# Today's Topic

• Pretty Printing

Like parsing a typical demo-application

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# What's it all about?

A pretty printer is...

• a tool (often a library of routines) designed for converting a *tree* into plain *text* 

Essential goals...

• a minimum number of lines while preserving and reflecting the structure of the tree by indentation

# **Pretty Printing**

#### **Pretty Printing**

...like lexical and syntactical analysis another typical application for demonstrating the elegance of functional programming.

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

2

# "Good" Pretty-Printer

...distinguished by properly balancing

- Simplicity of usage
- Flexibility of the format
- "Prettiness" of output

### Reference

The following presentation is based on...

• Philip Wadler. A Prettier Printer. In Jeremy Gibbons, Oege de Moor (Eds.), The Fun of Programming. Palgrave MacMillan, 2003.

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# Why "prettier" than "pretty"?

Wadler considers his "Prettier Printer" an improvement of the pretty printer library proposed by John Hughes, which is widely recognized as a standard.

• The design of a pretty-printer library. In Johan Jeuring, Erik Meijers (Hrsg.), Advanced Functional Programming, LNCS 925, Springer, 1995.

Hughes' library enjoys the following characteristics:

- Two ways to concatenate documents (horizontal and vertical), one of which
  - vertical: without unit
  - horizontal: with right-unit (only)
- ca. 40% more code, ca. 40% slower as Wadler's proposal

# **Distinguishing Feature**

...of the "Prettier Printer" proposed by Philip Wadler:

- There is only a single way to concatenate documents, which is
  - associative
  - with left-unit and right-unit

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

6

# A Simple Pretty Printer: The Basis

Characteristic: For each document there is only one possible layout (e.g., no attempt is made to compress structure onto a single line).

The basic operators needed are:

#### Convention:

• Arguments of text are free of *newline* characters

## **A Simple Implementation**

Implement...

• doc as strings (i.e. as data type String)

with...

- (<>) ...concatenation of strings
- nil ...empty string
- text ...identity on strings
- line ...new line
- ullet nest i ...indentation: adding i spaces (after each line break by means of line)  $\leadsto$  essential difference to Hughes' pretty printer allowing to drop one concatenation operator
- layout ...identity on strings

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

\_

11

# And its desired output

A text, where indentation reflects the tree structure...

# Example

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

10

# **Implementation**

The below implementation achieves this...

# Another possibly wanted output of B

```
aaa[
bbbbbb[
ccc,
dd
],
eee,
fffff[
gg,
hhh,
ii
]
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

#### 13

15

# An implementation producing the latter output

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

14

# **A** Normal Form of Documents

Normal form...

• text alternating with line breaks nested to a given indentation

#### Note:

• Documents can always be reduced to normal form

# Normal Forms: An Example 1(3)

The document...

```
text "bbbbb" <> text "[" <>
nest 2 (
          line <> text "ccc" <> text "," <>
          line <> text "dd"
) <>
line <> text "]"
```

# Normal Forms: An Example 2(3)

...has the normal form:

```
text "bbbbb[" <>
nest 2 line <> text "ccc," <>
nest 2 line <> text "dd" <>
nest 0 line <> text "]"
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# Why does it work

...because of the properties (laws) the functions enjoy.

More on this next...

# Normal Forms: An Example 3(3)

```
bbbbb[
ccc,
dd
```

...and prints as follows:

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

18

# Properties of the Functions – Laws 1(2)

We have:

17

19

```
text (s ++ t) = text s <> text t
                                          (text is homomorphism from
text ""
                                           string concatenation to
                                           document concatenation)
nest (i+j) x
                 = nest i (nest j x)
                                          (nest is homomorphism from
nest 0 x
                 = x
                                           addition to composition)
nest i (x \leftrightarrow y) = nest i x \leftrightarrow nest i y (nest distributes through
nest i nil
                 = nil
                                           document concatenation)
nest i (text s) = text s
                                   (Nesting is absorbed by text)
```

# Properties of the Functions – Laws 2(2)

#### Meaning

• The above laws are sufficient to establish that documents can always be transformed into normal form (first four laws: application left to right; last three laws: application right to left)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

21

# The Implementation of Doc

#### Intuition

...represent documents as a concatenation of items, where each item is a text or a line break indented to a given amount.

...realized as a sum type (the algebra of documents):

 $\ldots$  with the following relationships of the constructors to document operators:

### Further Properties – Laws

...on the relationship of documents and their layouts

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

22

# Example

The normal form (considered already previously)...

```
text "bbbbb[" <>
nest 2 line <> text "ccc," <>
nest 2 line <> text "dd" <>
nest 0 line <> text "]"

...has the representation:
   "bbbbb[" 'Text' (
   2 'Line' ("ccc," 'Text' (
   2 'Line' ("dd," 'Text' (
   0 'Line' ("]," 'Text' Nil)))))
```

# **Derived Implementations 1(2)**

...of the document operators from the above equations:

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

## On the Correctness

...of the derived implementations:

• Remaining equations: Similar reasoning

# **Derived Implementations 2(2)**

```
nest i (s 'Text' x) = s 'Text' nest i x
nest i (j 'Line' x) = (i+j) 'Line' nest i x
nest i Nil = Nil
layout (s 'Text' x) = s ++ layout x
layout (i 'Line' x) = '\n' : copy i ' ' ++ layout x
layout Nil = ""
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

26

# **Documents with Multiple Layouts**

- *Up to now...* documents are equivalent to a string (i.e., have a fixed single layout)
- *Next...* documents are equivalent to a set of strings (i.e., may have multiple layouts)

where each string correponds to a layout.

All what is needed to render this possible: A new function

```
group :: Doc -> Doc
```

#### Informally:

group returns the set with one new element added, which represents the layout in which everything is compressed on one line, when applied to a document representing a set of layouts.

25

## **Preferred Layouts**

Technically, this also requires...

• layout is replaced by pretty

```
pretty :: Int -> Doc -> String
```

• pretty's integer-argument specifies the preferred maximum line length of the output (and hence the prettiest layout out of the set of alternatives at hand)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

29

# Implementation of the new Functions

The following supporting functions are required:

```
-- Forming the union of two sets of layouts
(<|>) :: Doc -> Doc -> Doc
-- Replacement of each line break (including subsequent
-- indentation) by a single space
flatten :: Doc -> Doc
```

- Observation ...a document always represents a non-empty set of layouts
- Requirements
  - ...in (x <| y) all layouts of x and y enjoy the same flat layout (mandatory invariant of <| y)
  - ...each first line in x is at least as long as each first line in y (second invariant)

# Example

```
Using...
```

#### This ensures

- Output in one line where possible (i.e. length  $\leq$  30)
- Insertion of sufficiently many line breaks in order to avoid exceeding the given maximum line length

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

30

# Properties (Laws) of (<|>)

 $\ldots$ operators on simple documents are extended pointwise through union:

```
(x <|> y) <> z = (x <> z) <|> (y <> z)

x <> (y <|> z) = (x <> y) <|> (x <> z)

nest i (x <|> y) = nest i x <|> nest i y
```

# Properties (Laws) of flatten

...the interaction of flatten with other document operators:

```
flatten (x <|> y) = flatten x

flatten (x <> y) = flatten x <> flatten y

flatten nil = nil

flatten (text s) = text s

flatten line = text " " -- the most interesting case

flatten (nest i x) = flatten x
```

33

35

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

### **Normal Form**

Based on the previous laws each document can be reduced to a *normal form* of the form

where each xi is in the normal form of simple documents (which was introduced previously).

# Implementation of group

...by means of flatten and (<>), the implementation of group can be given:

group 
$$x = flatten x < |> x$$

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

21

# Selecting a "best" Layout out of a Set of Layouts

...by defining an ordering relation on lines in dependence of the given maximum line length

Out of two lines...

- which do not exceed the maximum length, select the longer one
- of which at least one exceeds the maximum length, select the shorter one

*Note*: Sometimes we have to pick a layout where some line exceeds the limit. However, this is done only, if this is unavoidable.

## The Adapted Implementation of Doc

The new implementation of Doc. Quite similar to the original one...

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

#### 00)

37

# Example 1(8)

```
The document...
```

```
group(
    group(
        group( text "hello" <> line <> text "a")
        <> line <> text "b")
        <> line <> text "c")
<> line <> text "d")
```

# Relationship of Constructors and Document Operators

The following relationships hold...

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

38

# Example 2(8)

...has the following possible layouts:

# Example 3(8)

Task: ...print the above document under the constraint that the maximum line length is 5

ightharpoonup the right-most layout of the previous slide is requested

Initial considerations:

- ...Factoring out "hello" of all the layouts in x and y

  "hello" 'Text' ((" " 'Text' x) 'Union' (0 'Line' y))
- ...Defining additionally the interplay of (<>) and nest with Union

```
(x 'Union' y) \Leftrightarrow z = (x \Leftrightarrow z) 'Union' (y \Leftrightarrow z)
nest k (x 'Union' y) = nest k x 'Union' nest k y
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

41

# Example 5(8)

Considerations on correctness...

Derivation of group (i 'Line' x) (see line two) (preserving the invariant required by union)

```
group (i 'Line' x)
= { Definition of Line }
group (nest i line <> x)
= { Definition of group}
flatten (nest i line <> x) <|> (nest i line s <> x)
= { Definition of flatten }
  (text " " <> flatten x) <|> (nest i line <> x)
= { Definition of Text, Union, Line }
  (" " 'Text' flatten x) 'Union' (i 'Line' x)
```

# Example 4(8)

Implementations of group and flatten can easily be derived...

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

42

# Example 6(8)

Correctness considerations...

Derivation of group (s 'Text' x) (see line three)

```
group (s 'Text' x)
= { Definition Text }
group (text s <> x)
= { Definition group}
flatten (text s <> x) <|> (text s <> x)
= { Definition flatten }
  (text s <> flatten x) <|> (text s <> x)
= { <> distributes through <|> }
  text s <> (flatten x <|> x)
= { Definition group }
  text s <> group x
= { Definition Text }
  s 'Text' group x
```

# Example 7(8)

Selecting the "best" layout...

#### Remark:

- best ...converts a "union"-afflicted document into a "union"-free document
- Argument w ...maximum line length
- $\bullet$  Argument  ${\bf k}$  ...already consumed letters (including indentation) on current line

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

45

47

### A more efficient variant

...by means of a new implementation of documents

#### Remark:

• In distinction to the previous document type we here use capital letters in order to avoid name clashes with the previous definitions

# Example 8(8)

Check, if the first document line stays within the maximum line length...

```
fits w x | w<0 = False
fits w Nil = True
fits w (s 'Text' x) = fits (w - length s) x
fits w (i 'Line' x) = True</pre>
```

Last but not least, the output routine (layout remains unchanged): Select the best layout and convert it to a string...

```
pretty w x = layout (best w 0 x)
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

46

# **Implementing the Document Operators**

Defining the operators to build a document: Straightforward...

# Implementing group and flatten

As before, we require the following invariants:

- ullet ...in (x :<|> y) all layouts in x and y flatten to the same layout
- ...no first line in x is shorter than any first line in y

Definitions of group and flatten are then straightforward:

```
group x = flatten x :<|> x

flatten NIL = NIL
flatten (x :<> y) = flatten x:<> flatten y
flatten (NEST i x) = NEST i (flatten x)
flatten (TEXT s) = TEXT s
flatten LINE = TEXT " "
flatten (x :<|> y) = flatten x
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

40

## Selecting the "best" Layout

Generalizing the function "best" by composing the old function with the representation function...

```
be w k z = best w k (rep z) (Hypothesis)

best w k x = be w k [(0,x)]
```

where the definition is derived from the old one...

# **Representation Function**

...generating the document from an indentation-afflicted document

```
rep z = fold (<>) nil [nest i x | (i,x) <- z ]
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

50

# In Preparation of further Applications 1(3)

...first some useful convenience functions:

```
x <+> y
x </> y
= x <> text " " <> y
= x <> line <> y

folddoc f []
folddoc f [x]
folddoc f (x:xs)
= f x (folddoc f xs)

spread
stack
= folddoc (<+>)
stack
```

# In Preparation of further Applications 2(3)

...some additional auxiliary functions:

53

55

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# In Preparation of further Applications 3(3)

```
fill, a variant of fillwords \rightarrow ...collapses a list of documents to a single document fill [] = nil fill [x] = x fill (x:y:zs) = (flatten x <+> fill (flatten y : zs)) :<|> (x </> fill (y : zs)
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

54

# Application 1(2)

Printing XML-documents (simplified syntax)...

# Application 2(2)

Continuation...

```
quoted s = "\"" ++ s ++ "\""
showTag n a = text n <> showFill showAtts a
showFill f [] = nil
showFill f xs = bracket "" (fill (concat (map f xs))) ""
```

# XML Example 1

...for a given maximum line length of 30 letters:

```
  Here is some
  <em> emphasized </em> text.
  Here is a
    <a
      href="http://www.eg.com/"
      > link </a>
      elsewhere.
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

#### 57

# XML Example 3:

```
...after dropping of flatten in fill:
```

```
  Here is some <em>
     emphasized
  </em> text. Here is a <a
     href="http://www.eg.com/"
  > link </a> elsewhere.
```

...start and close tags are crammed together with other text  $\sim$  less beautifully than before.

# XML Example 2

...for a given maximum line length of 60 letters:

```
  Here is some <em> emphasized </em> text. Here is a
  <a href="http://www.eg.com/" > link </a> elsewhere.
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

58

# Overview of the Code 1(11)

Source: Philip Wadler. A Prettier Printer. In Jeremy Gibbons, Oege de Moor (Eds.), The Fun of Programming. Palgrave MacMillan, 2003.

### Overview of the Code 2(11)

```
= NIL
nil
                          = x :<> y
x <> y
                          = NEST i x
nest i x
text s
                          = TEXT s
line
                          = I.TNF.
                          = flatten x : < | > x
group x
flatten NIL
                          = NIL
flatten (x :<> y)
                          = flatten x:<> flatten y
flatten (NEST i x)
                          = NEST i (flatten x)
flatten (TEXT s)
                          = TEXT s
                          = TEXT " "
flatten LINE
flatten (x : < | > y)
                          = flatten x
```

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# Overview of the Code 4(11)

```
fits w \times | w < 0
                          = False
fits w Nil
                          = True
fits w (s 'Text' x)
                          = fits (w - length s) x
fits w (i 'Line' x)
                          = True
pretty w x
                          = layout (best w 0 x)
-- Utility functions
                          = x <> text " " <> y
x <+> y
x </> y
                          = x <> line <> y
folddoc f ∏
                          = nil
folddoc f [x]
                          = x
folddoc f (x:xs)
                          = f x (folddoc f xs)
```

# Overview of the Code 3(11)

```
lavout Nil
layout (s 'Text' x)
                          = s ++ layout x
                          = '\n': copy i ' ' ++ layout x
layout (i 'Line' x)
                          = [x | <- [1..i]]
copy i x
best w k x
                          = be w k \lceil (0,x) \rceil
be w k []
                          = Nil
be w k ((i.NIL):z)
                          = be w k z
be w k ((i,x : <> y) : z) = be w k ((i,x) : (i,y) : z)
be w k ((i, NEST j x) : z) = be w k ((i+j),x) : z)
be w k ((i,TEXT s) : z) = s 'Text' be w (k+length s) z
be w k ((i,LINE) : z) = i 'Line' be w i z
be w k ((i.x : \langle | \rangle y) : z) = better w k (be w k ((i.x) : z))
                                          (be w k (i,v) : z))
better w k x y
                       = if fits (w-k) x then x else y
```

# Overview of the Code 5(11)

```
spread
                   = folddoc (<+>)
stack
                   = folddoc (</>)
bracket 1 x r
                   = group (text 1 <>
                             nest 2 (line \langle \rangle x) \langle \rangle
                             line <> text r)
x <+/> y
                   = x <> (text " " :<|> line) <> y
fillwords
                   = folddoc (<+/>) . map text . words
fill []
                   = nil
fill [x]
fill (x:y:zs)
                   = (flatten x <+> fill (flatten y : zs))
                      :<|> (x </> fill (y : zs)
```

61

# Overview of the Code 6(11)

65

67

# Overview of the Code 7(11)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

66

# Overview of the Code 8(11)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

# Overview of the Code 9(11)

# Overview of the Code 10(11)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

69

71

# Further Readings 1(2)

On an imperative Pretty Printer

• Derek Oppen. *Pretty-printing*. ACM Transactions on Programming Languages and Systems, 2(4):465-483, 1980.

...and a functional realization of it:

• Olaf Chitil. *Pretty printing with lazy dequeues*. In ACM SIGPLAN Haskell Workshop, 183-201, Florence, Italy, 2001. Universiteit Utrecht UU-CS-2001-23.

# Overview of the Code 11(11)

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)

70

# Further Readings 2(2)

Overview on the evolution of a Pretty Printer Library and origin of the development of the *Prettier Printers* proposed by Phil Wadler.

• John Hughes. *The design of a pretty-printer library*. In Johan Jeuring, Erik Meijers (Eds.), *Advanced Functional Programming*, LNCS 925, Springer, 1995.

...a variant implemented in the Glasgow Haskell Compiler

• Simon Peyton Jones. *Haskell pretty-printer library*. http://www.haskell.org/libraries/#prettyprinting, 1997.

# Final lecture...

• Thu, June 19, 2008, lecture time: 4.15 p.m. to 5.45 p.m., lecture room on the ground floor of the building Argentinierstr. 8

Seventh (and final) assignment (as well as previous assignments)...

• Please check out the homepage of the course for details.

Advanced functional Programming (SS 2008) / Part 7 (Thu, 06/12/08)