

# Deflating

A mathematical diversion

Karst Koymans / Jeroen van der Ham

Informatics Institute  
University of Amsterdam

Tuesday, September 11, 2012

- 1 Definitions and conventions
- 2 Text files and their representation
- 3 Initiators
- 4 Terminators
- 5 Separators or delimiters

# Outline

- 1 Definitions and conventions
- 2 Text files and their representation
- 3 Initiators
- 4 Terminators
- 5 Separators or delimiters

# Definitions

- Deflating is a synonym for
  - Serialization
  - Marshalling
- The opposite notion is inflating
  - Deserialization
  - Unmarshalling

# Definitions

- $A$  is a given set (alphabet)
- $A^*$  is the set of finite sequences of elements in  $A$
- $\langle \rangle$ , the empty sequence, is in  $A^*$
- If  $\alpha$  is in  $A^*$  and  $a$  is in  $A$ , then  $\langle \alpha, a \rangle$  is in  $A^*$
- $a, b, c, \dots \in A$
- $\alpha, \beta, \gamma, \dots \in A^*$
- $\rho, \sigma, \tau, \dots \in A^{**}$

# Conventions

- We write  $\langle a \rangle$  in stead of  $\langle \langle \rangle, a \rangle$
- We write  $\langle a, b \rangle$  in stead of  $\langle \langle a \rangle, b \rangle$
- We write  $\langle a, b, c \rangle$  in stead of  $\langle \langle a, b \rangle, c \rangle$
- In general we write  $\langle a_1, a_2, a_3, \dots, a_n, a_{n+1} \rangle$  in stead of  
 $\langle \langle a_1, a_2, a_3, \dots, a_n \rangle, a_{n+1} \rangle$  or  
 $\langle \dots \langle \langle \langle \rangle, a_1 \rangle, a_2 \rangle, a_3 \rangle, \dots, a_n \rangle, a_{n+1} \rangle$

# Concatenation

- $* : A^* \times A^* \rightarrow A^*$ 
  - $\alpha * \langle \rangle = \alpha$
  - $\alpha * \langle \beta, b \rangle = \langle \alpha * \beta, b \rangle$

# Outline

- 1 Definitions and conventions
- 2 Text files and their representation**
- 3 Initiators
- 4 Terminators
- 5 Separators or delimiters



# Representation of text files

- Text files are two-dimensional, hence elements in  $A^{**}$
- Text file with lines 'abc', 'd', '', 'ef' is represented as  
 $\langle\langle a, b, c \rangle, \langle d \rangle, \langle \rangle, \langle e, f \rangle\rangle$

# Adding a specific line marker symbol

- We use  $\wr$  as a line marker symbol
- We extend our alphabet  $A$  with this extra symbol  $\wr \notin A$
- We define  $B = A \cup \{ \wr \}$

# Outline

- 1 Definitions and conventions
- 2 Text files and their representation
- 3 Initiators**
- 4 Terminators
- 5 Separators or delimiters

# Deflating with an initial line marker

- Define a mapping  $f_{\text{ini}} : A^{**} \rightarrow B^*$  by
  - $f_{\text{ini}}(\langle \rangle) = \langle \rangle$
  - $f_{\text{ini}}(\langle \rho, \alpha \rangle) = f_{\text{ini}}(\rho) * \langle \wr \rangle * \alpha$
- In  $f_{\text{ini}}$   $\wr$  is used as an initiator of lines
- $f_{\text{ini}}$  is injective, but not surjective (why?)

# Examples of deflating with initial line marker

$$f_{\text{ini}}(\langle\langle a, b, c \rangle, \langle d \rangle, \langle \rangle, \langle e, f \rangle\rangle) = \\ \langle \varrho, a, b, c, \varrho, d, \varrho, \varrho, e, f \rangle$$

$$f_{\text{ini}}(\langle \rangle) = \langle \rangle$$

$$f_{\text{ini}}(\langle\langle \rangle\rangle) = f_{\text{ini}}(\langle\langle \rangle, \langle \rangle\rangle) = f_{\text{ini}}(\langle \rangle) * \langle \varrho \rangle * \langle \rangle = \langle \varrho \rangle$$

# Outline

- 1 Definitions and conventions
- 2 Text files and their representation
- 3 Initiators
- 4 Terminators**
- 5 Separators or delimiters

# Deflating with a terminal line marker

- Define a mapping  $f_{\text{term}} : A^{**} \rightarrow B^*$  by
  - $f_{\text{term}}(\langle \rangle) = \langle \rangle$
  - $f_{\text{term}}(\langle \rho, \alpha \rangle) = f_{\text{term}}(\rho) * \alpha * \langle \mathfrak{d} \rangle$
- In  $f_{\text{term}}$   $\mathfrak{d}$  is used as a terminator of lines
- $f_{\text{term}}$  is injective, but not surjective (why?)

# Examples of deflating with terminator

$$f_{\text{term}}(\langle\langle a, b, c \rangle, \langle d \rangle, \langle \rangle, \langle e, f \rangle\rangle) = \\ \langle a, b, c, \varnothing, d, \varnothing, \varnothing, e, f, \varnothing \rangle$$

$$f_{\text{term}}(\langle \rangle) = \langle \rangle$$

$$f_{\text{term}}(\langle\langle \rangle\rangle) = f_{\text{term}}(\langle\langle \rangle, \langle \rangle\rangle) = \\ f_{\text{term}}(\langle \rangle) * \langle \rangle * \langle \varnothing \rangle = \langle \varnothing \rangle$$



# Outline

- 1 Definitions and conventions
- 2 Text files and their representation
- 3 Initiators
- 4 Terminators
- 5 Separators or delimiters**

# Deflating with a separator or delimiter

- Let  $A^{*+} = A^{**} - \{\langle \rangle\}$
- Define a mapping  $f_{\text{sep}} : A^{*+} \rightarrow B^*$  by
  - $f_{\text{sep}}(\langle \alpha \rangle) = \alpha$
  - $f_{\text{sep}}(\langle \rho, \beta \rangle) = f_{\text{sep}}(\rho) * \langle \wr \rangle * \beta$ , where  $\rho \neq \langle \rangle$
- In  $f_{\text{sep}}$   $\wr$  is used inbetween lines
- $f_{\text{sep}}$  is bijective (why?)
- $\#markers = \#lines - 1$

# Examples of deflating with separator

$$f_{\text{sep}}(\langle\langle a, b, c \rangle, \langle d \rangle, \langle \rangle, \langle e, f \rangle\rangle) = \\ \langle a, b, c, \varrho, d, \varrho, \varrho, e, f \rangle$$

$f_{\text{sep}}(\langle \rangle)$  is not defined

$$f_{\text{sep}}(\langle\langle \rangle\rangle) = \langle \rangle$$

$$f_{\text{sep}}(\langle \alpha, \beta \rangle) =$$

$$f_{\text{sep}}(\langle\langle \langle \rangle, \alpha \rangle, \beta \rangle) =$$

$$f_{\text{sep}}(\langle\langle \langle \rangle, \alpha \rangle) * \langle \varrho \rangle * \beta =$$

$$f_{\text{sep}}(\langle \alpha \rangle) * \langle \varrho \rangle * \beta =$$

$$\alpha * \langle \varrho \rangle * \beta$$