Deflating

A mathematical diversion

Karst Koymans / Jeroen van der Ham

Informatics Institute
University of Amsterdam

Tuesday, September 11, 2012

Table of Contents

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
- $<>$, the empty sequence, is in A^*
- If α is in A^* and a is in A, then $\langle \alpha, a \rangle$ is in A^*
- $a, b, c, \ldots \in A$
- $\alpha, \beta, \gamma, \ldots \in A^*$
- $\rho, \sigma, \tau, \ldots \in A^{**}$

Conventions

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

Concatenation

- $* : A^* \times A^* \to A^*$
  - $\alpha * <> = \alpha$
  - $\alpha * \langle \beta, b \rangle = \langle \alpha * \beta, b \rangle$

Representation of text files

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

Adding a specific line marker symbol

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

Deflating with an initial line marker

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

Examples of deflating with initial line marker

$f_{\text{ini}}(<< a, b, c >, < d >, <>, < e, f >>) =$
$< \lrcorner, a, b, c, \lrcorner, d, \lrcorner, \lrcorner, e, f >$

$f_{\text{ini}}(<>) = <>$

$f_{\text{ini}}(<<>>) = f_{\text{ini}}(\langle <>, <> \rangle) = f_{\text{ini}}(<>) * < \lrcorner > * <> =$
$< \lrcorner >$

Deflating with a terminal line marker

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

Examples of deflating with terminator

$f_{\text{term}}(<< a, b, c >, < d >, <>, < e, f >>) =$
$< a, b, c, \lrcorner, d, \lrcorner, \lrcorner, e, f, \lrcorner >$

$f_{\text{term}}(<>) = <>$

$f_{\text{term}}(<<>>) = f_{\text{term}}(\langle <>, <> \rangle) =$
$f_{\text{term}}(<>) * <> * < \lrcorner > = < \lrcorner >$

Deflating with a separator or delimiter

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

Examples of deflating with separator

$f_{sep}(<< a, b, c >, < d >, <>, < e, f >>) =$
$\quad < a, b, c, \wr, d, \wr, \wr, e, f >$

$f_{sep}(<>)$ is not defined

$f_{sep}(<<>>) = <>$

$f_{sep}(< \alpha, \beta >) =$
$\quad f_{sep}(\langle \langle <>, \alpha \rangle, \beta \rangle) =$
$\quad f_{sep}(\langle <>, \alpha \rangle) * < \wr > * \beta =$
$\quad f_{sep}(< \alpha >) * < \wr > * \beta =$
$\quad \alpha * < \wr > * \beta$