

# Generating Random Text

Finding Follow Set

# Markov Models

- Markov-zero model generated random text
  - Characters chosen based on frequencies in training text
- What's needed for Markov-one, -two, ...?
  - Find occurrences of 't', or 'th' in training text
  - Follow characters to generate text at random
- Key algorithm: finding all follow characters
  - Develop for Markov-one, then generalize

# Step 1: Work an Instance

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Use example of training text shown
- Find characters that follow 't'

# Step 1: Work an Instance

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Use example of training text shown
- Find characters that follow 't'
  - 'h'/1,

# Step 1: Work an Instance

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Use example of training text shown
- Find characters that follow 't'
  - 'h'/1,
  - 't'/13,

# Step 1: Work an Instance

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Use example of training text shown
- Find characters that follow 't'
  - 'h'/1,
  - 't'/13,
  - 'e'/14



# Step 1: Work an Instance

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t	
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	

- Use example of training text shown
- Find characters that follow 't'
  - 'h'/1,
  - 't'/13,
  - 'e'/14
  - What about nothing that follows last 't'?

# Step 2: Write Down What You Did

follows:

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty



# Step 2: Write Down What You Did

follows:

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0

# Step 2: Write Down What You Did

follows: 'h'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows

# Step 2: Write Down What You Did

follows: 'h'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1

# Step 2: Write Down What You Did

follows: 'h', 't'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows



# Step 2: Write Down What You Did

**follows:** 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 2: Write Down What You Did

**follows:** 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14



# Step 2: Write Down What You Did

follows: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 2: Write Down What You Did

follows: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 2: Write Down What You Did

`follows`: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 `follows` list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to `follows`
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to `follows`
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to `follows`
- 8 find first occurrence of 't' starting at index 14



# Step 2: Write Down What You Did

`follows`: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 `follows` list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to `follows`
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to `follows`
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to `follows`
- 8 find first occurrence of 't' starting at index 14

# Step 2: Write Down What You Did

`follows`: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 `follows` list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to `follows`
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to `follows`
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to `follows`
- 8 find first occurrence of 't' starting at index 14

# Step 2: Write Down What You Did

**follows:** 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14



# Step 2: Write Down What You Did

**follows:** 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14



# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

**follows:** 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14



# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14



# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

follows: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- 1 follows list is initially empty
- 2 find first occurrence of 't' starting at index 0
- 3 add letter at index 1 to follows
- 4 find first occurrence of 't' starting at index 1
- 5 add letter at index 13 to follows
- 6 find first occurrence of 't' starting at index 13
- 7 add letter at index 14 to follows
- 8 find first occurrence of 't' starting at index 14

# Step 3: Find Patterns, Generalize

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Start search for 't' at index 0, list empty
- When do we stop loop/search?
  - When we don't find a 't'



# Step 3: Find Patterns, Generalize

follows: 'h', 't', 'e'

t	h	i	s		i	s		a	n		a	t	t	e	m	p	t
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

- Start search for 't' at index 0, list empty
- When do we stop loop/search?
  - When we don't find a 't'
  - When we run off end of text

# Step 4: Test Your Algorithm

KEY = 'a',

<b>a</b>	<b>n</b>		<b>a</b>	<b>p</b>	<b>p</b>	<b>l</b>	<b>e</b>		<b>a</b>		<b>d</b>	<b>a</b>	<b>y</b>
<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>

- A. **follows** is empty ArrayList, **pos** is zero
- B. While there is more searching to do
  1. **index** = first occurrence of KEY starting at **pos**
  2. if found, add value at [**index**+1] to **follows**
  3. **pos** = **index** + 1

# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

A. **follows** is empty ArrayList, **pos** is zero

B. While there is more searching to do

1. **index** = first occurrence of KEY starting at **pos**
2. if found, add value at [**index**+1] to **follows**
3. **pos** = **index** + 1

# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

A. **follows** is empty ArrayList, **pos** is zero

B. While there is more searching to do

1. **index** = first occurrence of KEY starting at **pos**
2. if found, add value at [**index**+1] to **follows**
3. **pos** = **index** + 1



# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

A. **follows** is empty ArrayList, **pos** is zero

B. While there is more searching to do

1. **index** = first occurrence of KEY starting at **pos**

2. if found, add value at [**index**+1] to **follows**

3. **pos** = **index** + 1

# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

- A. **follows** is empty ArrayList, **pos** is zero
- B. While there is more searching to do
  1. **index** = first occurrence of KEY starting at **pos**
  2. if found, add value at [**index**+1] to **follows**
  3. **pos** = **index** + 1

# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

- A. **follows** is empty ArrayList, **pos** is zero
- B. While there is more searching to do
  - 1. **index** = first occurrence of KEY starting at **pos**
  - 2. if found, add value at [**index**+1] to **follows**
  - 3. **pos** = **index** + 1

# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

- A. **follows** is empty ArrayList, **pos** is zero
- B. While there is more searching to do
  1. **index** = first occurrence of KEY starting at **pos**
  2. if found, add value at [**index**+1] to **follows**
  3. **pos** = **index** + 1
- C. Your answer is **follows**



# Step 4: Test Your Algorithm

KEY = 'a',

a	n		a	p	p	l	e		a		d	a	y
0	1	2	3	4	5	6	7	8	9	10	11	12	13

- A. **follows** is empty ArrayList, **pos** is zero
- B. While there is more searching to do
1. **index** = first occurrence of KEY starting at **pos**
  2. if found, add value at [**index**+1] to **follows**
  3. **pos** = **index** + 1
- C. Your answer is **follows**

Answer: 'n', 'p', ' ', 'y'