

Compact formats for storing directed graphs

Michael Torpey

University of St Andrews

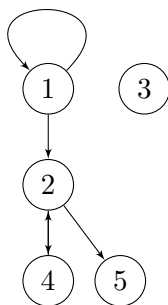
2015-01-28

Representing digraphs

How do we represent this digraph?

- ▶ Edge set:
 $\{(1, 1), (1, 2), (2, 4), (2, 5), (4, 2)\}$
- ▶ Adjacencies:
 $(\{1, 2\}, \{4, 5\}, \{ \}, \{2\}, \{ \})$
- ▶ Adjacency matrix:

	1	2	3	4	5
1	✓	✓			
2				✓	✓
3					
4		✓			
5					



Storing as text

- ▶ Edge set: $\{(1, 1), (1, 2), (2, 4), (2, 5), (4, 2)\}$
- ▶ Adjacencies list: $(\{1, 2\}, \{4, 5\}, \{ \}, \{2\}, \{ \})$
- ▶ Adjacency matrix

Print the edge set:

1>1_
1>2_
2>4_
2>5_
4>2

(19 characters)
($\sim 4e$)

Print the adjacencies:

1,2_
4,5_
_
2_

(11 characters)
($\sim 2e + v$)

Print a matrix:

11000
00011
00000
01000
00000

(25 characters)
($\sim v^2$)

The matrix approach: *Digraph6*

1	1	0	0	0
0	0	0	1	1
0	0	0	0	0
0	1	0	0	0
0	0	0	0	0

- ▶ 10000 10010 00000 01000 01000
- ▶ 100001 001000 000010 000100 000000
- ▶ 33 8 2 4 0
- ▶ 96 71 65 67 63
- ▶ ' G A C ?
- ▶ +D'GAC? (7 characters)
- ▶ Always characters in the set:
?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^
_`abcdefghijklmnopqrstuvwxyz{|}~

The adjacencies/edge-set approach: *Disparse6*

- ▶ Edge set: $\{(1, 1), (1, 2), (2, 4), (2, 5), (4, 2)\}$
- ▶ (1,3) (0,0) (0,1) (1,3) (1,4)
- ▶ 1 3 0 1 1 5 0 0 1 0 1 3 0 1 1 1
- ▶ 1 011 0 001 1 101 0 000 1 000 1 011 0 001 1 001
- ▶ 101100 011101 000010 001011 000110 011111
- ▶ 44 29 2 11 6 31
- ▶ 107 92 65 74 69 94
- ▶ k \ A J E ^
- ▶ .Dk\AJE^ (9 characters)

Choosing a representation

- ▶ *Digraph6* gives a string of $\sim v^2$ bits.
- ▶ *Disparse6* gives a string of $\sim \frac{3}{2}e(\lceil \log_2 v \rceil + 1)$ bits.
- ▶ When writing a graph, calculate both of these and choose the smaller.