

i2html(1) and html2i(1)

2010-11-07 1.1

1	Introduction	2
1.1	NAME	2
1.2	SYNOPSIS	2
1.3	DESCRIPTION	2
1.4	OPTIONS	2
1.5	EXIT STATUS	2
1.6	KNOWN BUGS	2
1.7	OPEN ISSUES	2
1.8	AVAILABILITY	2
1.9	AUTHOR	3
1.10	SEE ALSO	3
2	Items	3
2.1	Heading items	3
2.2	Caption items	3
2.3	Bibliography and equation items	4
2.4	Paragraph items	4
2.5	List items	4
2.6	Quotation items	5
2.7	Pre-formatted items	5
2.8	Picture items	6
2.9	Latex items	6
2.10	Man items	7
2.11	Interrupt items	7
2.12	Footnotes	7
3	Modes	8
4	Arrangement	14
4.1	One-dimensional arrangement	14
4.2	Two-dimensional arrangement (tables)	14
5	Bibliography	15

1 Introduction

1.1 NAME

`i2html` — `i(1)` to *HTML* converter

`html2i` — *HTML* to `i(1)` converter

1.2 SYNOPSIS

```
i2html [ -u ] [ -h ] [ -? ] [ -- ] [ iInputFile ]  
html2i [ -u ] [ -h ] [ -? ] [ -- ] [ htmlInputFile ]
```

1.3 DESCRIPTION

The `i2html(1)` utility reads the `iInputFile` (or standard input if none is given) and produces on standard output the *HTML* equivalent of its input. The `html2i(1)` utility reads the `htmlInputFile` (or standard input if none is given) and produces on standard output the `i(1)` equivalent of its input.

1.4 OPTIONS

`-u` The `i2html(1)` input is interpreted as being *UTF-8* encoded with option `-u` and as being *ISO-8859-1* encoded without option `-u`. The `html2i(1)` output is *UTF-8* encoded with option `-u` and is *ISO-8859-1* encoded without option `-u`.

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

`-?` See option `-h`.

`--` Indicate end of options.

1.5 EXIT STATUS

The `i2html(1)` and `html2i(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 KNOWN BUGS

There are no known bugs.

1.7 OPEN ISSUES

There are no open issues.

1.8 AVAILABILITY

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

1.9 AUTHOR

Copyright (c) 2010 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>.

1.10 SEE ALSO

[i\(1\)](#), [i2i\(1\)](#), [inbr\(1\)](#), [i2latex\(1\)](#), [i2man\(1\)](#)

2 Items

2.1 Heading items

Heading items are enclosed by `<h3>` and `</h3>` tags in *HTML*.

2.1.1 *HTML* Syntax

```
<h3>...</h3>
```

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

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

2.1.3 *HTML* example

```
<html><head></head><body><h3>#1 Heading one</h3>
<h3>#1.1 Heading two</h3>
<h3>#1.1.1 Heading three</h3>
<h3>#1.1.1.1 Heading four</h3>
<h3>Heading five</h3>
<p></body></html>
```

2.2 Caption items

Table caption items are enclosed by `<caption>` and `</caption>` tags in *HTML*. Figure caption items are tackled the same way in *HTML* as heading items (see [2.1](#)).

2.2.1 *HTML* Syntax

```
<caption>...</caption>
```

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

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

2.2.3 *HTML* example

```
<html><head></head><body>  
<table border="1"><caption>#T1 Table caption</caption>  
</table>  
<h3>#F1 Figure caption</h3>  
<p></body></html>
```

2.3 Bibliography and equation items

Bibliography and equation items are tackled the same way in *HTML* as heading items (see 2.1).

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 *HTML* example

```
<html><head></head><body>  
<h3>#B1 Hemingway, Ernest: The old man and the sea. 1952.</h3>  
<h3>#B2 Orwell, George: Animal farm. 1945</h3>  
<h3>#E1 Equation</h3>  
<p></body></html>
```

2.4 Paragraph items

Paragraph items start with `<p>` or `
` tags in *HTML*.

2.4.1 *HTML* Syntax

```
<p>  
<br>
```

2.4.2 *i*(1) example

```
First paragraph  
Second paragraph
```

```
Third paragraph
```

2.4.3 *HTML* example

```
<html><head></head><body>  
First paragraph  
<br>Second paragraph  
<p>Third paragraph  
<p></body></html>
```

2.5 List items

List items are enclosed by `` and `` tags in *HTML*. Subsequent list items are grouped in clusters and enclosed by `` and `` tags in *HTML*. List items can also be nested, whereas the maximum nesting level is two. The outermost list item can contain `
` tags and pre-formatted items.

2.5.1 *HTML* Syntax

```
<ul>...</ul>
<li>...</li>
```

2.5.2 *i(1)* example

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

2.5.3 *HTML* example

```
<html><head></head><body>
<ul><li>one</li>
<li>two</li>
<li>three
<br>four
<ul><li>five</li>
<li>six</li></ul></li>
<li>seven
<pre> eight  nine</pre></li></ul>
<p></body></html>
```

2.6 Quotation items

Quotation items are enclosed by `<blockquote>` and `</blockquote>` tags in *HTML*.

2.6.1 *HTML* Syntax

```
<blockquote>...</blockquote>
```

2.6.2 *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.
```

2.6.3 *HTML* example

```
<html><head></head><body>
Ernest Hemingway opened his book with the following words...
<blockquote>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.</blockquote>
<p></body></html>
```

2.7 Pre-formatted items

Pre-formatted items are enclosed by `<pre>` and `</pre>` tags in *HTML*. At the beginning of every line, there is an additional space character.

2.7.1 *HTML* Syntax

```
<pre>...</pre>
```

2.7.2 *i(1)* example

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

2.7.3 *HTML* example

```
<html><head></head><body>  
<pre> for( i = 0; i &lt; k; i++ ) {  
    printf("&quot;i=%d\n&quot;;, i);  
}</pre>  
<p></body></html>
```

2.8 Picture items

Picture items are enclosed by `<pre title="pic">` and `</pre>` tags in *HTML*. At the beginning of every line, there is an additional space character.

2.8.1 *HTML* Syntax

```
<pre title="pic">...</pre>
```

2.8.2 *i(1)* example

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

2.8.3 *HTML* example

```
<html><head></head><body>  
<pre title="pic"> +-----+  
|       |  
+-----+</pre>  
<p></body></html>
```

2.9 Latex items

Latex items are enclosed by `<pre title="latex">` and `</pre>` tags in *HTML*. At the beginning of every line, there is an additional space character.

2.9.1 *HTML* Syntax

```
<pre title="latex">...</pre>
```

2.9.2 *i(1)* example

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

2.9.3 *HTML* example

```
<html><head></head><body>
<h3>#F1 Photo</h3>
<!--intr-->
<pre title="latex"> }\includegraphics[scale=0.25]
  {photo.jpg}\end{ifigure}</pre>
<pre title="man"> [ photo.jpg ]
  .fi</pre>
<p></body></html>
```

2.10 Man items

Man items are enclosed by `<pre title="man">` and `</pre>` tags in *HTML*. At the beginning of every line, there is an additional space character.

2.10.1 *HTML* Syntax

```
<pre title="man">...</pre>
```

2.10.2 *i(1)* example

Refer to chapter 2.9.2.

2.10.3 *HTML* example

Refer to chapter 2.9.3.

2.11 Interrupt items

Interrupt tokens are `<!--intr-->` tags in *HTML*.

2.11.1 *HTML* Syntax

```
<!--intr-->
```

2.11.2 *i(1)* example

Refer to chapter 2.9.2.

2.11.3 *HTML* example

Refer to chapter 2.9.3.

2.12 Footnotes

Footnotes start with ``(`` and end with ``)`` in *HTML*.

2.12.1 *HTML* Syntax

```
<b></b>...<b></b>
```

2.12.2 *i(1)* example

```
    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. #(opening of #B1#)
```

2.12.3 *HTML* example

```
<html><head></head><body>
<blockquote>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. <b></b>opening of #B1<b></b></blockquote>
<p></body></html>
```

3 Modes

The byte mode is not relevant for the conversion from *i(1)* to *HTML* and vice versa. The byte adaption depends on whether option `-u` is set or not:

- Without option `-u`, the byte adaption is done according to tables 1 and 2.
- With option `-u`, the byte adaption for characters 0x01...0x7F is done according to table 1. The other characters contain more than one byte in *UTF-8* and are adapted according to their unicode scalar value.

Table 1: Byte adaption (characters 0x01...0x7F)

<i>i(1)</i>	<i>HTML</i>	
Hex	ISO-8859-1 and UTF-8 representation	Tag
0x01		
0x02		
0x03		
0x04		
0x05		
0x06		
0x07		
0x08		
0x09	tabulator	
0x0A	new line	
0x0B		
0x0C		
0x0D	carriage return	
0x0E		
0x0F		
0x10		
0x11		
0x12		
0x13		
0x14		
0x15		
0x16		
0x17		
0x18		
0x19		
0x1A		
0x1B		
0x1C		
0x1D		
0x1E		

i(1)		HTML
Hex	<i>ISO-8859-1</i> and <i>UTF-8</i> representation	Tag
0x1F		
0x20	space	
0x21	!	!
0x22	"	";
0x23	#	#
0x24	\$	\$
0x25	%	%
0x26	&	&;
0x27	'	'
0x28	((
0x29))
0x2A	*	*
0x2B	+	+
0x2C	,	,
0x2D	-	-
0x2E	.	.
0x2F	/	/
0x30	0	0
0x31	1	1
0x32	2	2
0x33	3	3
0x34	4	4
0x35	5	5
0x36	6	6
0x37	7	7
0x38	8	8
0x39	9	9
0x3A	:	:
0x3B	;	;
0x3C	<	<;
0x3D	=	=
0x3E	>	>;
0x3F	?	?
0x40	@	@
0x41	A	A
0x42	B	B
0x43	C	C
0x44	D	D
0x45	E	E
0x46	F	F
0x47	G	G
0x48	H	H
0x49	I	I
0x4A	J	J
0x4B	K	K
0x4C	L	L
0x4D	M	M

i(1)		HTML
Hex	<i>ISO-8859-1</i> and <i>UTF-8</i> representation	Tag
0x4E	N	N
0x4F	O	O
0x50	P	P
0x51	Q	Q
0x52	R	R
0x53	S	S
0x54	T	T
0x55	U	U
0x56	V	V
0x57	W	W
0x58	X	X
0x59	Y	Y
0x5A	Z	Z
0x5B	[[
0x5C	\	\
0x5D]]
0x5E	^	^
0x5F	_	_
0x60	`	`
0x61	a	a
0x62	b	b
0x63	c	c
0x64	d	d
0x65	e	e
0x66	f	f
0x67	g	g
0x68	h	h
0x69	i	i
0x6A	j	j
0x6B	k	k
0x6C	l	l
0x6D	m	m
0x6E	n	n
0x6F	o	o
0x70	p	p
0x71	q	q
0x72	r	r
0x73	s	s
0x74	t	t
0x75	u	u
0x76	v	v
0x77	w	w
0x78	x	x
0x79	y	y
0x7A	z	z
0x7B	{	{
0x7C		

<i>i</i> (1)		<i>HTML</i>
Hex	<i>ISO-8859-1</i> and <i>UTF-8</i> representation	Tag
0x7D	}	}
0x7E	~	~
0x7F		

Table 2: Byte adaption (without option -u characters 0x80...0xFF)

<i>i</i> (1)		<i>HTML</i>
Hex	<i>ISO-8859-1</i> representation	Tag
0x80		€
0x81		
0x82		‚
0x83		ƒ
0x84		„
0x85		…
0x86		†
0x87		‡
0x88		ˆ
0x89		‰
0x8A		Š
0x8B		‹
0x8C		Œ
0x8D		
0x8E		Ž
0x8F		
0x90		
0x91		‘
0x92		’
0x93		“
0x94		”
0x95		•
0x96		–
0x97		—
0x98		˜
0x99		™
0x9A		š
0x9B		›
0x9C		œ
0x9D		
0x9E		ž
0x9F		Ÿ
0xA0		
0xA1	ı	¡
0xA2	ç	¢
0xA3	£	£
0xA4	ı	¤

i(1)		<i>HTML</i>
Hex	<i>ISO-8859-1</i> representation	Tag
0xA5	¥	¥
0xA6	¦	¦
0xA7	§	§
0xA8	¨	¨
0xA9	©	©
0xAA	ª	ª
0xAB	«	«
0xAC	¬	¬
0xAD		­
0xAE	®	®
0xAF	¯	¯
0xB0	°	°
0xB1	º	±
0xB2	²	²
0xB3	³	³
0xB4	´	´
0xB5	µ	µ
0xB6	¶	¶
0xB7	·	·
0xB8	¸	¸
0xB9	¹	¹
0xBA	º	º
0xBB	»	»
0xBC	¼	¼
0xBD	½	½
0xBE	¾	¾
0xBF	¿	¿
0xC0	À	À
0xC1	Á	Á
0xC2	Â	Â
0xC3	Ã	Ã
0xC4	Ä	Ä
0xC5	Å	Å
0xC6	Æ	Æ
0xC7	Ç	Ç
0xC8	È	È
0xC9	É	É
0xCA	Ê	Ê
0xCB	Ë	Ë
0xCC	Ì	Ì
0xCD	Í	Í
0xCE	Î	Î
0xCF	Ï	Ï
0xD0	Ð	Ð
0xD1	Ñ	Ñ
0xD2	Ò	Ò
0xD3	Ó	Ó

i(1)		<i>HTML</i>
Hex	<i>ISO-8859-1</i> representation	Tag
0xD4	ō	Ô
0xD5	õ	Õ
0xD6	ö	Ö
0xD7	×	×
0xD8	∅	Ø
0xD9	Û	Ù
0xDA	Ū	Ú
0xDB	Ů	Û
0xDC	Û	Ü
0xDD	Ÿ	Ý
0xDE	Ɔ	Þ
0xDF	ƒ	ß
0xE0	à	à
0xE1	á	á
0xE2	â	â
0xE3	ã	ã
0xE4	ä	ä
0xE5	å	å
0xE6	æ	æ
0xE7	ç	ç
0xE8	è	è
0xE9	é	é
0xEA	ê	ê
0xEB	ë	ë
0xEC	ì	ì
0xED	í	í
0xEE	î	î
0xEF	ï	ï
0xF0	ð	ð
0xF1	ñ	ñ
0xF2	ò	ò
0xF3	ó	ó
0xF4	ô	ô
0xF5	õ	õ
0xF6	ö	ö
0xF7	÷	÷
0xF8	ø	ø
0xF9	ù	ù
0xFA	ú	ú
0xFB	û	û
0xFC	ü	ü
0xFD	ý	ý
0xFE	þ	þ
0xFF	ÿ	ÿ

4 Arrangement

4.1 One-dimensional arrangement

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

In *HTML*, tables are enclosed by `<table border="1">` and `</table>` tags and columns are enclosed by `<td>` and `</td>` tags. The `colspan` attribute in the `<td>` tag indicates the number of columns it spans. Rows are separated from each other by `<tr>` tags, whereas the content of the `I_TOK_HLINE` token is assigned to the `title` attribute.

4.2.1 *HTML* Syntax

```
<table border="1">...</table>
<tr>
<td>...</td>
```

4.2.2 *i(1)* example of a table without a head

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

4.2.3 *HTML* example of a table without a head

```
<html><head></head><body><table border="1">
<tr title=":=====:=====:=====">
<td colspan="1">one</td>
<td colspan="1">two</td>
<td colspan="1">three</td>
<tr title=":-----:-----:-----">
<td colspan="1">four</td>
<td colspan="2">five</td>
</table>
<p></body></html>
```

4.2.4 *i(1)* example of a table with a head

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

4.2.5 *HTML* example of a table with a head

```
<html><head></head><body><table border="1">
<tr title=":-----:-----:-----">
<td colspan="1">one</td>
<td colspan="1">two</td>
<td colspan="1">three</td>
<tr title=":=====:=====.=====">
<td colspan="1">four</td>
<td colspan="2">five</td>
</table>
<p></body></html>
```

4.2.6 *i*(1) example of a table with various items

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

4.2.7 *HTML* example of a table with various items

```
<html><head></head><body><table border="1">
<tr title=":-----:-----:-----">
<td colspan="1">one</td>
<td colspan="1">two</td>
<td colspan="1">three</td>
<tr title=":=====:=====.=====">
<td colspan="1">
pre
<ul><li>list</li>
<li>LIST</li></ul></td>
<td colspan="2">
PRE
<blockquote>quote</blockquote></td>
<tr title=":-----:-----:-----">
<td colspan="1">
<ul><li>list</li>
<li>LIST</li></ul>
post</td>
<td colspan="2">
<blockquote>quote</blockquote>
POST</td>
</table>
<p></body></html>
```

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.

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