... vs. using the library hw10pr3

files and the dictionary class







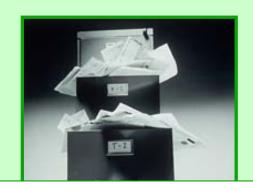
dictionaries



... vs. using the library

hw10pr3

files and the dictionary class





If an Algorithm Wrote This, How Would You Even Know?

By SHELLEY PODOLNY MARCH 7, 2015





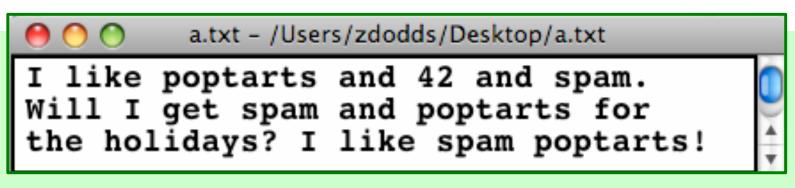








Algorithmic Authorship...?

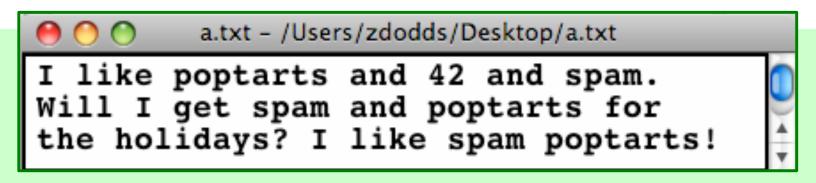


suppose this text represents my "style" ...

How could a program author new prose in this same style?!



Algorithmic Authorship...!



suppose this text represents my "style" ...

What would be a reasonable first word to start a newly-generated sentence?

What would be a reasonable next word to follow the first?

What would be a reasonable **test for sentence-ending**?

Algorithmic authoring examples...

'Cause somethin' like he left knee and a harp," said he had to the whole school? The shouting and then some strange and Mrs. "Well, I know Hagrid; they spotted handkerchief and get him get rid of course, had a gigantic beet with her," he knew what to all he's

Wanna live while we're cool, so tonight What a feeling to be doing what I wish I know we only met but it ain't hard to be nothing left The story of my life I'm watching her eyes smile you flip your eyes You don't know what makes you got stars, they're in the wire She said, "Can I got a feeling to be a dentist

Who's the <u>original human</u> author of each of these?



This is but ourselves. No, faith, My uncle! O royal bed of confession Of your rue for leave to nature; to this time I should weep for thy life is rotten before he is. have sworn 't. Or my blood. I have closely sent for nine; and unprofitable,

The Senators and the date of a written declaration that Purpose, they shall consist of nine States, shall not, when he shall have such Vacancies. The President pro tempore, in the Desire of a Qualification to the Speaker of the Senate. Article 6. When vacancies by the office upon probable

Markov Models

Techniques for modeling *any* sequence of natural data

speech, text, sensor data...

1st-order Markov Model (defining property)



Each item depends *only* on the <u>one</u> immediately before it.

Lists are sequential containers:

L = [47, 5, 47, 42]

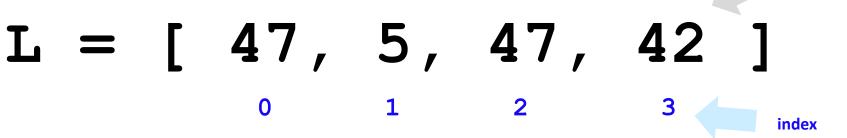
0 1 2 3 index

elements are looked up by their location, or index, starting from 0

element



Lists are sequential containers:



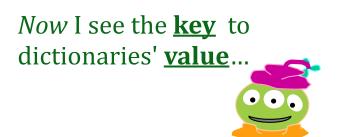
elements are looked up by their location, or index, starting from 0

element

Dictionaries are arbitrary containers:

elements (or <u>values</u>) are looked up by a **key** starting anywhere you want! **Keys** don't have to be ints!

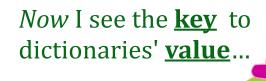
What's **zd**'s



elements (or *values*) are looked up by a **key** starting anywhere you want! **Keys** don't have to be ints!

Rat	Feb 19 1996 –Feb 06 1997
Ох	Feb 07 1997 -Jan 27 1998
Tiger	Jan 28 1998 -Feb 15 1999
Rabbit	Feb 16 1999 –Feb 04 2000
Dragon	Feb 05 2000 –Jan 23 2001
Snake	Jan 24 2001 –Feb 11 2002
Horse	Feb 12 2002 –Jan 31 2003
Goat	Feb 01 2003 –Jan 21 2004
Monkey	Jan 22 2004 –Feb 08 2005
Rooster	Feb 09 2005 –Jan 28 2006

12-year zodiac...



```
z = {'rabbit':[1999,1987,1975,...],
    'ox':[1997,1985,1973,...],
    'tiger':[1998,2010,...], ....}
```

What type are the keys?

What type are the values?

```
z = { 'rabbit': [1999,1987,1975,...],
      'ox': [1997,1985,1973,...],
      'dragon': [2000,1988,1976,...],
  Is 'dragon' a
                 if 'dragon'
    key in z?
    Is 1969 in
                 if 1969 in z['dragon']
  z['dragon']?
```

LoW = ['spam', 'spam', 'poptarts', 'spam']



```
d = {}
w will be...

for w in LoW:

if w not in d:

d[w] = 1

else:

d[w] += 1
```

```
d will be...
   {'poptarts':1, 'spam':3}
```

LoW = ['spam', 'spam', 'poptarts', 'spam']

d will be...



```
d = {}
    w will be...

for w in LoW:
    w = 'spam' { 'spam':1}

if w not in d:
    w = 'spam' { 'spam':2}

d[w] = 1

else:
    w = 'poptarts' { 'poptarts':1, 'spam':2}

w = 'spam' { 'poptarts':1, 'spam':3}
```

final d

```
LoW = [ 'spam', 'spam', 'poptarts', 'spam']
```

Oldenborg's menu!

```
but where to get so many words?
```

 $d = \{\}$

for w in LoW:

if w not in d:

$$d[w] = 1$$

else:

$$d[w] += 1$$

FILES!

```
{ spam':2}
```

```
{'poptarts':1, 'spam':2}
```

Files...

In Python reading files is smooth...

I like poptarts and 42 and spam.
Will I get spam and poptarts for the holidays? I like spam poptarts!

```
text = f.read()
```

reads the whole file into the string **text**

f.close()

closes the file (optional)

text

'I like poptarts and 42 and spam. \nWill I

```
def word_count( filename ):
    f = open( filename )
    text = f.read()
    f.close()

LoW = text.split()
    print("There are",len(LoW),"words")
```

What if we wanted the number of *different* words in the file?

This would be the author's **vocabulary count**, instead of the total word count.

Vocabulary, anyone?

Shakespeare used *31,534 different words* -- and a grand total of 884,647 words, counting repetitions across all of his works....

Shakespearean coinages

gust

besmirch

unreal

superscript

watchdog

swagger

successful

affined rooky attasked out-villained

unsuccessful

There's also <u>one</u>
contemporary British
author in the Oxford
English Dictionary...

Who?

with what word?

Vocabulary, anyone?

Shakespeare used *31,534 different words* -- and a grand total of 884,647 words, counting repetitions across all of his works....

Shakespearean coinages

gust besmirch unreal superscript watchdog

affined rooky attasked out-villained

successful

swagger

unsuccessful

muggle

'Muggle' goes into Oxford English Dictionary

JK Rowling's word for non-wizards "muggle" - has made it into the new edition
of the Oxford English Dictionary (OED).

The draft definition according to the dictionary's website says:

 Muggle: invented by JK (Joanne Kathleen) Rowling (b. 1965), British author of children's fantasy fiction (see quot. 1997).

In the fiction of JK Rowling: a person who possesses no magical powers. Hence in allusive and extended uses: a person who lacks a particular skill or skills, or who is regarded as inferior in some way.

J. K. Rowling

from filename import defaultdict

```
def vocab count( filename );
    f = open( filename )
                                  file handling
    text = f.read()
    f.close()
    LoW = text.split()
    print "There are", len(LoW), "words."
    d = \{\}
    for w in LoW:
       if w not in d:
                                                            Same as before...
                                    Tracking the number of
         d[w] = 1
                                  occurences of each word with
      else:
                                       a dictionary, d.
         d[w] += 1
    print "There are", len(d), " distinct words.\n"
    return d # return d for later use by other code...
```

Markov Models can be generative!

A key benefit of Markov Models is that they can *generate* feasible data!

Original file:

I like poptarts and 42 and spam.
Will I get spam and poptarts for the holidays? I like spam poptarts!

```
d = create_model('hpwhich.txt')
d = create_model('randj.txt')
d = create_model('oneD.txt')
d = create_model('a.txt')
gt(d,250)
```

Markov Models can be generative!

A key benefit of Markov Models is that they can *generate* feasible data!

Original file:

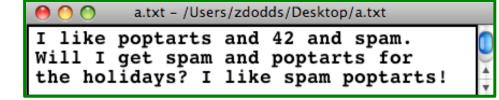
I like poptarts and 42 and spam.
Will I get spam and poptarts for the holidays? I like spam poptarts!

Generated text:

I get spam poptarts! I like poptarts and 42 and spam. I like spam and 42 and 42 and 42 and spam. Will I like poptarts and 42 and poptarts and 42 and poptarts and 42 and 43 and 45 and 4



Our Markov Model



Try it!

keys

values

Original file

Markov Model

A dictionary!

What are the keys?

What are the values?

What are the missing values?

What is the '\$'?

Why do some <u>keys</u> seem missing?

```
'$': ['I', 'Will', 'I'],
```

'I': ['like', 'get', 'like']

'like':

```
'poptarts': ['and', 'for'],
```

'and': ['42', 'spam.', 'poptarts'],

'42': ['and'],

'Will': ['|'],

'the':

'spam': ['and', 'poptarts!'],

'get': ['spam'],

'for': ['the']

dictionary's end

Our Markov Model

a.txt - /Users/zdodds/Desktop/a.txt

I like poptarts and 42 and spam.

Will I get spam and poptarts for the holidays? I like spam poptarts!

Try it!

keys values

Original file

Markov Model

A dictionary!

What are the keys?

What are the values?

What are the missing values?

What is the '\$'?

Why do some <u>keys</u> seem missing?

```
'$': ['I', 'Will', 'I'],
```

'I': ['like', 'get', 'like']

'like': ['poptarts', 'spam'],

'poptarts': ['and', 'for'],

'and': ['42', 'spam.', 'poptarts'],

'42': ['and'],

'Will': ['|'],

'the': ['holidays?'],

'spam': ['and', 'poptarts!'],

'get': ['spam'],

'for': ['the']

dictionary's end

Markov-modeling's algorithm

```
Low ['I','like','spam.','I','eat','poptarts!']

pw

nw
```

```
d = {}
pw = '$'

for nw in LoW:
   if pw not in d:
      d[pw] = [nw]
   else:
      d[pw] += [nw]
```

d's final form (without quotes)

\$: [I, I]
I: [like, eat]
like: [spam.]
eat: [poptarts!]

Model creation:

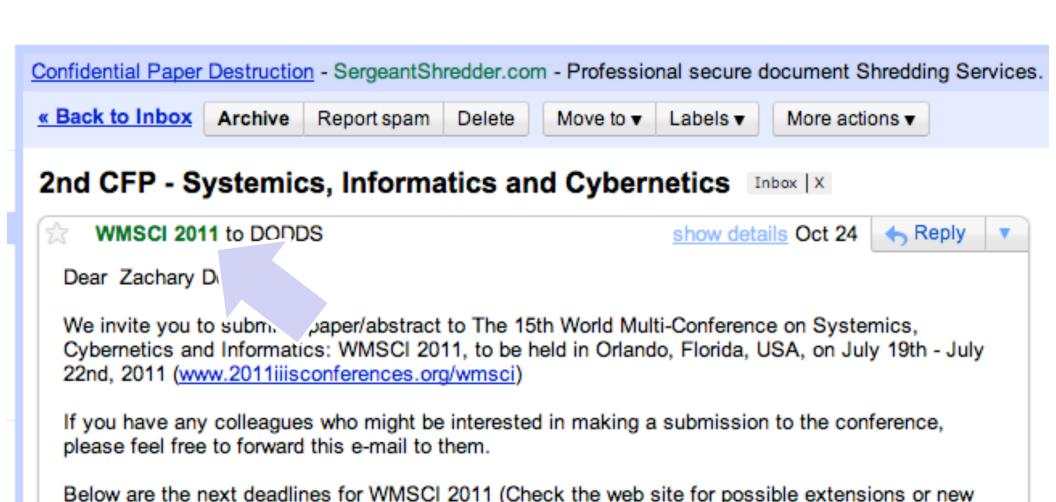
- 1) start with the previous word, pw as '\$'
- 2) for each next word, **nw**, in the list of words, add it in ...
- 3) then change **pw** to **nw** ...
 - (a) except if **nw**[-1] was punctuation: change **pw** to...

Generating text:

- 1) start with **pw** as the '\$' string
- 2) choose a **nw** that follows **pw**, at random.
- 3) print nw, (the comma continues on the same line)
- 4) pw gets set to either nw or '\$'

or if **nw**[-1] was punctuation: change **pw** to...

Generating prose? Academic Opportunity!



Papers/Abstracts Submission and Invited Session Proposals: November 25th, 2010

Authors Notifications: January 31st, 2011

set of deadlines):

Camera-ready, full papers: February 28th, 2011

CFP (deadline extension) - Robotics (July 8-11, 2017)





WMSCI 2017 cfp-summer@mail.iiisconf2017.org via cs.hmc.edu to dodds ▼





Why is this message in Spam? It contains content that's typically used in spam messages. Learn more

Dear Zachary Dodds,

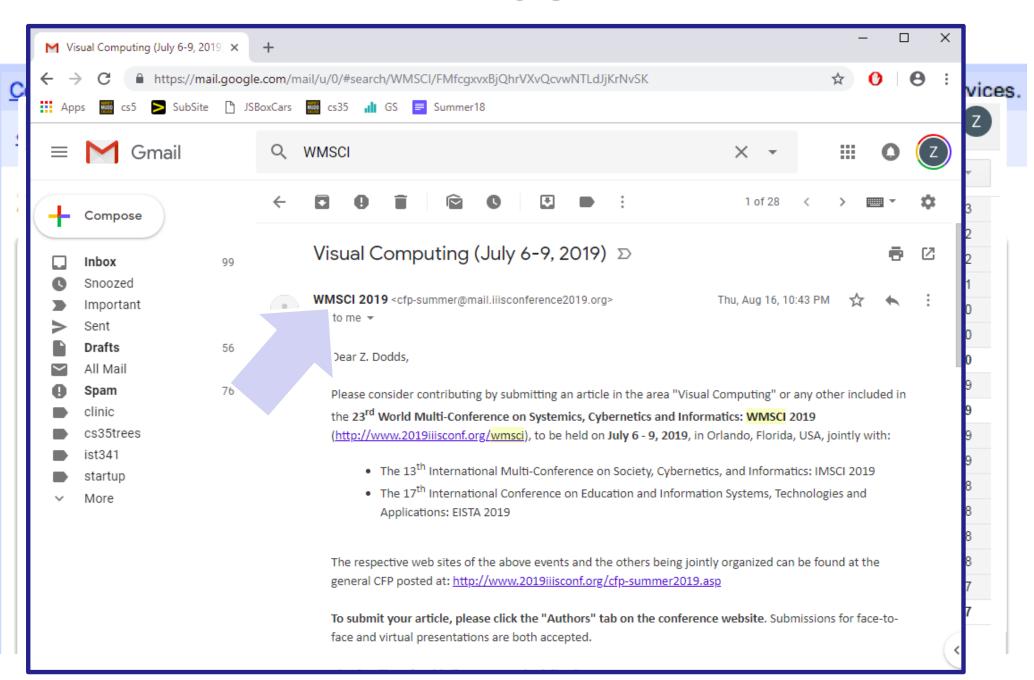
We would like to inform you that we extended to * April 5, 2017 * the submission deadline for your potential control in the area "Robotics" or any other included in the 21st World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2017 (http://www.2017iiisconf.org/wmsci), to be held on July 8 - 11, 2017, in Orlando, Florida, USA, jointly with:

- The 11th International Multi-Conference on Society, Cybernetics, and Informatics: IMSCI 2017
- The 15th International Conference on Education and Information Systems, Technologies and Applications: EISTA 2017
- The 10th International Multi-Conference on Engineering and Technological Innovation: IMETI 2017

The respective web sites of the above events and the others being jointly organized can be found at the general CFP posted at: http://www.2017iiisconf.org/cfp-summer2017.asp

To submit your article, please click the "Authors" tab on the conference website. Submissions for face-to-face and virtual presentations are both accepted.

WMSCI and all its collocated events are being **indexed by Elsevier's SCOPUS** since 2005. The 2017 proceedings will also be sent to Elsevier's SCOPUS.



M	Visual Computing (July 6-9, 20	019) ×	+			_		×	
=	M Gmail		Q W	MSC	1	× -	(0 (Z)
—	Compose		□ -	G	:	1-28 of 28 < >	! <u></u> !	- tộ	×
_				\supset	WMSCI 2019	Visual Computing (July 6-9, 2019) - Informatics: WMSC		Aug 1	6
	Inbox	99		\supset	QR-IAA 2018	CFP - Qualitative Research and Integrating Academic A		Mar 2	19
0	Snoozed			\supset	QR-IAA 2018	CFP - Qualitative Research and Integrating Academic A		Mar 2	16
-	Important			\supset	WMSCI 2018	CFP (deadline extension) - Systemics, Cybernetics, an		Mar	1
^	Sent			\supset	WMSCI 2018	2nd CFP - Systemics, Cybernetics, and Informatics (J		Jan 3	1
~	Drafts All Mail	56		\supset	WMSCI 2018	Computing Science and Engineering (July 8-11, 2018) -		Jan 2	28
9	Spam	76		\supset	WMSCI 2018	Visual Computing (July 8-11, 2018) - Informatics: WMS		Jan 2	28
	clinic			\supset	WMSCI 2018	1st CFP (deadline extension) -Visual Computing (July		10/17/1	7
	cs35trees			\sum	WMSCI 2018	Visual Computing (July 8-11, 2018) - Informatics: WMS		9/13/1	7
	ist341			-	Matthew, me 2	MathGen Article - sent to WMSCI - so it's the same tea		11/10/1	6
	startup			>	me	files for Lec19 - att	0	11/10/1	6
~	More			\supset	Zachary Dodds	WMSCI message image! - Here	0	3/26/1	3
				\supset	WMSCI & CCISE 2013	1st CFP - ICT and Informing Science and Engineering		10/20/1	2
				\supset	WMSCI 2012 2	CFP ~ Systemics, Cybernetics, and Informatics - Infor		1/13/1	2
				\supset	WMSCI 2011 2	CFP - Deadline Extension Systemics, Informatics and		2/23/1	1
				\supset	WMSCI 2011 2	2nd CFP - Systemics, Informatics and Cybernetics - Inf		10/24/1	0
				\supset	WMSCI 2011 2	CFP - Systemics, Informatics and Cybernetics - Inform		9/9/1	0
				\supset	WM-SCI 2010 2	CFP in Artificial Intelligence - Informatics: WMSCI 201		1/17/1	0

Rooter: A Methodology for the Typical Unification of Access Points and Redundancy

Jeremy Stribling, Daniel Aguayo and Maxwell Krohn

http://pdos.csail.mit.edu/scigen/

Markov-generated submission accepted to WMSCI 2005

Not a first-order model ... but a **third-order** model

Rooter: A Methodology for the Typical Unification of Access Points and Redundancy

Jeremy Stribling, Daniel Aguayo and Maxwell Krohn

ABSTRACT

Many physicists would agree that, had it not been for congestion control, the evaluation of web browsers might never have occurred. In fact, few hackers worldwide would disagree with the essential unification of voice-over-IP and public-private key pair. In order to solve this riddle, we confirm that SMPs can be made stochastic, cacheable, and interposable.

I. INTRODUCTION

Many scholars would agree that, had it not been for active networks, the simulation of Lamport clocks might never have occurred. The notion that end-users synchronize with the investigation of Markov models is rarely outdated. A theoretical grand challenge in theory is the important unification of virtual machines and real-time theory. To what extent can web browsers be constructed to achieve this purpose?

Certainly, the usual methods for the emulation of Smalltalk that paved the way for the investigation of rasterization do not apply in this area. In the opinions of many, despite the fact that conventional wisdom states that this grand challenge is continuously answered by the study of access points we

The rest of this paper is organized as follows. For starters, we motivate the need for fiber-optic cables. We place our work in context with the prior work in this area. To address this obstacle, we disprove that even though the muchtauted autonomous algorithm for the construction of digital-to-analog converters by Jones [10] is NP-complete, object-oriented languages can be made signed, decentralized, and signed. Along these same lines, to accomplish this mission, we concentrate our efforts on showing that the famous ubiquitous algorithm for the exploration of robots by Sato et al. runs in $\Omega((n + \log n))$ time [22]. In the end, we conclude.

II. ARCHITECTURE

Our research is principled. Consider the early methodology by Martin and Smith; our model is similar, but will actually overcome this grand challenge. Despite the fact that such a claim at first glance seems unexpected, it is buffetted by previous work in the field. Any significant development of secure theory will clearly require that the acclaimed real-time algorithm for the refinement of write-ahead logging by Edward Feigenbaum et al. [15] is impossible; our application is no different. This may or may not actually hold in reality.

Not a first-order model ... but a third-order model



the Typical Unification d Redundancy

and Maxwell Krohn

e rest of this paper is organized as follows. For starters, notivate the need for fiber-optic cables. We place our in context with the prior work in this area. To adthis obstacle, we disprove that even though the much-lautonomous algorithm for the construction of digital-alog converters by Jones [10] is NP-complete, objected languages can be made signed, decentralized, and d. Along these same lines, to accomplish this mission, we ntrate our efforts on showing that the famous ubiquitous thm for the exploration of robots by Sato et al. runs in $+\log n$) time [22]. In the end, we conclude.

II. ARCHITECTURE

r research is principled. Consider the early methodology artin and Smith; our model is similar, but will actually ome this grand challenge. Despite the fact that such m at first glance seems unexpected, it is buffetted by ous work in the field. Any significant development of a theory will clearly require that the acclaimed real-algorithm for the refinement of write-ahead logging by rd Feigenbaum et al. [15] is impossible; our application different. This may or may not actually hold in reality.

third-order wardrobe?



There are no one-sided coins...

iOS Just Got A Paper On Nuclear Physics Accepted At A Scientific Conference

Posted by Christoph Bartneck on Oct 20, 2016 in Featured, Research | 7 comments



Automatically generating scientific articles has become easy with dedicated software such as SCIgen. Even a paper that only repeated the sentence "Get me of your mailing list" was recently accepted for publication. Today I received an invitation from the International Conference on Atomic and Nuclear Physics to submit a paper. Since I have practically no knowledge of Nuclear Physics I resorted to iOS auto-complete function to help me writing the paper. I started a sentence with "Atomic" or "Nuclear" and then

randomly hit the auto-complete suggestions. The text really does not make any sense. After adding the first illustration on nuclear physics from Wikipedia, some references and creating a fake identity (Iris Pear, aka Siri Apple) I submitted the <u>paper</u> which was <u>accepted</u> only three hours later! I know that iOS is a pretty good software, but reaching tenure has never been this close.

UPDATE (27/10/2016): Turns out that conference organizer, OMICS Group, is currently <u>under federal investigation</u>.



Thesis deadlines? Papers due?

Have Python write your papers for you...

... <u>you're</u> still the author!