-regular languages* over unary alphabet.

$$\{a^{4^n} : n \in \mathbb{N}\}$$

# Result

## Theorem (Disproving the conjecture)

Conjunctive grammars generate *non-regular languages* over unary alphabet.

$$\{a^{4^n} : n \in \mathbb{N}\}$$

## Theorem (Extension)

For every regular language  $R \subseteq \{0, 1, \dots, k-1\}^*$  language

$$\{a^n : \exists w \in R \text{ } w \text{ read as a number is } n\}$$

is a unary conjunctive language.

# Result

## Theorem (Disproving the conjecture)

Conjunctive grammars generate *non-regular languages* over unary alphabet.

$$\{a^{4^n} : n \in \mathbb{N}\}$$

## Theorem (Extension)

For every regular language  $R \subseteq \{0, 1, \dots, k-1\}^*$  language

$$\{a^n : \exists w \in R \text{ w read as a number is } n\}$$

is a unary conjunctive language. *Positional notation.*

# Language

## Remark

We identify  $a^n$  with  $n$  and work with sets of **integers**.

# Language

## Remark

We identify  $a^n$  with  $n$  and work with sets of **integers**.

## Solution

$$L_1 = \{1 \cdot 4^n : n \in \mathbb{N}\},$$

$$L_2 = \{2 \cdot 4^n : n \in \mathbb{N}\},$$

$$L_3 = \{3 \cdot 4^n : n \in \mathbb{N}\},$$

$$L_{12} = \{6 \cdot 4^n : n \in \mathbb{N}\}.$$



# Language

## Remark

We identify  $a^n$  with  $n$  and work with sets of **integers**.

## Solution

$$\begin{aligned}L_1 &= \{1 \cdot 4^n : n \in \mathbb{N}\}, \\L_2 &= \{2 \cdot 4^n : n \in \mathbb{N}\}, \\L_3 &= \{3 \cdot 4^n : n \in \mathbb{N}\}, \\L_{12} &= \{6 \cdot 4^n : n \in \mathbb{N}\}.\end{aligned}$$

## Equations

$$\begin{aligned}B_1 &= (B_2 B_2 \cap B_1 B_3) \cup \{1\}, \\B_2 &= (B_{12} B_2 \cap B_1 B_1) \cup \{2\}, \\B_3 &= (B_{12} B_{12} \cap B_1 B_2) \cup \{3\}, \\B_{12} &= (B_3 B_3 \cap B_1 B_2).\end{aligned}$$

# Language

## Remark

We identify  $a^n$  with  $n$  and work with sets of **integers**.

## Solution

$$\begin{aligned}L_1 &= \{1 \cdot 4^n : n \in \mathbb{N}\}, \\L_2 &= \{2 \cdot 4^n : n \in \mathbb{N}\}, \\L_3 &= \{3 \cdot 4^n : n \in \mathbb{N}\}, \\L_{12} &= \{6 \cdot 4^n : n \in \mathbb{N}\}.\end{aligned}$$

## Equations

$$\begin{aligned}B_1 &= (B_2 B_2 \cap B_1 B_3) \cup \{1\}, \\B_2 &= (B_{12} B_2 \cap B_1 B_1) \cup \{2\}, \\B_3 &= (B_{12} B_{12} \cap B_1 B_2) \cup \{3\}, \\B_{12} &= (B_3 B_3 \cap B_1 B_2).\end{aligned}$$

This effectively manipulates the **positional notation**.

# What needs to be proved

- By general knowledge there is a **unique  $\epsilon$ -free solution**.

# What needs to be proved

- By general knowledge there is a **unique  $\epsilon$ -free solution**.
- Vector of sets  $(\dots, L_i, \dots)$  is  $\epsilon$ -free.

# What needs to be proved

- By general knowledge there is a **unique  $\epsilon$ -free solution**.
- Vector of sets  $(\dots, L_i, \dots)$  is  $\epsilon$ -free.
- We need to show that it is a solution.

# What needs to be proved

- By general knowledge there is a **unique  $\epsilon$ -free solution**.
- Vector of sets  $(\dots, L_i, \dots)$  is  $\epsilon$ -free.
- We need to show that it is a solution.

## Example

For example  $L_1$ , the rule is

$$B_1 = (B_2 B_2 \cap B_1 B_3) \cup \{1\}$$

So we want to prove that

$$L_1 = (L_2 L_2 \cap L_1 L_3) \cup \{1\}$$

## Details—what is in $B_2B_2$

Proof.

What are the possible non-zero **symbols** in  $B_2B_2$ ?

## Details—what is in $B_2B_2$

Proof.

What are the possible non-zero **symbols** in  $B_2B_2$ ?

$$\begin{array}{r} \phantom{+} \phantom{2} \phantom{0\dots0} \phantom{0} \phantom{0\dots0} \\ \phantom{+} \phantom{2} \phantom{0\dots0} 2 \phantom{0\dots0} 0 \phantom{0\dots0} \\ + \phantom{2} 2 \phantom{0\dots0} 0 \phantom{0\dots0} 0 \phantom{0\dots0} \\ \hline 2 \phantom{0\dots0} 0 \phantom{0\dots0} 2 \phantom{0\dots0} 0 \phantom{0\dots0} \end{array}$$



## Details—what is in $B_2B_2$

Proof.

What are the possible non-zero **symbols** in  $B_2B_2$ ?

$$\begin{array}{r} \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \hline + \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \hline 2 \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \end{array}$$

$$\begin{array}{r} \phantom{+} \phantom{2} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \\ \hline + \phantom{2} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \\ \hline 1 \phantom{0} \phantom{0\dots 0} \end{array}$$

## Details—what is in $B_2B_2$

Proof.

What are the possible non-zero **symbols** in  $B_2B_2$ ?

$$\begin{array}{r} \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ + \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ \hline \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \end{array}$$

$$\begin{array}{r} \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ + \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ \hline \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \\ \phantom{+} \phantom{2} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \phantom{2} \phantom{0\dots 0} \end{array}$$

Either only 1 or  $\{2, 2\}$ .



## Details—what is in $B_1 B_3$

Proof.

What are the possible non-zero **symbols** in  $B_1 B_3$ ?

## Details—what is in $B_1 B_3$

Proof.

What are the possible non-zero **symbols** in  $B_1 B_3$ ?

$$\begin{array}{r} \phantom{+} \phantom{1} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{1} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ + \phantom{1} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \hline 1 \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \end{array}$$







## Details—what is in $B_1 B_3$

### Proof.

What are the possible non-zero **symbols** in  $B_1 B_3$ ?

$$\begin{array}{r} \phantom{+} \phantom{1} \phantom{0\dots 0} \phantom{0} \phantom{0\dots 0} \\ \phantom{+} \phantom{1} \phantom{0\dots 0} 3 \phantom{0\dots 0} \\ + \phantom{1} \phantom{0\dots 0} 0 \phantom{0\dots 0} \\ \hline 1 \phantom{0\dots 0} 3 \phantom{0\dots 0} \end{array}$$

$$\begin{array}{r} \phantom{+} \phantom{3} \phantom{0\dots 0} \\ \phantom{+} 1 \phantom{0\dots 0} \\ + \phantom{3} \phantom{0\dots 0} \\ \hline 1 \phantom{0} \phantom{0\dots 0} \end{array}$$

Either only 1 or  $\{1, 3\}$ .

We compare this with 1 or  $\{2, 2\}$  from  $B_2 B_2$ .

The only possibility is only 1.





## First step: $ij0^*$

### Theorem

For every  $k$  and  $0 < i, j < k$  languages

$$\{a^n : \exists w \in i0^* \text{ } w \text{ read as a number is } n\}$$

$$\{a^n : \exists w \in ij0^* \text{ } w \text{ read as a number is } n\}$$

are unary conjunctive languages.

### Idea

Done in the **same way**.

## First step: $ij0^*$

### Theorem

For every  $k$  and  $0 < i, j < k$  languages

$$\{a^n : \exists w \in i0^* \text{ } w \text{ read as a number is } n\}$$

$$\{a^n : \exists w \in ij0^* \text{ } w \text{ read as a number is } n\}$$

are unary conjunctive languages.

### Idea

Done in the **same way**.

- Nonterminal  $B_{i,j}$  for each language.

## First step: $ij0^*$

### Theorem

For every  $k$  and  $0 < i, j < k$  languages

$$\{a^n : \exists w \in i0^* \text{ } w \text{ read as a number is } n\}$$

$$\{a^n : \exists w \in ij0^* \text{ } w \text{ read as a number is } n\}$$

are unary conjunctive languages.

### Idea

Done in the *same way*.

- Nonterminal  $B_{i,j}$  for each language.
- We focus on the *leading symbols*—the only non-zero symbols in  $ij0^*$ , that is  $i$  and  $j$ .

## First step: $ij0^*$

### Theorem

For every  $k$  and  $0 < i, j < k$  languages

$$\{a^n : \exists w \in i0^* \text{ } w \text{ read as a number is } n\}$$

$$\{a^n : \exists w \in ij0^* \text{ } w \text{ read as a number is } n\}$$

are unary conjunctive languages.

### Idea

Done in the *same way*.

- Nonterminal  $B_{i,j}$  for each language.
- We focus on the *leading symbols*—the only non-zero symbols in  $ij0^*$ , that is  $i$  and  $j$ .
- Intersections of concatenations *filter out* wrong combination of leading symbols.

## Second step: any regular language

### Theorem

For every  $k$  and  $R \subset \{0, \dots, k-1\}^*$

$$\{a^n : \exists w \in R \text{ } w \text{ read as a number is } n\}$$

is a unary conjunctive language.

## Second step: any regular language

### Theorem

For every  $k$  and  $R \subset \{0, \dots, k-1\}^*$

$$\{a^n : \exists w \in R \text{ w read as a number is } n\}$$

is a unary conjunctive language.

### Idea

Let  $\langle \{0, \dots, k-1\}, Q, q_0, F, \delta \rangle$  recognizes  $R$ .

## Second step: any regular language

### Theorem

For every  $k$  and  $R \subset \{0, \dots, k-1\}^*$

$$\{a^n : \exists w \in R \text{ w read as a number is } n\}$$

is a unary conjunctive language.

### Idea

Let  $\langle \{0, \dots, k-1\}, Q, q_0, F, \delta \rangle$  recognizes  $R$ .

We introduce nonterminal  $B_{i,j,q}$  for language

$$\{ijw : \delta(q_0, w) = q\}$$

## Second step: any regular language

### Theorem

For every  $k$  and  $R \subset \{0, \dots, k-1\}^*$

$$\{a^n : \exists w \in R \text{ w read as a number is } n\}$$

is a unary conjunctive language.

### Idea

Let  $\langle \{0, \dots, k-1\}, Q, q_0, F, \delta \rangle$  recognizes  $R$ .

We introduce nonterminal  $B_{i,j,q}$  for language

$$\{ijw : \delta(q_0, w) = q\}$$

Information the indices carry:

- leading symbol  $i$
- second leading symbol  $j$
- $q$ —the computation of  $M$  on the rest of the word



# Productions for $B_{i,j,q}$

## Example

$$B_{i,j,q} \rightarrow \left( \&_{n=1}^4 B_{i-1,j+n} B_{k-n,x,q'} \right)$$

where  $x, q'$  such that  $q \in \delta(q', x)$

# Productions for $B_{i,j,q}$

## Example

$$B_{i,j,q} \rightarrow \left( \&_{n=1}^4 B_{i-1,j+n} B_{k-n,x,q'} \right)$$

where  $x, q'$  such that  $q \in \delta(q', x)$

$$\begin{array}{rcccc}
 & & & & \text{state } q' \\
 & & & k-n & x \underbrace{\dots} \\
 + & i-1 & j+n & 00 & \dots 0 \\
 \hline
 & i & j & \underbrace{x \dots}_{\text{state } q} & 
 \end{array}$$

# The result

## Definition

Conjunctive grammar is a CFG **extended by intersection** in the body of the rules.

# The result

## Definition

Conjunctive grammar is a CFG **extended by intersection** in the body of the rules.

## Theorem

*For every regular language  $R \subseteq \{0, 1, \dots, k - 1\}$  language*

$$\{a^n : \exists w \in R \text{ wread as a number is } n\}$$

*is a unary conjunctive language.*

# The result

## Definition

Conjunctive grammar is a CFG **extended by intersection** in the body of the rules.

## Theorem

*For every regular language  $R \subseteq \{0, 1, \dots, k - 1\}$  language*

$$\{a^n : \exists w \in R \text{ wread as a number is } n\}$$

*is a unary conjunctive language.*

In particular it generates non-regular languages.

# The result

## Definition

Conjunctive grammar is a CFG **extended by intersection** in the body of the rules.

## Theorem

*For every regular language  $R \subseteq \{0, 1, \dots, k - 1\}$  language*

$$\{a^n : \exists w \in R \text{ wread as a number is } n\}$$

*is a unary conjunctive language.*

In particular it generates non-regular languages.

We effectively manipulate **positional notation**.

## Related topics and following work

- **Unambiguity** of the language. The construction for  $R = ij0^*$  can be made unambiguous. What happens in general?

## Related topics and following work

- **Unambiguity** of the language. The construction for  $R = ij0^*$  can be made unambiguous. What happens in general?
- The result can be extended to a **larger** class of languages [A. Jez, A. Okhotin, CSR 2007].



## Related topics and following work

- **Unambiguity** of the language. The construction for  $R = ij0^*$  can be made unambiguous. What happens in general?
- The result can be extended to a **larger** class of languages [A. Jez, A. Okhotin, CSR 2007].
- Instead of grammars we can focus on sets of integers. Equations on sets of integers using  $\cap$ ,  $\cup$  and  $+$  defined as

$$A + B = \{a + b : a \in A, b \in B\}.$$

[A. Jez, A. Okhotin, TALE 2007].

## Open questions

- **General properties** of conjunctive grammars

## Open questions

- **General properties** of conjunctive grammars
  - ▶ closure under **complementation**

## Open questions

- **General properties** of conjunctive grammars
  - ▶ closure under **complementation**
  - ▶ better **recognition** (space/time)

## Open questions

- **General properties** of conjunctive grammars
  - ▶ closure under **complementation**
  - ▶ better **recognition** (space/time)
  - ▶ inherent **ambiguity**

## Open questions

- **General properties** of conjunctive grammars
  - ▶ closure under **complementation**
  - ▶ better **recognition** (space/time)
  - ▶ inherent **ambiguity**
- **Unambiguity** of the constructed unary languages

## Open questions

- **General properties** of conjunctive grammars
  - ▶ closure under **complementation**
  - ▶ better **recognition** (space/time)
  - ▶ inherent **ambiguity**
- **Unambiguity** of the constructed unary languages
- Closure under **complementation** in the unary case.