

一般教育演習

プログラミングで問題を解く：

集計から人工知能まで

瀧川 一学

工学部 情報理工学コース

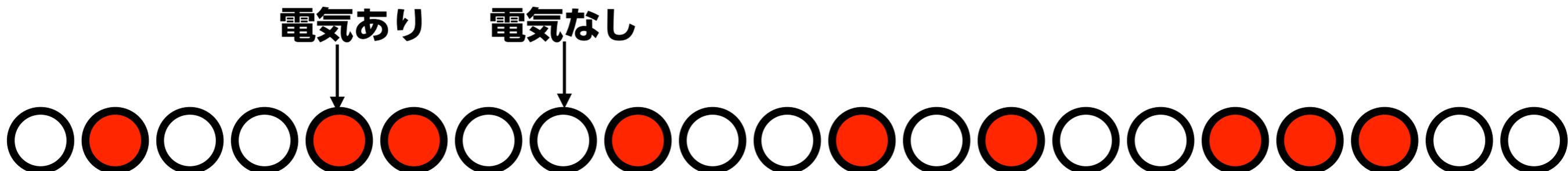
今日のお題：Pythonの基本

- フォロアーアップQ & A：
 - デジタル表現と文字化けについて
- Notebookをダウンロードして実行してみる
- Python言語の基本を学ぶ
 - 前回の復習 + 関数 + リスト + PEP8
- ミニレポートタイム

0と1の世界

コンピュータはプログラムもデータも全て
「0と1で表現」して保持します。

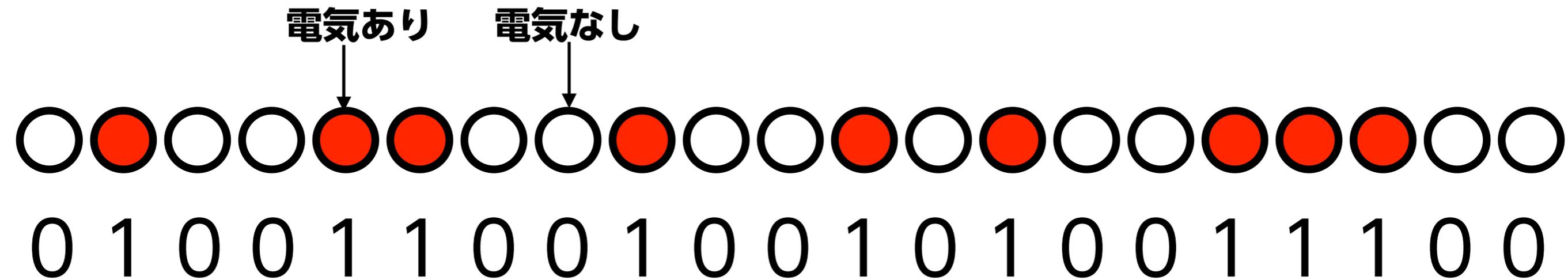
「0と1で表現」って何??



物理的には情報を記憶しておく領域(メモリ、ハードディスク、レジスタ)は
とにかく、素子が並んでいて、電気(電圧)のありなしを制御している

0と1の世界

「0と1で表現」って何??



この状態パターンをとりあえず0と1で表しましょうというだけ
(本質的には電気じゃなくても基石でも赤白カードでも何でもよい)

0と1の世界

数字の場合は基本的には**二進数**にする

0	00000	9	01001
1	00001	10	01010
2	00010	11	01011
3	00011	12	01100
4	00100	13	01101
5	00101	14	01110
6	00110	15	01111
7	00111	16	10000
8	01000	17	10001

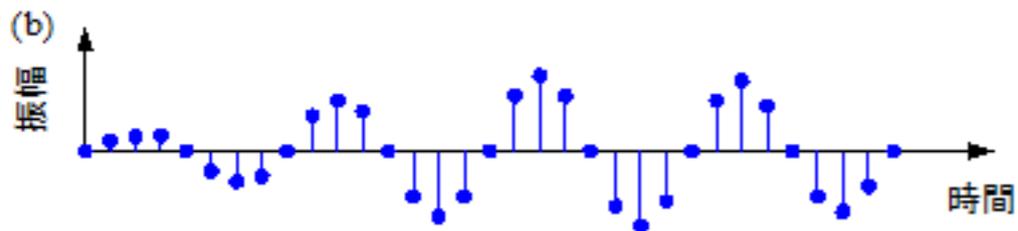
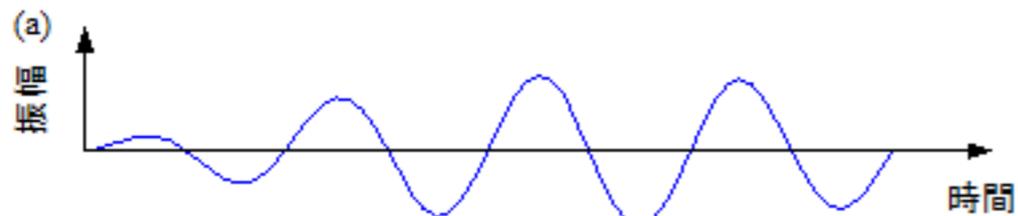
0と1の世界

2進数(b)は長くなるので16進表記(0x)をよく使う

(1) 0~15までの値			(2) 16以上の代表値		
10進数	16進数	2進数	10進数	16進数	2進数
0	= 0x0	= 0b	16	= 0x10	= 1 0000b
1	= 0x1	= 1b	32	= 0x20	= 10 0000b
2	= 0x2	= 10b	64	= 0x40	= 100 0000b
3	= 0x3	= 11b	128	= 0x80	= 1000 0000b
4	= 0x4	= 100b	255	= 0xFF	= 1111 1111b
5	= 0x5	= 101b	256	= 0x100	= 1 0000 0000b
6	= 0x6	= 110b	512	= 0x200	= 10 0000 0000b
7	= 0x7	= 111b	1024	= 0x400	= 100 0000 0000b (1K)
8	= 0x8	= 1000b	65535	= 0xFFFF	= 1111 1111 1111 1111b
9	= 0x9	= 1001b	65536	= 0x10000	= 1 0000 0000 0000 0000b (64K)
10	= 0xA	= 1010b	1048576	= 0x100000	= 1 0000 0000 0000 0000 0000b (1M)
11	= 0xB	= 1011b			
12	= 0xC	= 1100b			
13	= 0xD	= 1101b			
14	= 0xE	= 1110b			
15	= 0xF	= 1111b			

その1：音のファイル

デジタル化する



CD音質なら44.1kHz
=1秒間に441000点



数字になるのではありません
それをただ並べる

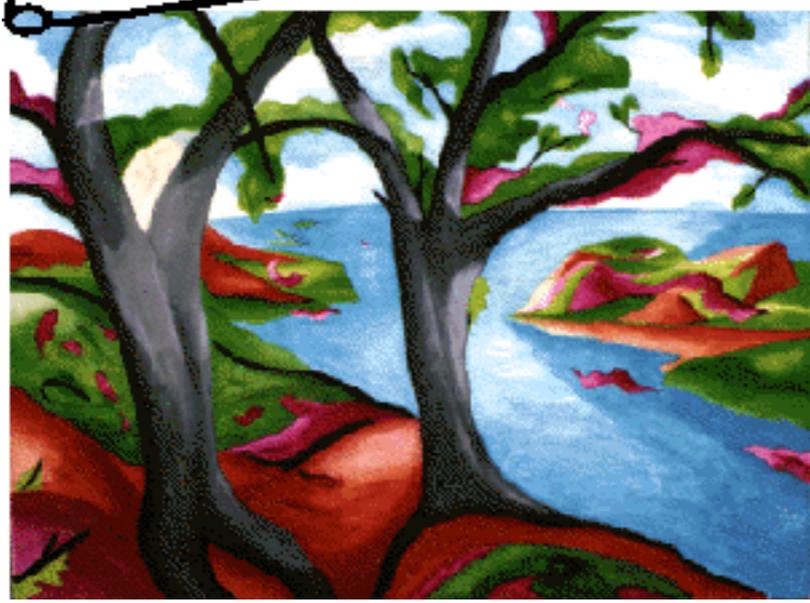
その2 : 画像のファイル

0.2235	0.1294	Blue	0.4196	0.2235	0.2588	0.2588
0.5804	0.2902	0.0627	0.2902	0.2902	0.4824	0.0627
0.5804	0.0627	0.0627	0.0627	0.2235	0.2588	0.2588
0.5176	0.1922	0.0627	Green	0.1922	0.2588	0.2588
0.5176	0.1294	0.1608	0.1294	0.1294	0.2588	0.2588
0.5176	0.1608	0.0627	0.1608	0.1922	0.2588	0.2588
0.5490	0.2235	0.5490	Red	0.1412	0.1765	0.1765
0.5490	0.3882	0.5176	0.5804	0.5804	0.1765	0.1765
0.2588	0.2902	0.2588	0.2235	0.4824	0.2235	0.2235
0.2235	0.1608	0.2588	0.2588	0.1608	0.2588	0.2588
0.2588	0.1608	0.2588	0.2588	0.2588	0.2588	0.2588

各ピクセルは例えば
赤緑青の光の三原色
の割合の数字



数字になるのであとは
それをただ並べる



その3 : 文字(テキスト)

各々の文字に番号が決まっているだけ

10進数	16進数	2進数	文字
65	41	1000001	A
66	42	1000010	B
67	43	1000011	C
68	44	1000100	D
69	45	1000101	E
70	46	1000110	F
71	47	1000111	G
72	48	1001000	H
73	49	1001001	I
74	4A	1001010	J
75	4B	1001011	K
76	4C	1001100	L

その3：文字(テキスト)

日本語特有の問題「文字化け」

各文字の番号の割付ルールが複数ある
(間違えて解釈すると読めない=文字化け)

- UTF-8
- Shift JIS
- ISO-2022-JP
- EUC

その3 : 文字(テキスト)

2進数で書くと長くなるので16進表記をよく使う

■文字集合と文字コード

愛

文字集合 JIS X 0208
では 16区 6点 の位
置に存在する文字

文字集合 Unicode で
は、U+611B の位置に
存在する文字

ISO-2022-JP(いわゆる
JISコード)で表すと、
0x3026

EUC-JPで表すと、
0xB0A6

Shift_JISで表すと、
0x88A4

UTF-16で表すと、
0x611B

UTF-8で表すと、
0xE6849B

参考) [\[byte_codes.ipynb\]](#)

▼ 1 コンピュータ内部の表現について

▼ 1.1 漢字/日本語

```
In [3]: s = '愛'  
        bytes = s.encode('utf-8')
```

```
In [4]: bytes
```

```
Out[4]: b'\xe6\x84\x9b'
```

16進で表記

```
In [13]: print(*["{:x}".format(x) for x in bytes], sep='')
```

e6849b

(長くなるけど)敢えて2進で書けば

```
In [14]: print(*["{:b}".format(x) for x in bytes], sep='')
```

111001101000010010011011

作業：授業のホームページから

「[practice02.ipynb](#)」をダウンロードして実行してみよう。

<http://art.ist.hokudai.ac.jp/~takigawa/prog/>

Pythonが起動しているフォルダにコピーし、
クリックして開く！

 jupyter

Logout

Files

Running

Clusters

Nbextensions

Select items to perform actions on them.

Upload

New ▾



0



Name ↓

Last Modified

files

7 days ago

python

7 days ago

 myfirstcode.ipynb

Running seconds ago

今日のお題：Pythonの基本

- フォロアーアップQ & A：
 - デジタル表現と文字化けについて
- Notebookをダウンロードして実行してみる
- Python言語の基本を学ぶ
 - 前回の復習 + 関数 + リスト + PEP8
- ミニレポートタイム