

# ispeggra(1), html2txt(1) and txt2i(1)

2012-04-06 1.0

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	NAME . . . . .	2
1.2	SYNOPSIS . . . . .	2
1.3	DESCRIPTION . . . . .	2
1.4	OPTIONS . . . . .	2
1.5	EXIT STATUS . . . . .	3
1.6	EXAMPLES . . . . .	3
1.7	KNOWN BUGS . . . . .	3
1.8	OPEN ISSUES . . . . .	3
1.9	AVAILABILITY . . . . .	3
1.10	AUTHOR . . . . .	3
1.11	SEE ALSO . . . . .	3
<b>2</b>	<b>Items</b>	<b>4</b>
2.1	Heading items . . . . .	4
2.2	Caption items . . . . .	4
2.3	Bibliography and equation items . . . . .	4
2.4	Paragraph items . . . . .	5
2.5	List items . . . . .	5
2.6	Quotation items . . . . .	6
2.7	Pre-formatted items . . . . .	6
2.8	Picture items . . . . .	6
2.9	Latex items . . . . .	7
2.10	Man items . . . . .	7
2.11	Interrupt items . . . . .	7
2.12	Footnotes . . . . .	7
<b>3</b>	<b>Modes</b>	<b>8</b>
<b>4</b>	<b>Arrangement</b>	<b>8</b>
4.1	One-dimensional arrangement . . . . .	8
4.2	Two-dimensional arrangement (tables) . . . . .	8
<b>5</b>	<b>Bibliography</b>	<b>9</b>

# 1 Introduction

## 1.1 NAME

`ispegra` — `i(1)` spelling/grammar checker

`html2txt` — *HTML* to unformatted text converter

`txt2i` — unformatted text to `i(1)` converter

## 1.2 SYNOPSIS

```
ispegra [ -u ] [ -d diffCmd ] -c checkCmd [--] iFile
html2txt -f tmpFile [ -u ] [ -h ] [ -? ] [ -V ] [ -- ] [ htmlInputFile ]
txt2i [ -h ] [ -? ] [ -V ] [ -- ] [ txtInputFile ]
```

## 1.3 DESCRIPTION

The `ispegra(1)` utility opens the spelling/grammar checker given by the `checkCmd` with the unformatted text parts of the `iFile` and stores the corrected outcome in the `iFile`. The content of the `iFile` follows the rules of `i(1)`. The meaning of unformatted text is explained in chapters 2, 3 and 4. The `diffCmd` allows to track the modifications which have been carried out in the `iFile`. The `ispegra(1)` utility is based on the `html2txt(1)` and `txt2i(1)` utilities.

The `html2txt(1)` utility reads the `htmlInputFile` (or standard input if none is given) and produces on standard output the unformatted text equivalent of its input. Merely use the `html2txt(1)` utility in combination with the `i2html(1)` utility. In order to be capable of restoring the original `i(1)` formatting, the `html2txt(1)` utility writes supplementary information to the `tmpFile` (and further files with the same base name). Avoid calling the `html2txt(1)` utility directly; use the `ispegra(1)` utility instead.

The `txt2i(1)` utility reads the `txtInputFile` (or standard input if none is given) and supplementary output of the `html2txt(1)` utility and produces on standard output the `i(1)` equivalent of its input. Avoid calling the `txt2i(1)` utility directly; use the `ispegra(1)` utility instead.

IMPORTANT: Only under the following circumstances the `ispegra(1)` and `txt2i(1)` utilities produce proper `i(1)` output:

- There are no lines added to or removed from the unformatted text.
- There are no characters with a special meaning to `i(1)` inserted into the unformatted text.

Use the `diffCmd` to verify that none of these conditions is accidentally violated.

## 1.4 OPTIONS

`-c checkCmd` The spelling/grammar check command.

`-d diffCmd` The file compare command.

`-f tmpFile` The base name of the files (various extensions) which contain the necessary information to restore the original `i(1)` formatting.

`-h` Give a bit of help about the command line arguments and options.

**-u** The input is interpreted as being *UTF-8* encoded with the `-u` option and as being *ISO-8859-1* encoded without the `-u` option.

**-v** Give version.

**-?** See option `-h`.

**--** Indicate end of options.

## 1.5 EXIT STATUS

The `ispeggra(1)`, `html2txt(1)` and `txt2i(1)` utilities exit with value 0 if the processing was successful. The occurrence of an error is indicated by an exit value 1 and an error message on standard error.

## 1.6 EXAMPLES

Execute the following command in order to use the `ispeggra(1)` utility on Mac OS X to check the spelling and grammar of the `foo1` file with the *OpenOffice* application and to track the changes with the `sdiff(1)` utility:

```
ispeggra -u -c "open -W -a /Applications/OpenOffice.org.app" -d sdiff foo1
```

Execute the following command in order to use the `ispeggra(1)` utility to check the spelling of the `foo2` file with the `aspell(1)` interactive spell checker utility and to track the changes with the `diff(1)` utility:

```
ispeggra -u -c "aspell -x -d en_GB -c" -d diff foo2
```

## 1.7 KNOWN BUGS

There are no known bugs.

## 1.8 OPEN ISSUES

There are no open issues.

## 1.9 AVAILABILITY

This document is part of the `i` project which is available on-line at the following site: <http://i2i.sourceforge.net>.

## 1.10 AUTHOR

Copyright (c) 2012 Peter Schiess

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. The GNU Free Documentation License is available on-line at the site <http://www.gnu.org/licenses/fdl.html>.

## 1.11 SEE ALSO

`i(1)`, `i2html(1)`, `diff(1)`, `sdiff(1)`

## 2 Items

### 2.1 Heading items

Apart from the differences listed in table 1, [i\(1\)](#) and unformatted text heading items are identical.

Table 1: Heading item differences

<a href="#">i(1)</a>	Unformatted text
Line breaks	No line breaks

#### 2.1.1 [i\(1\)](#) example

```
#1 Heading one
#1.1 Heading two
#1.1.1 Heading three
#1.1.1.1 Heading four
Heading five
```

#### 2.1.2 Unformatted text example

```
#1 Heading one
#1.1 Heading two
#1.1.1 Heading three
#1.1.1.1 Heading four
Heading five
```

### 2.2 Caption items

Apart from the differences listed in table 2, [i\(1\)](#) and unformatted text caption items are identical.

Table 2: Caption item differences

<a href="#">i(1)</a>	Unformatted text
Line breaks	No line breaks

#### 2.2.1 [i\(1\)](#) example

```
#T1 Table caption
#F1 Figure caption
```

#### 2.2.2 Unformatted text example

```
#T1 Table caption
#F1 Figure caption
```

### 2.3 Bibliography and equation items

Apart from the differences listed in table 3, [i\(1\)](#) and unformatted text bibliography and equation items are identical.

Table 3: Bibliography and equation item differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks

### 2.3.1 `i(1)` example

#B1 Hemingway, Ernest: The old man and the sea. 1952.  
 #B2 Orwell, George: Animal farm. 1945  
 #E1 Equation

### 2.3.2 Unformatted text example

#B1 Hemingway, Ernest: The old man and the sea. 1952.  
 #B2 Orwell, George: Animal farm. 1945  
 #E1 Equation

## 2.4 Paragraph items

Apart from the differences listed in table 4, `i(1)` and unformatted text paragraph items are identical.

Table 4: Paragraph item differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks
Indentation	No indentation

### 2.4.1 `i(1)` example

First paragraph  
 Second paragraph  
 Third paragraph

### 2.4.2 Unformatted text example

First paragraph  
 Second paragraph  
 Third paragraph

## 2.5 List items

Apart from the differences listed in table 5, `i(1)` and unformatted text list items are identical.

Table 5: List item differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks
Indentation	No indentation
List characters	No list characters
Pre-formatted items	No pre-formatted items

### 2.5.1 `i(1)` example

```
- one
- two
- three
  four
  - five
  - six
- seven
; eight  nine
```

### 2.5.2 Unformatted text example

```
one
two
three
four
five
six
seven
```

## 2.6 Quotation items

Apart from the differences listed in table 6, `i(1)` and unformatted text quotation items are identical.

Table 6: Quotation item differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks

### 2.6.1 `i(1)` example

```
Ernest Hemingway opened his book with the following words...
    He was an old man who fished alone in a skiff in the Gulf Stream
    and he had gone eighty-four days now without taking a fish.
```

### 2.6.2 Unformatted text example

```
Ernest Hemingway opened his book with the following words...
He was an old man who fished alone in a skiff in the Gulf Stream and he had gone eighty-four days
```

## 2.7 Pre-formatted items

Pre-formatted items are suppressed in unformatted text.

### 2.7.1 `i(1)` example

```
; for( i = 0; i < k; i++ ) {
:   printf("i=%d\n", i);
: }
```

## 2.8 Picture items

Picture items are suppressed in unformatted text.

### 2.8.1 `i(1)` example

```
, +-----+
: |       |
: +-----+
```

## 2.9 Latex items

Latex items are suppressed in unformatted text.

### 2.9.1 `i(1)` example

```
#F1 Photo
#I
#L }\includegraphics[scale=0.25]
#L {photo.jpg}\end{ifigure}
#M [ photo.jpg ]
#M .fi
```

## 2.10 Man items

Man items are suppressed in unformatted text.

### 2.10.1 `i(1)` example

Refer to chapter 2.9.1.

## 2.11 Interrupt items

Interrupt items are suppressed in unformatted text.

### 2.11.1 `i(1)` example

Refer to chapter 2.9.1.

## 2.12 Footnotes

Apart from the differences listed in table 7, `i(1)` and unformatted text footnotes are identical.

Table 7: Footnote differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks

### 2.12.1 `i(1)` example

To be or not to be. #(Shakespeare, William: Hamlet#)

### 2.12.2 Unformatted text example

To be or not to be. #(Shakespeare, William: Hamlet#)

### 3 Modes

The byte mode is not relevant for the conversion from `i(1)` to unformatted text and vice versa. No byte adaption is carried out.

## 4 Arrangement

### 4.1 One-dimensional arrangement

One-dimensionally arranged content can contain all kinds of items (see chapter 2.1 to 2.11) and footnotes (see chapter 2.12).

### 4.2 Two-dimensional arrangement (tables)

Two-dimensionally arranged content (tables) can only contain paragraph, list and quotation items (see chapter 2.4 to 2.6), and footnotes (see chapter 2.12).

Apart from the differences listed in table 8, `i(1)` and unformatted text tables are identical.

Table 8: Table differences

<code>i(1)</code>	Unformatted text
Line breaks	No line breaks
Indentation	No indentation
List characters	No list characters
Table borders	No table borders

#### 4.2.1 `i(1)` example of a table without a head

```

:=====:=====:=====
: one  : two  : three
:-----:-----:-----
: four : five

```

#### 4.2.2 Unformatted text example of a table without a head

```

one
two
three
four
five

```

#### 4.2.3 `i(1)` example of a table with a head

```

:-----:-----:-----
: one  : two  : three
:=====:=====:=====
: four : five

```

#### 4.2.4 Unformatted text example of a table with a head

```

one
two
three
four
five

```



#### 4.2.5 `i(1)` example of a table with various items

```
:-----:-----:-----
: one      : two : three
:=====:=====,=====
: pre      : PRE
: - list   :      quote
: - LIST   :
:-----:-----,-----
: - list   :      quote
: - LIST   : POST
: post     :
```

#### 4.2.6 Unformatted text example of a table with various items

```
one
two
three
pre
list
LIST
PRE
quote
list
LIST
post
quote
POST
```

## 5 Bibliography

**HTML** HyperText Markup Language 4.0 Specification.  
<http://www.w3.org/TR/1998/REC-html40-19980424>

**ISO-8859-1** ISO/IEC 8859-1 Information technology. 8-bit single-byte coded graphic character sets. Part 1. Latin alphabet No. 1. 1998.

**OpenOffice (TM)** The free and open productivity suite.  
<http://www.openoffice.org>

**UTF-8** RFC 3629. UTF-8, a transformation format of ISO 10646. 2003.

**aspell(1)** The free and open source interactive spell checker.  
<http://aspell.net>