Object-Oriented Software Construction

Bertrand Meyer
Lecture 15:

Some design principles
Container structures

CONTAINER

BOX

COLLECTION

finite

infinite

UNBOUNDED

COUNTABLE

TABLE

BAG

SET

TRAVERSABLE

HIERARCHICAL

LINEAR

RESIZABLE

INDEXABLE

CURSOR_STRUCTURE

DISPENSER

SEQUENCE

STRING

HASH_TABLE

STACK

QUEUE

RESIZABLE

COUNTABLE

COLLECTION

CONTAINER

finite

infinite

UNBOUNDED

INDEXABLE

CURSOR_STRUCTURE

DISPENSER

SEQUENCE

STRING

HASH_TABLE

STACK

QUEUE
The notion of dispenser

- Examples:
  - Stacks
  - Queues
  - Priority queues
  - ...
How big should a class be?

- The first question is how to measure class size. Candidate metrics:
  - Source lines.
  - Number of features.

- For the number of features the choices are:
  - With respect to information hiding:
    - Internal size: includes non-exported features.
    - External size: includes exported features only.
  - With respect to inheritance:
    - Immediate size: includes new (immediate) features only.
    - Flat size: includes immediate and inherited features.
    - Incremental size: includes immediate and redeclared features.
The features of a class

- Immediate
- Redefined
- Redeclared
- Effected
- New in class
- From parent
- Inherited
- Changed
- Was deferred
- Had an implementation
- Kept
- Unchanged
- Was deferred
- New in class
- From parent
- Inherited
- Changed
- Had an implementation
- Kept
- Unchanged

Most useful measure is incremental size. Easy to measure.
Incremental size

Most useful measure is incremental size. Easy to measure.
The shopping list approach

- If a feature may be useful, it probably is.
- An extra feature cannot hurt if it is designed according to the spirit of the class (i.e. properly belongs in the underlying abstract data type), is consistent with its other features, and follows the principles of this presentation.
- No need to limit classes to “atomic” features.
Some statistics from EiffelBase

- Percentages, rounded.
- 149 classes, 1823 exported features.
  (Includes Lex and Parse.)

<table>
<thead>
<tr>
<th>Feature Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 5 features</td>
<td>45</td>
</tr>
<tr>
<td>6 to 10 features</td>
<td>17</td>
</tr>
<tr>
<td>11 to 15 features</td>
<td>11</td>
</tr>
<tr>
<td>16 to 20 features</td>
<td>9</td>
</tr>
<tr>
<td>21 to 40 features</td>
<td>13</td>
</tr>
<tr>
<td>41 to 80 features</td>
<td>4</td>
</tr>
<tr>
<td>81 to 142 features</td>
<td>1</td>
</tr>
</tbody>
</table>
Some statistics from EiffelVision

- Percentages, rounded. 546 classes, 3666 exported features.

<table>
<thead>
<tr>
<th>Feature Range</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 5 features</td>
<td>68</td>
</tr>
<tr>
<td>6 to 10 features</td>
<td>12</td>
</tr>
<tr>
<td>11 to 15 features</td>
<td>7</td>
</tr>
<tr>
<td>16 to 20 features</td>
<td>4</td>
</tr>
<tr>
<td>21 to 40 features</td>
<td>6</td>
</tr>
<tr>
<td>41 to 78 features</td>
<td>2</td>
</tr>
</tbody>
</table>
Including non-exported features

- Percentage rounded. All features (about 7600).

<table>
<thead>
<tr>
<th>Feature Range</th>
<th>Base</th>
<th>Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 5 features</td>
<td>37</td>
<td>55</td>
</tr>
<tr>
<td>6 to 10 features</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>11 to 15 features</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>16 to 20 features</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>21 to 40 features</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>41 to 80 features</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>81 or more features</td>
<td>2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

- Ratio of total features to exported features: 1.27 (EiffelBase), 1.44 (EiffelVision)
Minimalism revisited

- The language should be small (ETL: “The language design should provide a good way to express every operation of interest; it should avoid providing two.”)

- The library, in contrast, should provide as many useful facilities as possible.

- Key to a non-minimalist library:
  - Consistent design.
  - Naming.
  - Contracts.

- Usefulness and power.
The size of feature interfaces

- More relevant than class size for assessing complexity.
- Statistics from Base, Lex and Parse (exported features only):

<table>
<thead>
<tr>
<th>Number of features</th>
<th>1823</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of queries</td>
<td>59%</td>
</tr>
<tr>
<td>Percentage of commands</td>
<td>41%</td>
</tr>
<tr>
<td>Average number of arguments to a feature</td>
<td>0.4</td>
</tr>
<tr>
<td>Maximum number</td>
<td>3</td>
</tr>
<tr>
<td>No argument</td>
<td>60%</td>
</tr>
<tr>
<td>One argument</td>
<td>37%</td>
</tr>
<tr>
<td>Two arguments</td>
<td>3%</td>
</tr>
<tr>
<td>Three arguments</td>
<td>0.3%</td>
</tr>
</tbody>
</table>
### Including non-exported features:

<table>
<thead>
<tr>
<th>Average number of arguments to a feature</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum number</td>
<td>6</td>
</tr>
<tr>
<td>No argument</td>
<td>57%</td>
</tr>
<tr>
<td>One argument</td>
<td>36%</td>
</tr>
<tr>
<td>Two arguments</td>
<td>5%</td>
</tr>
<tr>
<td>Three arguments</td>
<td>1%</td>
</tr>
<tr>
<td>Four arguments</td>
<td>0.6%</td>
</tr>
<tr>
<td>Five or six arguments</td>
<td>0.2%</td>
</tr>
</tbody>
</table>
The size of feature interfaces (cont’d)

- Statistics from EiffelVision (546 classes, exported features only):

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of features</td>
<td>3666</td>
</tr>
<tr>
<td>Percentage of queries</td>
<td>39%</td>
</tr>
<tr>
<td>Percentage of commands</td>
<td>61%</td>
</tr>
<tr>
<td>Average number of arguments to a feature</td>
<td>0.7</td>
</tr>
<tr>
<td>Maximum number</td>
<td>7</td>
</tr>
<tr>
<td>No argument</td>
<td>49%</td>
</tr>
<tr>
<td>One arguments</td>
<td>32%</td>
</tr>
<tr>
<td>Two arguments</td>
<td>15%</td>
</tr>
<tr>
<td>Three arguments</td>
<td>3%</td>
</tr>
<tr>
<td>Four arguments</td>
<td>0.4%</td>
</tr>
<tr>
<td>Five to seven arguments</td>
<td>0.4%</td>
</tr>
</tbody>
</table>
Operands and options

- Two possible kinds of argument to a feature:
  - Operands: values on which feature will operate.
  - Options: modes that govern how feature will operate.

- Example: printing a real. The number is an operand; format properties (e.g. number of significant digits, width) are options.

$$\text{print} \ (\text{real\_value}, \text{number\_of\_significant\_digits}, \text{zone\_length}, \text{number\_of\_exponent\_digits}, \ldots)$$

$$\text{my\_window\_display} \ (\text{x\_position}, \text{y\_position}, \text{height}, \text{width}, \text{text}, \text{title\_bar\_text}, \text{color}, \ldots)$$
Recognizing options from operands

- Two criteria to recognize an option:
  - There is a reasonable default value.
  - During the evolution of a class, operands will normally remain the same, but options may be added.
Operands and options

- The Option Principle:
  - The arguments of a feature should only be operands.

- Options should have default values, with procedures to set different values if requested.

- For example:

  ```python
  my_window.set_background_color("dd-blue")
  ...
  my_window.display
  ```
### Useful checklist for options:

<table>
<thead>
<tr>
<th>Option</th>
<th>Default</th>
<th>Set</th>
<th>Accessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window color</td>
<td>White</td>
<td>set_background_color</td>
<td>background_color</td>
</tr>
<tr>
<td>Hidden?</td>
<td>No</td>
<td>set_visible</td>
<td>hidden</td>
</tr>
<tr>
<td></td>
<td></td>
<td>set_hidden</td>
<td></td>
</tr>
</tbody>
</table>
## Naming (1)

<table>
<thead>
<tr>
<th>Class</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARRAY</td>
<td>enter, entry</td>
</tr>
<tr>
<td>STACK</td>
<td>push, top, pop</td>
</tr>
<tr>
<td>QUEUE</td>
<td>add, oldest, remove_oldest</td>
</tr>
<tr>
<td>HASH_TABLE</td>
<td>insert, value, delete</td>
</tr>
</tbody>
</table>
## Naming (2)

<table>
<thead>
<tr>
<th>Class</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARRAY</td>
<td>put</td>
</tr>
<tr>
<td>STACK</td>
<td>put</td>
</tr>
<tr>
<td>QUEUE</td>
<td>put</td>
</tr>
<tr>
<td>HASH_TABLE</td>
<td>put</td>
</tr>
</tbody>
</table>
Naming rules

- Achieve consistency by systematically using a set of standardized names.

- Emphasize commonality over differences.

- Differences will be captured by:
  - Signatures (number and types of arguments and result).
  - Assertions.
  - Comments.
Some standard names

- Queries:
  - \texttt{count}
  - \texttt{item}
  - \texttt{infix "@"}
  - \texttt{to\textunderscore external, to\textunderscore c, from\textunderscore external}

- Commands:
  - \texttt{make} -- For creation
  - \texttt{put, extend, replace, force}
  - \texttt{remove, prune, wipe\textunderscore out}

- Boolean queries:
  - \texttt{writable, readable, extendible, prunable}
  - \texttt{empty, full, capacity}

- Array access:
  - \texttt{a.item(i) or a @ i}

- Rejected names:
  - \texttt{if s.addable then}
    - \texttt{s.add(v)}
  - \texttt{end}
  - \texttt{if s.deletable then}
    - \texttt{s.delete(v)}
  - \texttt{end}

- Usual invariants:
  - \texttt{empty = (count = 0)}
  - \texttt{full = (count = capacity)}
Grammatical rules

- Procedures (commands): verbs, infinitive form. Examples: *make, put, display*.

- Boolean queries: adjectives, e.g. *full*. Also (especially in case of potential ambiguity) names of the form *is_some_property*. Example: *is_first*.
  - In all cases, you should usually choose the form of the property that is false by default at initialization (making it true is an event worth talking about). Example: *is_erroneous*.

- Other queries: nouns or adjectives. Examples: *count, error_window*.

- Do not use verbs for queries, in particular functions; this goes with the command-query separation principle (prohibition of side-effects in functions).
Feature categories

\begin{verbatim}
class C inherit

... feature -- Category 1
    ... Feature declarations

feature {\texttt{A, B}} -- Category 2
    ... Feature declarations

feature \{\texttt{NONE}\} -- Category n
    ... Feature declarations

invariant

... end
\end{verbatim}
Feature categories (cont’d)

- Standard categories (the only ones in EiffelBase):
  - Initialization
  - Access
  - Measurement
  - Comparison
  - Status report
  - Status setting
  - Cursor movement
  - Element change
  - Removal
  - Resizing
  - Transformation
  - Conversion
  - Duplication
  - Basic operations
  - Obsolete
  - Inapplicable
  - Implementation
  - Miscellaneous
Obsolete features and classes

- A central problem in the computer field: how to reconcile progress with the protection of the installed base?
- Obsolete features and classes support smooth evolution.
- In class `ARRAY`:

  ```
  enter (i: V; v: T) is
  obsolete "Use `put (value, index)`"
  do
    put (v, i)
  end
  ```
class \texttt{ARRAY\_LIST} [G] \texttt{obsolete}

"[

Use \texttt{MULTI\_ARRAY\_LIST} instead (same semantics, but new name ensures more consistent terminology). Caution: do not confuse with \texttt{ARRAYED\_LIST} (lists implemented by one array each).

]"

\texttt{inherit}

\texttt{MULTI\_ARRAY\_LIST} [G]

\texttt{end}
End of lecture 15