Format Strings	$\{A\}$ $\Delta$ format B [C0 C1]
	{A} Δf B [C0 C1]

Format strings provide a simple way to display a mixture of text, APL variables, and arbitrary expressions in a straightforward way, including (where required) precise formatting of native multidimensional and multi-line objects.<sup>1</sup> They are similar to **f-strings** in python<sup>2</sup> and similar constructions in other languages, but designed for handing APL arrays.

#### **Preview**

Before we get started, we'll share two examples to give you a sense of what format strings are all about

Example 1: Here is a most simplistic example to give you a flavor of format strings:

Example 2: And here is an example that uses some more of the power of  $\Delta \mathbf{f}$  and APL:

```
Names←'John Jones' 'Mary Smith' 'Jan Ito'
Salaries←125000 132000 85000

□ Here we apply APL ↑ to Names,
□ and use a shortcut to APL □FMT specs to format Salaries.

Δf'Name{8}Salary±↓{↑Names} {⊂£⊃,G⊂ZZ9,999⊃$Salaries}'
Name Salary
John Jones £125,000
Mary Smith £132,000
Jan Ito £ 85,000
```

#### **Arguments**

#### {A} (Options):

**A=1** <u>default</u>:  $\Delta$ **format or**  $\Delta$ **f** returns a formatted version of its right-argument string. If a single line, a character vector is returned; if more than one line, a character matrix is returned. (If multidimensional objects are formatted, they are mapped into the matrix result, with appropriate spacing, just as for  $\Box$ **FMT**).

**A=0:**  $\Delta \mathbf{f}$  returns a parsed string of pseudo-code that shows exactly how the right argument was interpreted. The pseudo-code is normally not executable, but useful for debugging.

**A=2:** Echos the  $\Delta f$  command and argument string (**C**) in a canonical form:

```
Δformat 'string' ['pos1' 'pos2' ...]
```

¹ In this documentation, we'll use the term *format string*, with  $\Delta f$  standing in for the synonyms,  $\Delta f$  and  $\Delta f$ . Several field types, including **Code Fields**— {}-format strings— will be demonstrated. Name and Quote Fields, which begin with  $\pm$ , have similar options, but may not appear in the final release of  $\Delta f$  (see Appendix A).

<sup>□</sup>**IO** □**ML←0 1** will be used throughout.

**Help Info:** For help information within the  $\Delta f$  command, enter:  $\Delta f$  ' $\Box$ ?'

For typical initialization of format functions, see **Appendix A** (below).

<sup>&</sup>lt;sup>2</sup> Format strings here are not based on any particular language's constructions.

## **B** (Format String):

The primary right argument to  $\Delta \mathbf{f}$  is a single character vector (after any evaluation, catenation, etc.), including both literal text along with special characters like  $\{\}$ ,  $\pm$ , \$, and  $\Box$ , to indicate APL variables and code to evaluate, along with specifications for formatting their output.

## C (Positional and Immediate Variables).:

The primary right argument, **B**, may contain references to external objects by name or through positional objects passed to the  $\Delta f$  function, directly after the **format string**. Each of these positional values **C0**, **C1**, ..., is handled in one of two ways, depending on whether it is of the form  $\omega nnn$  or  $\alpha nnn$ , where nnn is an integer such that 0 refers to the first item *after* the format string, 1 to the second, and so on.

**Positional Variables** are denoted by  $\omega 0$ ,  $\omega 1$ , ..., which are treated like any named APL variable referenced within the format string and may be of any expected type; and

**Immediate Variables** are denoted by  $\alpha 0$ ,  $\alpha 1$ , ..., whose text is immediately substituted into the format string *before* any other processing; these must be convertible to character vectors (or scalars), via  $\overline{\alpha}0$ , etc. These are useful for setting headers, footers, and formatting specifications based, for example, on data with changing length or type.

If a positional or immediate variable does not refer to an actual argument to  $\Delta \mathbf{f}$ , its value defaults to its name (e.g.  $\omega \mathbf{1}$  by default has ' $\omega \mathbf{1}$ ' as its literal value).

#### Result:

 $\Delta$ f by default returns the formatted expression that results from evaluating the expressions and specifications in **B**. If valid, that expression is *guaranteed* to be a simple character vector, if the result fits on a single APL line; otherwise, the expression is a character matrix.

# Specifications within B (the format string)

Fields: Format Strings can be divided into fields. Each field will be formatted as a 2-dimensional, rectangular object, one catenated to the next, left to right.

#### ∆f 'This is an example ∮◊of two lines.'

This is an example Of two lines.

Code Fields: Code to be evaluated at run-time is entered within braces {...}. Code may include any standard APL, including dfns, with some extensions described below; double-quoted strings will be recognized (making string constants easy to enter) and appropriately converted to single-quoted strings on execution. Finally, Code Fields may not consist of a single unadorned integer (surrounded by 0 or more spaces), e.g. {15}; that syntax is reserved for Space Fields, below.

In this example, we have three independent fields, the **Literal Field** on the left ends when the **Code Field** (with its left brace) starts, and the **Literal Field** on the right begins, when the **Code Field** ends (with the right brace).

```
\Delta f 'One \pm \delta two \pm \delta three { \uparrow "cat" "dog" "mouse"} four \pm \delta five two dog five three mouse
```

Note how the first "field," consisting of "One...three" is self-contained vertically, as are the other two fields, "{...}" and " four...five". {} alone represents a null field, whose only function is to separate simple text fields or other fields:

```
Af 'One ± 0 two ± 0 three { bar} cat ± 0 dog ± 0 mouse { bar ← ↑3ρ ⊂ " | "} four ± 0 five'

One | cat | four

two | dog | five

three | mouse |
```

Null/Space Fields:  $\Delta f$  also supports space fields. The sequence {nnn} will insert an nnn-space wide field, where nnn is a simple non-negative integer:

three mouse

14 25

Here, the left-most field is separated from the middle field by 3 spaces, which is separated in turn from the right-most field by 2 spaces. To evaluate code that consists of a single integer, enter anything in the field besides an integer,  $^3$  e.g.  $\{2.0\}$ ,  $\{,2\}$ ,  $\{\vdash 2\}$ , etc.

A pair of single literal braces may be entered into a string by doubling to  $\{\{...\}\}$ .

```
\Delta f 'This is {{a}} test.' This is {a} test.
```

A 0-length space field **{0}** or its variant **{}** can be used to separate each rectangular (2-dimensional) field from the next, without adding more spacing [Note: our next example will do this more elegantly]:

```
Δf '1±02±03{}4±05±06'
```

When bare braces are used like this {} to conclude one field before beginning another, we call that a Null Field, equivalent to {0}, a zero-length Space Field.

```
\Delta f ' \{ 71+13 \} \{ 74+13 \} ' \cap Same output as above!
```

Unicode and Numeric pseudo-variables: Literal braces, { and }, may also be entered as □UCS pseudo-variables, □U123 and □U125, or □U7BX and □U7DX:

<sup>&</sup>lt;sup>3</sup> Extra spaces are also ignored within {nnn} fields, i.e. { 5 } is the same as {5}. Note that negative integers like { 3} are invalid in Space Fields; use a literal field or add other code: { -3}.

```
\Delta f 'This is \square U123a\square U125 \square U7BXthe\square U7DX test.' This is {a} {the} test.
```

More generally, any Unicode character may be entered via decimal  $\square U ddd$  or hexadecimal  $\square U dhhx$  (or  $\square U dhhx$ ), where d is in [0-9] and h in [0-9a-fA-F].  $\square N ddd$  specifies an integer in decimal which is displayed as hexadecimal, and  $\square N dhh$  specifies hexadecimal to decimal:

```
\Delta f '\alpha 0 in hex is \Box N\alpha 0. And \alpha 1x in decimal is \Box N\alpha 1X' 45 '2D' 45 in hex is 2D And 2Dx in decimal is 45
```

Code Expressions: The general form for a code field is:

```
{[[fmt1|fmt2]+ '$'] code} or { [[fmt1|fmt2]+ '$$'] code}
```

Where *fmt1* consists of the standard formatting codes of APL **FMT** without quotes and *fmt2* consists of special *extensions* to APL **FMT** for left- and right-justification and centering. The latter may be used with *any* APL extension, but the former must be within **FMT**'s domain. Let's start with **fmt1**, **FMT**-code **expressions**:

#### □FMT-Code Expressions<sup>5</sup>:

```
Δf '{I4$1+ι3}{I3, □ <□,I2,□> □,I2$3 3ρ4+ι9}'
2 5 < 6> 7
3 8 < 9> 10
4 11 <12> 13
```

Here, the single dollar-sign (\$) both delimits formatting from code, while denoting that the code is treated according to **DFMT** rules<sup>6</sup>, whereby a vector right argument, here **2 3 4**, is treated as a one-column matrix, as if:

```
'I4' □FMT 2 3 4
```

The expression 3  $3\rho4+\iota9$  is acted on by  $\Box$ FMT specifications 'I3, $\Box$  < $\Box$ ,I2, $\Box$ >  $\Box$ ,I2'; that is, a 3-digit integer, the text '<', a 2-digit integer, the text '>', and finally a 2-digit integer.

The double dollar sign (\$\$) functions identically to the single dollar sign (\$), *except* that a right-hand vector will be treated as if a one-row matrix, contrary to **DFMT** defaults, rather than as a one-column vector.

<sup>&</sup>lt;sup>4</sup> To escape □**U**ddd or □**U**hhhX, use □□**U...**: □□**U123** appears as the literal string □**U123**.

<sup>&</sup>lt;sup>5</sup> For delimiters, use those valid for □FMT, i.e. : ⊂ ⊃, ¨ ¨, □ □, < >, or □ □.

<sup>&</sup>lt;sup>6</sup> APL expressions *not* within the scope of a dollar-sign (\$) follow normal APL conventions, , i.e. not those of □FMT

□FMT-Code Extensions: There are currently three additional formatting expressions within Code Fields, which affect the evaluated code expression *code* to the right of the \$ (or \$\$):

Cnnn	C <i>nnn⊂c</i> ⊃	Center <b>code</b> in a field <i>nnn</i> characters wide.
L <i>nnn</i>	L <i>nnn</i> ⊂ <i>c</i> ⊃	Left-justify <i>code</i> in a field <i>nnn</i> characters wide.
R <i>nnn</i>	R <i>nnn⊂c</i> ⊃	Right-justify <b>code</b> in a field <i>nnn</i> characters wide.

When  $\neg c \neg$  is specified, a single character c will be used to pad the field; it may be specified as follows:

- A literal character, including a single digit, **except** an enclosing delimiter, a single or double quote, parenthesis, or brace (see 2nd and 3rd options<sup>7</sup>);
- A number *nn* of 2 *or more* digits, where *nn*>32, replaced on output by □ucs *nn*. Any value out of range will be treated as the **RC** (Unicode replacement character) below.
- Any of these 2-character names:

```
SQ single quote
' LP left parenthesis (
DQ double quote
" RP right parenthesis )
MD middle dot (□U183)
LB left brace {
SP space
'' RB right brace }
KS Kanji (wide) space (□U12288)
RC replacement character (□U65533)
```

Here's an example:

**Miscellaneous Features**: Here we describe several useful advanced features before diving into the fundamentals of advanced features of f-strings.

Headers and Footers: There F-strings allow for (multi-line) headers and footers, that are centered above or below the entire formatted object. The basic format uses the *option prefix*  $\Box$ , a letter [H|F], and a  $\Box$ FMT-style quoted specification:

```
□H⊂header⊃
□F⊂footer⊃
```

All special symbol constructors like  $\square Unnn$ ,  $\bullet \lozenge$ , and so on are available. In this example, we add a single-line header and a 2-line footer.

 $<sup>^7</sup>$  This is a limitation (and workaround) due to how format strings are scanned for  $^4$  () {}" in the prototype.

Positional  $\omega n$  and Immediate Variables  $\alpha n$ : In the example above, the format statement got long, so we catenated the header and footer strings on the  $\Delta f$  command line. An alternative is to use an *immediate variable*. Immediate variables are of the form  $\alpha \theta$  to  $\alpha \theta$ , where  $\alpha \theta$  instructs the format command to insert the *literal text* of the 0-th positional value to the right of the format string (and so on for  $\alpha n$  and the n-th positional value); that text is treated as a simple character scalar or vector<sup>8</sup> that is inserted into the format string **before** its other components are evaluated. Here we alter the format above to integrate the values directly into the string; while equivalent, this has the advantage of making the format statement itself more concise and easier to read:

Here's another example, where a header ( $\alpha$ 0) and a footer ( $\alpha$ 1) both contain embedded Code Fields with formatting specifications:

```
hdr←'⊡H"{C19⊂+⊃$" The ± ◊ Whole ± ◊ Story"}"'
ftr←'⊡F"{C19⊂⁻⊃$""}"'

Δf 'αθα1{F6.3$?2 3ρθ}' hdr ftr

+++++ The +++++
+++++ Whole ++++++
0.752 0.462 0.197
0.912 0.347 0.888
```

\_

<sup>&</sup>lt;sup>8</sup> If an immediate variable  $\alpha n$  cannot be converted to a character string via  $\pi \alpha n$ , an error occurs.

Positional vs. Immediate Variables: Positional variables of the form  $\omega 0$  have a different role than immediate variables  $\alpha 0^9$ . Each positional variable is treated as a full-fledged variable useful only in code expressions similarly to  $\omega$  in APL dfns; it can be of any shape or type that makes sense in the context. In the above example, we could use positional variables for Names, Scores, and Grades. Positional and immediate variables share the same numbering from left to right, based on absolute position to right of the format string (starting at 0). This example uses two immediate and three positional variables:

```
    Student Grades #3:
    This has same value as the Student Grades example above...
    Δf 'αθα1{↑ω2} {¬ω3} {↑ω4}' Hdr Ftr Names Scores Grades
```

Sequential Positional  $(\omega\omega)$  and Immediate  $(\alpha\alpha)$  Variables: As a shortcut, useful in some cases, instead of specifying a format string with positional or immediate variables in sequence  $(\omega 0 \ldots \omega 1 \ldots \omega 2 \ldots \alpha 3)$ ,  $\omega\omega$  or  $\omega$  can be used to refer to the *next* positional or immediate variable in sequence, respectively. That is, if  $\omega$ 5 was just referenced<sup>10</sup>, reading left to right, then a  $\omega\omega$  to its right will refer to  $\omega$ 6; a subsequent  $\omega\omega$  will refer to  $\omega$ 7. Immediate variables work similarly, based on prior positional variables (each has its own counter, so  $\omega\omega$  will never depend on prior immediate variables or vice versa).

Thus, cases (a) and (b) below are equivalent:

```
Name<'John Jones' \lozenge Addr<'41 Maiden Ln' \lozenge Tel<'555-1234' \lozenge Case (a) \Delta f 'Name \{\omega\omega\}, Address \{\omega\omega\}, Tel. Number \{\omega\omega\}' Name Addr Tel \lozenge Case (b) \Delta f 'Name \{\omega0\}, Address \{\omega1\}, Tel. Number \{\omega2\}' Name Addr Tel
```

Note that  $\alpha\alpha$  is used as an "escape" to enter a single  $\alpha$  into a literal string, but is only required when followed by a digit (i.e.  $\alpha\alpha$  alone is simply the literal ' $\alpha\alpha$ '):

```
\Delta f 'Alpha (\alpha), Alpha0 (\alpha\alpha0), Omega (\omega), and Question Mark (\alpha\omega)' '?' Alpha (\alpha), Alpha0 (\alpha0), Omega (\omega\omega), and Question Mark (?)
```

**Use of standard APL** ω (or α) in Code Fields: What if you want to ignore all this business about ω0, ωω, and so on in Code Fields? You can simply use standard APL ω as you normally would. Assuming □**I0**=**0**, (0⊃ω) will refer to the first item in ω, (-2↑ω) will gather the last two items, and ω will designate the entire right argument of  $Δ\mathbf{f}$ , after the format string (which is inaccessible).

```
\Delta f 'The sum of \{\{\{\omega\}\}\}\} is: \{+/\omega\} ' 10 20 30 40 A \{\{\}\}\} is: 100
```

<sup>&</sup>lt;sup>9</sup> The highest positional variable is ω**999**. By default, the highest immediate variable is α**9**, to make it easy to directly juxtapose immediate variables and other text, e.g. α**99** would map to the string 'K-9' if α**9** is 'K-'. The option □**A2** will allow immediate variables from α0 to α99. Only □**A1** (default) and □**A2** are supported.

<sup>&</sup>lt;sup>10</sup> When ωω or  $\alpha\omega$  is used before any other ω- or  $\alpha$ -reference, it is always ω0 or  $\alpha$ 0, respectively.

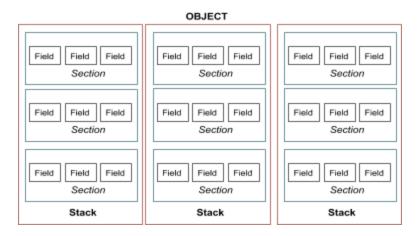
<sup>&</sup>lt;sup>11</sup> More generally, **ωω** always refers to **ωN+1**, if the prior positional variable was **ωN**, whether specified directly (e.g. as ω6) or a prior  $\omega\omega$ .

Order of Execution for Code Fields<sup>12</sup>: If there are multiple Code Fields within a format string, they are executed in APL order, right to left. Variables set in a Code Field (on the right) will be visible further left, but discarded after the format string has been returned. To ensure a variable is visible in the calling environment, either modify it in a Code Field (e.g. a, ←' ') or use global assignment (∘←). Here we assign a locally, then refer to it in the next Code Field to the left.

#### Advanced Features:

Fields, Sections, Stacks, and Objects:

Format is designed hierarchically, starting with fields (rectangular or character data) built of APL multidimensional and multi-line data and formatting specifications. Fields are catenated left-to-right into a 2-D grouping called a section that can operate as a unit. In simple cases, there is one section. In more complex cases, sections are input left to right, then displayed as a stack, top to bottom; they can be joined left-justified, centered, or right-justified. A set of sections, one on top of the other, constitutes a stack; there can be one stack in a formatted string, or several. Finally, stacks can be aligned together left-to-right, like fields, constituting a formatted object. An object can be decorated with a header and a footer, centered over the entire object. While no formatted object would sensibly be this complicated, this diagram maps out the hierarchy from field to section to stack to object:



Summary of Object-Building: We've shown the building blocks of several kinds of fields: literal fields, code fields, and null and space fields. Literal fields can be built into 2-D blocks using ≜◊ or □U unicode characters. Code fields are structured using formatting specifications (□FMT-based or extensions) on APL arrays. Null fields and space fields define horizontal size (in spaces) that automatically match the height (number of lines) of surrounding fields. As those fields are catenated left to right, they produce a section.

<sup>&</sup>lt;sup>12</sup> This applies as well to **Name Fields** and **Quote Fields** (see **Appendix A**).

Sections can be joined together into stacks, one over the other using the section-end **delimiter ≜** ★ after each stack (the delimiter may be omitted after the *last* section in the format string).

Finally, stacks can be joined left to right into a single object (with optional object-level header and **footer**) using the **stack-ending delimiter ±→** (the delimiter may be omitted after the *last* stack in the format string).

**Object-Building Delimiters** 

Level	Function	Symbols	Builds	Direction
Field	Separator	{}	Section	Horizontal
Section	Delimiter	•↑	Stack	Vertical
Stack	Delimiter	<u></u>	0bject	Horizontal

# **Advanced Miscellany**

 $1_{\Gamma} - - -$ 1 2 3

1

**Building Sections into Stacks:** When building a stack vertically, one section may be wider than another. Options **L**, **C**, and **R** are placed within the *upper* section (before the delimiter ★◆) to indicate that that section should be placed on the left (□L), central (□C), or right (**BR**) side of the next section down. By **default**, sections are stacked **left-**justified (**BL**).

Δf '□R{C11<1>\$box 2 3ρι6}±ψ□L{C15<2>\$box 2 3ρι6}±ψ□C{C21<3>\$box2 3ρι6}'

1

Format Strings Δformat, Δf

# **Escaping Formatting Characters:** Formatting characters can be escaped (displayed as literals) as follows:

Escape Sequence	Evaluates as	Comments	
{{	{	If not balanced, use <b>□U123</b> .	
}}	}	If not balanced, use <b>□U125</b> .	
<u> </u>	<u> </u>	E.g. <b>±±◊</b> displays as <b>±◊</b> .	
□□Uddd	□U <i>ddd</i>	Bare <b>□U</b> requires no escaping.	
□□Nddd	□N <i>ddd</i>	Bare □N requires no escaping.	
ΞΞA	ΞA	Escape required only for [A-Z].	
ααΘ	αΘ	$\alpha\alpha$ w/o following num is ' $\alpha\alpha$ ' .	
ω0, ω1,	ω0, ω1,	Is literal, outside Code Field.	
αN <b>or</b> αω	αN or αω	If $n$ -th pos'l arg omitted, $\alpha$ $N$ is literal ' $\alpha$ $N$ ', and $\alpha \omega$ is literal ' $\alpha \omega$ '.	

**Faux Space Option ( Esc)**: Sometimes it's easier to see what's happening in a literal string with spaces by substituting another, visible character for spaces on input. Likewise, it's sometimes useful to separate out different parts of the format string on input just for clarity without adding extra spacing on output. The **Faux Space** option allows you to address either or both of these goals for the entire format string. **ESc** specifies that the single character **c** will be replaced by a space, wherever it occurs; at the same time, all space characters in the format string are removed (whether in quotes, code, or otherwise). Again, the **Faux Space** applies to the *entire* format string.

# Appendix A

## Setting up run-time access to format strings

To make the format functions and run-time routines available, copy in namespace **format** from file **format.dyalog**. You can make format strings available for the current APL session using the following function:

If you prefer to load **format** permanently into, for example, the top-level namespace #, simply replace (only) the two underscored **\( \pi SE \)** above with #, or otherwise adjust for your use. Or use the **\( \begin{array}{c} \) Load** command:

```
]load format
)save
```

Loading file **format** will expose functions  $\Delta$ **format** and  $\Delta$ **f**, as well as the background utility function **formatPath**, which will point several other service routines to the active format namespace, minimizing the pollution of the user namespace.

**Name Fields and Quote Fields:** Similar to Code Fields, f-strings also support two other "experimental" fields, **Name** and **Quote Fields**. They are described only in Appendix A.

```
NF:
             ♠ [fmt $$? ] [simple_code | '(' any_code ')' ] name
             ★ [fmt $$? ] [simple_code | '(' any_code ')' ] quoted_string
      OF:
Where
                           a mixture of fmt1 and fmt2 specs as for Code Fields<sup>13</sup>;
      fmt
       $$?
                           either $ or $$ as for Code Fields;
      simple_code
                           any APL core functions or operators, integers, and spaces;
                            Names, including system names (IO, etc.), are not allowed.
       any_code
                           an arbitrary APL expression within balanced parentheses.
                           an APL variable name with optional indexing
       name
      quoted_string
                           a double-quoted or single-quoted string.
```

Using **simple\_code** allows selection, rotating, incrementing, and other simple operations, which can be distinguished from an APL variable expression (name plus index) or quoted string. A typical example of a **Name Field** might be:

<sup>&</sup>lt;sup>13</sup> Where **fmt1** consists of the standard formatting codes of APL □FMT, omitting quotes, and **fmt2** consists of special *extensions* **L**, **R**, **C** to APL □FMT for left- and right-justification and centering. The latter may be used with *any* APL extension, but the former must be within □FMT's domain.

Format Strings Δformat, Δf

```
AR←?2 2 3ρ0
    Δf '±C20$"This is a heading"±↓±F6.3$ 4 3ρAR'
This is a heading
0.092 0.150 0.093
0.096 0.230 0.567
0.360 0.228 0.489
0.561 0.317 0.573
```

Name Fields and Quote Fields are experimental features, which duplicate the functionality of Code Fields. In their most complex syntax, where (any\_code) is used, some may view it as unwieldy and less elegant or obvious than Code Fields, which use braces reminiscent of dfns. On the other hand, in simpler cases, e.g. printing simple values like the following, Name Fields and Quote Fields may seem very natural to APL users.

```
    Example 1
    AR←?2 2 3ρθ
    hdg←'">>> This is a heading<><"'
    Δf'±C30$αθ ±↓□IO is ±□IO, AR[1;1;] is ±F5.2$$AR[1;1;]' hdg
    >>> This is a heading<><
□IO is 1, AR[1;1;] is 0.48 0.69 0.81

    Example 2
    Δf 'The sum of ±ω is: {+/ω} ' 10 20 30 40

The sum of 10 20 30 40 is: 100
</pre>
```