Beyond context-free languages

Ashley Montanaro
ashley@cs.bris.ac.uk

Department of Computer Science, University of Bristol
Bristol, UK

21 March 2014

Introduction

It turns out that not all languages are context-free.

Introduction

It turns out that not all languages are context-free.

To prove that a language is not context-free, a tool which can be used is the pumping lemma for context-free languages.

Lemma

If $\mathcal{L}$ is a CFL, there exists an integer p (the pumping length) such that any string $s \in \mathcal{L}$ such that $|s| \geq p$ can be written as

$$s = uvxyz$$

where:

1. For all $i \geq 0$, $uv^ixy^iz \in \mathcal{L}$,
2. $|vy| > 0$,
3. $|vxy| \leq p$.

Pumping lemmas

Any string $s \in \mathcal{L}$ such that $|s| \geq p$ can be written as...

If $\mathcal{L}$ is context-free

$s = uvxyz$, where:

1. For all $i \geq 0$, $uv^i xy^i z \in \mathcal{L}$,
2. $|vy| > 0$,
3. $|vxy| \leq p$.

If $\mathcal{L}$ is regular

$s = xyz$, where:

1. For all $i \geq 0$, $xy^i z \in \mathcal{L}$,
2. $|y| > 0$,
3. $|xy| \leq p$.

Example

We show that $\mathcal{L} = \{\mathrm{a}^n \mathrm{b}^n \mathrm{c}^n \mid n \geq 0\}$ is not context-free.

Example

We show that $\mathcal{L} = \{\mathrm{a}^n\mathrm{b}^n\mathrm{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free.

Example

We show that $\mathcal{L} = \{\mathrm{a}^n\mathrm{b}^n\mathrm{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free. Let p be the pumping length and consider $s = \mathrm{a}^p\mathrm{b}^p\mathrm{c}^p \in \mathcal{L}$.

Example

We show that $\mathcal{L} = \{\mathrm{a}^n\mathrm{b}^n\mathrm{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free. Let p be the pumping length and consider $s = \mathrm{a}^p\mathrm{b}^p\mathrm{c}^p \in \mathcal{L}$.

- We can write $s = uvxyz$, and for all $i \geq 0$, $uv^ixy^iz \in \mathcal{L}$.

Example

We show that $\mathcal{L} = \{\mathtt{a}^n\mathtt{b}^n\mathtt{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free. Let p be the pumping length and consider $s = \mathtt{a}^p\mathtt{b}^p\mathtt{c}^p \in \mathcal{L}$.
- We can write $s = uvxyz$, and for all $i \geq 0$, $uv^ixy^iz \in \mathcal{L}$.
- We split into two cases:
  - If each of v and y contains only one kind of symbol, uv^2xy^2z cannot have equal numbers of $\mathtt{a}$'s, $\mathtt{b}$'s and $\mathtt{c}$'s; contradiction.

Example

We show that $\mathcal{L} = \{\mathtt{a}^n\mathtt{b}^n\mathtt{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free. Let p be the pumping length and consider $s = \mathtt{a}^p\mathtt{b}^p\mathtt{c}^p \in \mathcal{L}$.
- We can write $s = uvxyz$, and for all $i \geq 0$, $uv^ixy^iz \in \mathcal{L}$.
- We split into two cases:
  - If each of v and y contains only one kind of symbol, uv^2xy^2z cannot have equal numbers of a's, b's and c's; contradiction.
  - If one of v and y contains two types of symbol, uv^2xy^2z must have some symbols in the wrong order; contradiction.

Example

We show that $\mathcal{L} = \{\mathtt{a}^n\mathtt{b}^n\mathtt{c}^n \mid n \geq 0\}$ is not context-free.

- Towards a contradiction, assume that $\mathcal{L}$ is context-free. Let p be the pumping length and consider $s = \mathtt{a}^p\mathtt{b}^p\mathtt{c}^p \in \mathcal{L}$.
- We can write $s = uvxyz$, and for all $i \geq 0$, $uv^ixy^iz \in \mathcal{L}$.
- We split into two cases:
  - If each of v and y contains only one kind of symbol, uv^2xy^2z cannot have equal numbers of a's, b's and c's; contradiction.
  - If one of v and y contains two types of symbol, uv^2xy^2z must have some symbols in the wrong order; contradiction.

Another natural example of a non-context-free language is

$$\mathcal{L} = \{w\#w \mid w \in \{0, 1\}^*\}.$$

Summary and further reading

- The pumping lemma for context-free languages can be used to show that a language is not context-free.
- Just as with the pumping lemma for regular languages, applying the lemma can require some ingenuity…
- Further reading: Sipser §2.3.