Fonts For Forensics

May 20, 2010

Dr. Fred Cohen
President - California Sciences Institute
CEO – Fred Cohen & Associates

Forensic Fonts™

- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?
  - Which are continuation lines?

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?
  - Which are continuation lines?
  - Are those TAB or SPACE characters?

Demonstration

- A few questions about an exhibit from a real case
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?
  - Which are continuation lines?
  - Are those TAB or SPACE characters?
  - Is this a forgery? How can you tell?

Demonstration

- Here is an additional exhibit not available at trial... can you answer the same questions?
  - What does the 1st line of this document say?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?
  - Which are continuation lines?
  - Are those TAB or SPACE characters?
  - Is this a forgery? How can you tell?

The problem in words

- Based on the exhibits we know how to make, we cannot easily (or at all) tell or show others what is actually there.
  - The typical depiction is not accurate or precise as to what is present
  - A depiction that is accurate and precise is nearly unusable and certainly unclear
  - And this is only a text file! (or is it?)

Forensic Fonts™

- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

Exhibit F

- Examine the same content using a different visualization:
  - What does the 1st line of this document contain?
  - Is this a trace of an authentic electronic mail message?
  - Where is the end of the header?
  - Which are continuation lines?
  - Are those TAB or SPACE characters?
  - Is this a forgery? How can you tell?

Forensic Fonts™

- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

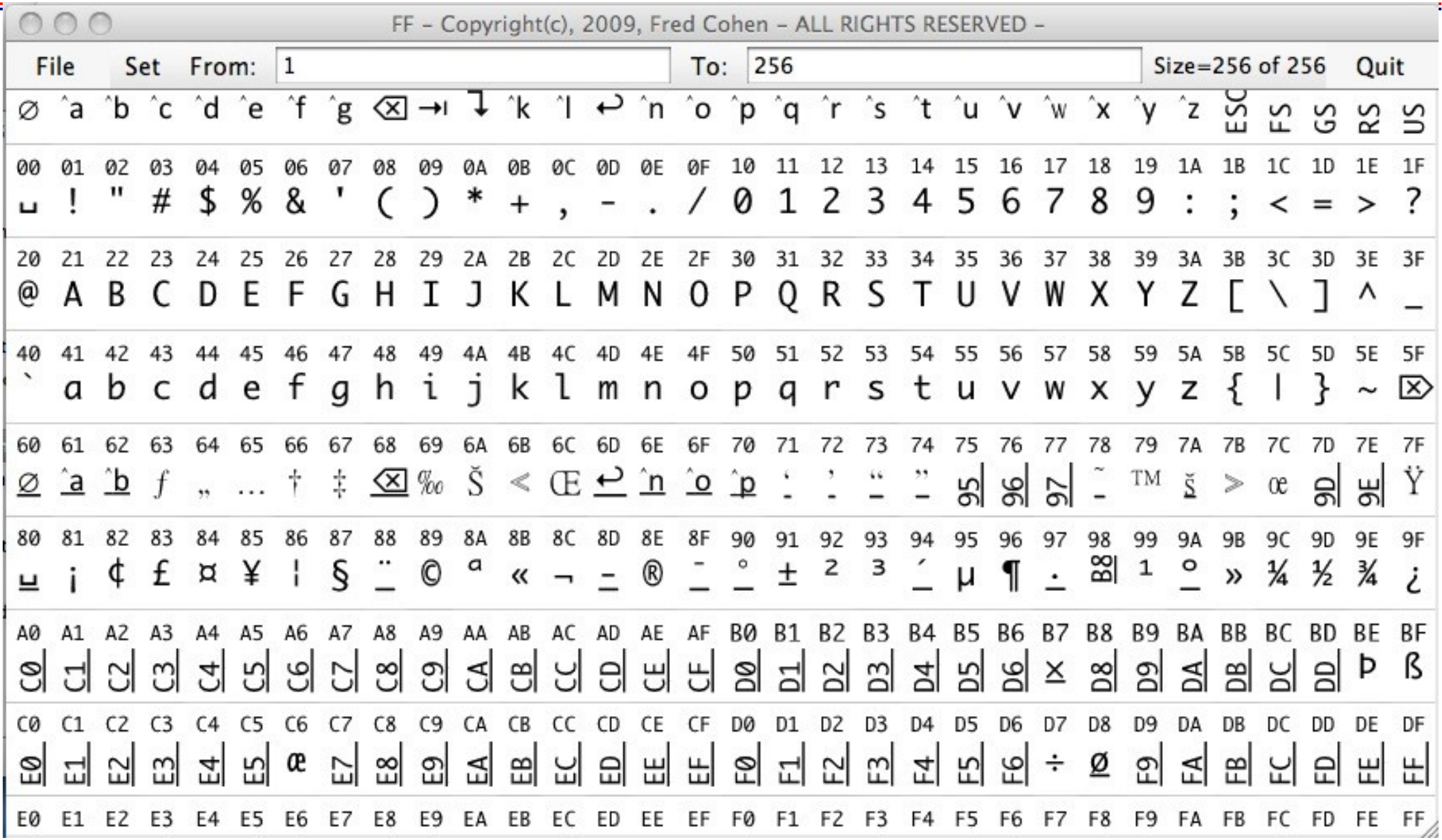
Desirable Properties

- Each symbol must be clearly different and readily distinguishable from all other symbols
- Each symbol must be displayable and printable so that a <space>, <tab>, <carriage-return>, <backspace>, <escape>, and other "non-printable" characters can be clearly seen on the printed page and on other displays.

Desirable Properties

- Each symbol should be familiar, with minimal added interpretation, so that it looks similar to what might appear on a display of the same symbol on a screen or printer in normal use.
- Each symbol should be able to be depicted so as to self-indicate the underlying bit pattern that produced it, so that it can be traced back to its original value.
- Each symbol should be depictable in the same width and height

A Forensic Font for ASCII



Another example

```
FF>diff test1 test2
1,4c1,4
< This is a test
< This is another test
< This is a different test
< This is still another test
---
> This is a test
> This is another test
> This is a different test
> This is still another test
FF>diff test1 test2 | ff
```

Still another example A database displayed with FF and width alignment

FF – Copyright(c), 2009, Fred Cohen – ALL RIGHTS RESERVED –

Set From: 37 To: 300 go Quit

^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^a	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^b
01	2F	00	00	EB	00	1F	00	08	00	00	01	01	2F	00	00	EB	00	1F	00	08	00	00	02
^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^c	^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^d
01	2F	00	00	DB	00	1F	00	08	00	00	03	01	2F	00	00	DB	00	1F	00	08	00	00	04
^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^e	^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^f
01	2F	00	00	DB	00	1F	00	08	00	00	05	01	2F	00	00	DB	00	1F	00	08	00	00	06
^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^g	^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	⌫
01	2F	00	00	DB	00	1F	00	08	00	00	07	01	2F	00	00	DB	00	1F	00	08	00	00	08
^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	→|	^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	↴
01	2F	00	00	DB	00	1F	00	08	00	00	09	01	2F	00	00	DB	00	1F	00	08	00	00	0A
^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^k	^a	/	∅	∅	DB	∅	US	∅	⌫	∅	∅	^l
01	2F	00	00	DB	00	1F	00	08	00	00	0B	01	2F	00	00	DB	00	1F	00	08	00	00	0C
^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	↩	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^n
01	2F	00	00	EB	00	1F	00	08	00	00	0D	01	2F	00	00	EB	00	1F	00	08	00	00	0E
^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^o	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^p
01	2F	00	00	EB	00	1F	00	08	00	00	0F	01	2F	00	00	EB	00	1F	00	08	00	00	10
^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^q	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^r
01	2F	00	00	EB	00	1F	00	08	00	00	11	01	2F	00	00	EB	00	1F	00	08	00	00	12
^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^s	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^t
01	2F	00	00	EB	00	1F	00	08	00	00	13	01	2F	00	00	EB	00	1F	00	08	00	00	14
^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^u	^a	/	∅	∅	EB	∅	US	∅	⌫	∅	∅	^v
01	2F	00	00	EB	00	1F	00	08	00	00	15	01	2F	00	00	EB	00	1F	00	08	00	00	16

Forensic Fonts™

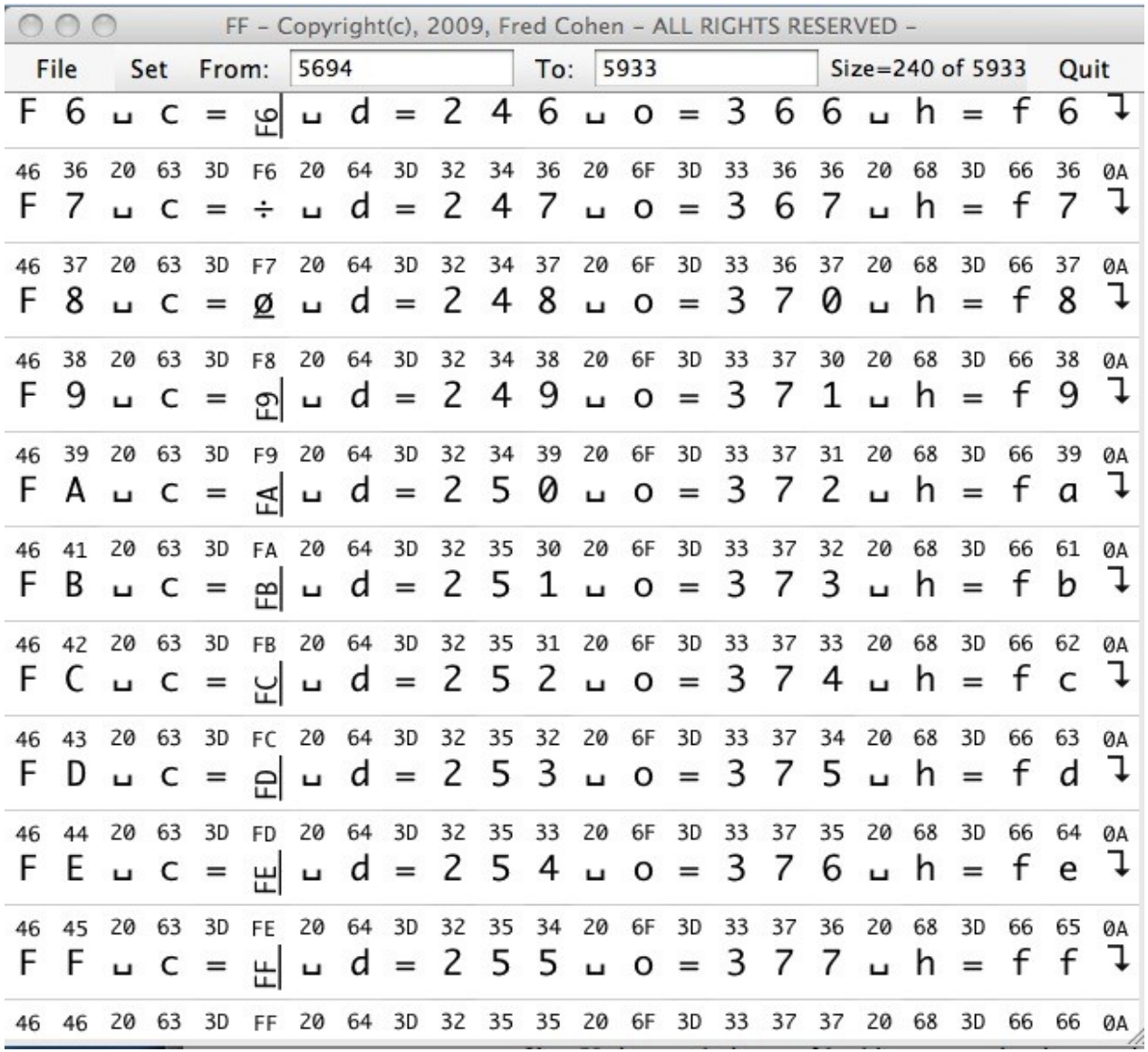
- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

Reliability of the tool

- Key desirable properties
  - The tool must not alter original writing
    - Achieved by controlling the code in the tool
  - Displayed information must be complete
    - Cannot be proven as well as property 1
    - In general, it is impossible (finite space/time)
    - Validation of failure modes is feasible / done
  - Displayed information must be accurate
    - Cannot be proven as well as property 1
    - Impossible to verify all possible sequences
    - But validation tests are available and done

Validation tests

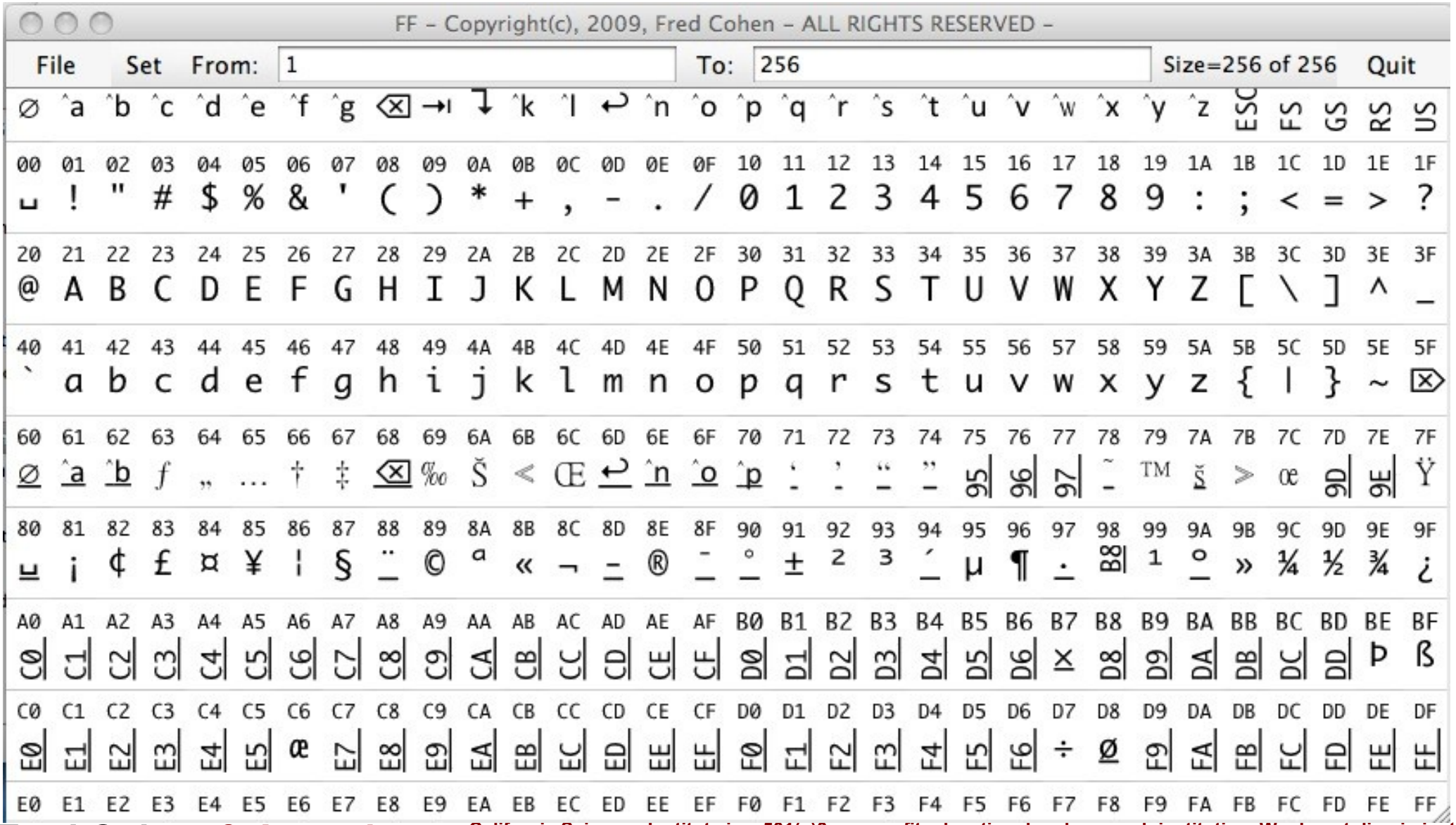
- Generate known patterns
- Run the patterns through FF
- Visually inspect results to verify output matches expected values
- Limited by visual inspection and sequence limits



F	6	␣	c	=	F6	␣	d	=	2	4	6	␣	o	=	3	6	6	␣	h	=	f	6	↵
46	36	20	63	3D	F6	20	64	3D	32	34	36	20	6F	3D	33	36	36	20	68	3D	66	36	0A
F	7	␣	c	=	÷	␣	d	=	2	4	7	␣	o	=	3	6	7	␣	h	=	f	7	↵
46	37	20	63	3D	F7	20	64	3D	32	34	37	20	6F	3D	33	36	37	20	68	3D	66	37	0A
F	8	␣	c	=	Ø	␣	d	=	2	4	8	␣	o	=	3	7	0	␣	h	=	f	8	↵
46	38	20	63	3D	F8	20	64	3D	32	34	38	20	6F	3D	33	37	30	20	68	3D	66	38	0A
F	9	␣	c	=	F9	␣	d	=	2	4	9	␣	o	=	3	7	1	␣	h	=	f	9	↵
46	39	20	63	3D	F9	20	64	3D	32	34	39	20	6F	3D	33	37	31	20	68	3D	66	39	0A
F	A	␣	c	=	FA	␣	d	=	2	5	0	␣	o	=	3	7	2	␣	h	=	f	a	↵
46	41	20	63	3D	FA	20	64	3D	32	35	30	20	6F	3D	33	37	32	20	68	3D	66	61	0A
F	B	␣	c	=	FB	␣	d	=	2	5	1	␣	o	=	3	7	3	␣	h	=	f	b	↵
46	42	20	63	3D	FB	20	64	3D	32	35	31	20	6F	3D	33	37	33	20	68	3D	66	62	0A
F	C	␣	c	=	FC	␣	d	=	2	5	2	␣	o	=	3	7	4	␣	h	=	f	c	↵
46	43	20	63	3D	FC	20	64	3D	32	35	32	20	6F	3D	33	37	34	20	68	3D	66	63	0A
F	D	␣	c	=	FD	␣	d	=	2	5	3	␣	o	=	3	7	5	␣	h	=	f	d	↵
46	44	20	63	3D	FD	20	64	3D	32	35	33	20	6F	3D	33	37	35	20	68	3D	66	64	0A
F	E	␣	c	=	FE	␣	d	=	2	5	4	␣	o	=	3	7	6	␣	h	=	f	e	↵
46	45	20	63	3D	FE	20	64	3D	32	35	34	20	6F	3D	33	37	36	20	68	3D	66	65	0A
F	F	␣	c	=	FF	␣	d	=	2	5	5	␣	o	=	3	7	7	␣	h	=	f	f	↵
46	46	20	63	3D	FF	20	64	3D	32	35	35	20	6F	3D	33	37	37	20	68	3D	66	66	0A

Validation tests 2

- main(){char i;int j;for (j=0;j<256;j++) {i=j;write(1,&j,1);}



Provenance and pointers

- Provenance and location information is also helpful in actual output

Created by fc on or about 2010-04-12@10:18:01.812

Forensic Font(TM) - Patent Pending is Copyright(c), 2009-10, Fred Cohen - ALL RIGHTS RESERVED

Created by fc (/Users/fc) from /Users/fc/src/java/FF/FF-FontGenerator with:

FF -ff ASCII -f F -i "-" -o "./-" -EOL 0A -W 21 -H 40 -BG "gray" -B T -T T -C T -LL 0 -TL 4 -from 1 -to -1

on or about 2010-04-12@10:18:01.812 in Mac OS X (v 10.5.8 - i386)

0	T	h	i	s	␣	i	s	␣	a	␣	t	e	s	t	↩	↴			
	54g	68g	69g	73g	20g	69g	73g	20g	61g	20g	74g	65g	73g	74g	0Dg	0Ag			
16	␣	→|	o	f	␣	t	h	e	␣	p	r	o	g	r	a	m	␣	→|	↴
	20g	09g	6Fg	66g	20g	74g	68g	65g	20g	70g	72g	6Fg	67g	72g	61g	6Dg	20g	09g	0Ag
35																			

End of -

Forensic Fonts™

- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

Describing the depictions

- Using the FF tool, which is the subject of a peer reviewed paper reliable depiction of forensic evidence in cases such as this[1], and that I have tested and used for some time and found to be reliable for the purposes at issue in this case, I generated the depiction found in exhibit [ITEM].
- This depiction starts with and otherwise contains information on the method I used to depict this exhibit, including the date, time, filename, and other related details that I will call provenance information.

Describing the depictions

- In the area depicting content, both the symbols normally printed on a screen or printer, and symbols that are present in traces but not normally depicted or differentiable are displayed, in the sequence they appears in [original writing].
- Each row starts with a number indicating the offset into the file containing [original writing] at which the sequence appears, starting with offset “0” to indicate the beginning of the file.

Describing the depictions

- In examining the content, I found that it was consistent with the “ASCII” character set and that it would be best understood for the purposes of my testimony by separating it into “lines”, each of which you can see, ends with a “carriage return” character followed by a “line feed character”. [or whatever other determination was used to present as depicted]

Forensic Fonts™

- What's the problem?
- What's the solution?
- Related properties and issues
- How reliable is it?
- How do we say it?
- Conclusions

One thing to ask yourself

- When on the stand, would you rather have the forensic font depiction to base your testimony on or not?
- I have started to use it in court and in examinations so that I can be certain of what I am looking at
  - Regardless of how they manged to prepare (read mangle) the depiction of traces
  - Especially when I am not certain of what I am looking at or seeing

Broader implications

- Clearly, we have a long way to go in understanding how to depict traces
  - Something as clarifying as this has been needed for a long time
  - Dealing with other symbols sets (ROT13, EBCDIC, etc.) is just the start (FF does it)
  - Depicting with direct traceback to original parts of traces and in a manner that can be printed for others to see is vital
  - The broader area of forensically sound and reliable visualization is clearly key to progress in digital forensics

Thank You

http://calsci.org/ - calsci at calsci.org
http://all.net/ - fc at all.net