

## Lab 1

---

Steven Bird, Ewan Klein, and Edward Loper

### 4 Writing Structured Programs

10. Write a program to sort words by length. Define a helper function `cmp_len` which uses the `cmp` comparison function on word lengths.

**Answer:**

11. Create a list of words and store it in a variable `sent1`. Now assign `sent2 = sent1`. Modify one of the items in `sent1` and verify that `sent2` has changed.
1. Now try the same exercise but instead assign `sent2 = sent1[:]`. Modify `sent1` again and see what happens to `sent2`. Explain.
  2. Now define `text1` to be a list of lists of strings (e.g. to represent a text consisting of multiple sentences. Now assign `text2 = text1[:]`, assign a new value to one of the words, e.g. `text1[1][1] = 'Monty'`. Check what this did to `text2`. Explain.
  3. Load Python's `deepcopy()` function (i.e. `from copy import deepcopy`), consult its documentation, and test that it makes a fresh copy of any object.

**Answer:**

Beginning Problem 4.11 Part 0

Before

```
True:"good" == "good"
True:"keeper" == "keeper"
True:"I" == "I"
True:"slightly" == "slightly"
True:"minerality" == "minerality"
True:"rather" == "rather"
True:"'d" == "'d"
True:". " == ". "
True:"tannins" == "tannins"
True:"at" == "at"
```

Lists are identical

After

```
True:"good" == "good"
True:"keeper" == "keeper"
True:"I" == "I"
True:"slightly" == "slightly"
True:"minerality" == "minerality"
True:"rather" == "rather"
True:"'d" == "'d"
True:"PROBLEM 0: VERY OBVIOUS REPLACEMENT" == "PROBLEM 0: VERY OBVIOUS REPLACEMENT"
True:"tannins" == "tannins"
True:"at" == "at"
Lists are identical
```

Beginning Problem 4.11 Part 1

Before

```
True:"nicely" == "nicely"
True:"." == "."
True:"dry" == "dry"
True:"Very" == "Very"
True:"acidity" == "acidity"
True:"**" == "**"
True:"these" == "these"
True:")" == ")"
True:"wound" == "wound"
True:"while" == "while"
Lists are identical
```

After

```
True:"nicely" == "nicely"
True:"." == "."
True:"dry" == "dry"
True:"Very" == "Very"
True:"acidity" == "acidity"
True:"**" == "**"
False:"PROBLEM 1: VERY OBVIOUS REPLACEMENT" == "these"
True:")" == ")"
True:"wound" == "wound"
True:"while" == "while"
Lists are not identical
```

Beginning Problem 4.11 Part 2

Before

```
True:"came" == "came"
True:"," == ","
True:"to" == "to"
True:"Persistent" == "Persistent"
True:"Bright" == "Bright"
```

```

True:"this" == "this"
True:"all" == "all"
True:"Down" == "Down"
True:"bit" == "bit"
True:"but" == "but"
Lists are identical
Caught a TypeError, because we tried to assign something to a character in a str

```

Beginning Problem 4.11 Part 3

Performing Deep Copy

Modifying Random Element

After

```

True:"that" == "that"
True:"a" == "a"
False:"PROBLEM 3: VERY OBVIOUS MODIFICATION" == "Rated"
True:"Dense" == "Dense"
True:"(" == "("
True:"nose" == "nose"
True:"in" == "in"
True:"good" == "good"
True:"structure" == "structure"
True:"very" == "very"
Lists are not identical

```

12. Initialize an n-by-m list of lists of empty strings using list multiplication, e.g. `word_table = [[''] * n] * m`. What happens when you set one of its values, e.g. `word_table[1][2] = "hello"`? Explain why this happens. Now write an expression using `range()` to construct a list of lists, and show that it does not have this problem.

**Answer:**

```

Building empty word table: word_table = [[''] * n] * m
[['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', '']]
Assigning to hopefully one spot: word_table[1][2] = 'hello'

```

```

[['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', 'hello', '', '', '', '']]

```

It looks like there's only one n-dimensional vector that's being referenced m-many times.  
 Building empty word table using range() for enumeration and append()

```

[['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', '']]

```

Assigning to hopefully one spot: word\_table[1][2] = 'hello'

```

[['', '', '', '', '', '', ''],
 ['', '', 'hello', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', ''],
 ['', '', '', '', '', '', '']]

```

13. Write code to initialize a two-dimensional array of sets called textttword\_vowels and process a list of words, adding each word to word\_vowels[l][v] where l is the length of the word and v is the number of vowels it contains.

**Answer:**

```

0 0 set()
0 1 set()

```

```

0 2 set()
0 3 set()
0 4 set()
0 5 set()
0 6 set()
1 0 {'[' , '-'}
1 1 {'A' , 'a'}
1 2 set()
1 3 set()
1 4 set()
1 5 set()
1 6 set()
2 0 {'XP' , ''''}
2 1 {'to' , 'in'}
2 2 set()
2 3 set()
2 4 set()
2 5 set()
2 6 set()
3 0 {"n't"}
3 1 {'and' , 'not' , 'web' , 'the' , 'tab'}
3 2 set()
3 3 set()
3 4 set()
3 5 set()
3 6 set()
4 0 set()
4 1 {'tall' , 'pref' , 'from' , 'when'}
4 2 {'like'}
4 3 set()
4 4 set()
4 5 set()
4 6 set()
5 0 set()
5 1 set()
5 2 {'boxes' , 'opens'}
5 3 set()
5 4 set()
5 5 set()
5 6 set()
6 0 set()
6 1 set()
6 2 {'window' , 'search' , 'tools-' , 'render'}
6 3 {'option'}
6 4 set()

```

```
6 5 set()
6 6 set()
7 0 set()
7 1 set()
7 2 {'closing', 'blocked', 'results'}
7 3 {'Manager', 'without', 'remains'}
7 4 {'outlook'}
7 5 set()
7 6 set()
8 0 set()
8 1 set()
8 2 set()
8 3 {'Download', 'rotating', 'rendered'}
8 4 set()
8 5 set()
8 6 set()
9 0 set()
9 1 set()
9 2 set()
9 3 set()
9 4 {'maximized'}
9 5 set()
9 6 set()
10 0 set()
10 1 set()
10 2 set()
10 3 {'Ctrl+mouse'}
10 4 {'completion'}
10 5 set()
10 6 set()
11 0 set()
11 1 set()
11 2 set()
11 3 set()
11 4 {'problematic'}
11 5 set()
11 6 set()
12 0 set()
12 1 set()
12 2 set()
12 3 set()
12 4 set()
12 5 set()
12 6 {'Autocomplete'}
```

14. Write a function `novel10(text)` that prints any word that appeared in the last 10% of a text that had not been encountered earlier.

**Answer:**

```
USING wine.txt FROM webtext
```

```
Unique 10% tail:
```

```
Blin
```

```
oaky
```

```
P-M
```

```
deliciously
```

```
matters
```

```
non-fruit
```

```
belnd
```

```
/ul
```

```
Valvona
```

```
These
```

```
M-C.
```

```
finesse
```

```
morning
```

```
Dom
```

```
Australian
```

```
chateauneuf
```

```
warm-vintage
```

```
lychee
```

```
failed
```

```
>
```

```
Lime
```

```
loads
```

```
Zind
```

```
struck
```

```
Appetizing
```

```
competition
```

```
ideally
```

```
Hill
```

```
Cough-sweet
```

```
success
```

```
till
```

```
Vincafe
```

```
gamay
```

```
Sold
```

```
rely
```

```
weakening
```

```
noticed
```

```
characteristically
```

```
Spicey
```

Coarsened  
putting  
afraid  
sherbert  
slavours  
literally  
matches  
Colheita  
dumb  
Koonunga  
faulty  
company  
brioche  
attributed  
foetid  
stand  
Grahams  
terms  
exceptional  
Penfolds  
wax  
softly  
minutes  
follows  
set  
30  
convince  
nibbles  
Class  
vulgar  
Promoted  
DP  
tarted  
Magnum  
chatracter  
old-style  
else  
average  
fast  
outstanding  
gently  
Aiming  
Slate  
Jobard  
sorting



bovril  
Juniper/Gin  
Volnay  
Depths  
shaped  
depths  
nothing  
Whatever  
happen  
taking  
overlying  
sees  
sour  
orange-rimmed  
assistants  
Gorgeously  
spot  
Rhone-like  
favourite  
contributor  
Butterscotch  
result  
Seepage  
VV  
C-R  
Finishes  
ages  
head  
PLeasant  
velvety  
tin  
metaphorically  
vanillan  
typicity  
Deutz  
Chocolatey  
unflawed  
resinous  
heaviness  
tropical-edged  
inbetween  
1978  
fragrantly  
persistence  
noted

kind  
ethereal  
semillon  
hell  
Rousanne  
muscular  
domain  
struggle  
winemaker  
marks  
alleviate  
regret  
Personally  
Beetroot  
Partly  
loved  
24  
seal  
contemplatively  
four-square  
kept  
Satisfying  
searing  
indication  
relationship  
Caramel  
Highly-wrought  
Provencal  
<  
sepcially  
Risky  
Rayas  
beyond-fruit  
Modern-styled  
beat  
variety  
Tokay  
non-Burgundian  
yummy  
reasobale  
Yummy  
loose  
paint-stripper  
Perignon  
blanced

pineapple  
arresting  
Written  
Acacia  
overawed  
turned  
surprisingly  
Dies  
MSR  
Riserva

15. Write a program that takes a sentence expressed as a single string, splits it and counts up the words. Get it to print out each word and the word's frequency, one per line, in alphabetical order.

**Answer:**

```
[jmcgover@MAINFRAME 1lab]$ python lab1-bk14.15.py
When it does it has amazing intensity and length .
amazing : 1
and : 1
does : 1
has : 1
intensity : 1
it : 2
length : 1
when : 1
[jmcgover@MAINFRAME 1lab]$ python lab1-bk14.15.py
Hee ha ha ..
ha : 2
hee : 1
```

16. Read up on Gematria, a method for assigning numbers to words, and for mapping between words having the same number to discover the hidden meaning of texts (<http://en.wikipedia.org/wiki/Gematria>, <http://essen.es.net/gemcal.htm>).

1. Write a function `gematria()` that sums the numerical values of the letters of a word, according to the letter values in `letter_vals`:
2. Process a corpus (e.g. `nltk.corpus.state_union`) and for each document, count how many of its words have the number 666.
3. Write a function `decode()` to process a text, randomly replacing words with their Gematria equivalents, in order to discover the "hidden meaning" of the text.

**Answer:**

17. Write a function `shorten(text, n)` to process a text, omitting the `n` most frequently occurring words of the text. How readable is it?

**Answer:** Somewhat readable. Kind of funny.

SCENE 1 : [ wind ] [ clop clop clop ] KING ARTHUR : Whoa there ! [ clop clop clop ] SOLDIER # 1 : Halt ! Who goes there ? ARTHUR : It is I , Arthur , son of Uther Pendragon , from the castle of Camelot . King of the Britons , defeator of the Saxons , sovereign of all England ! SOLDIER # 1 : Pull the other one ! ARTHUR : I am , ... and this is my trusty servant Patsy . We have ridden the length and breadth of the land in search of knights who will join me in my court at Camelot . I must speak with your lord and master . SOLDIER # 1 : What ? Ridden on a horse ? ARTHUR : Yes ! SOLDIER # 1 : You 're using coconuts ! ARTHUR : What ? SOLDIER # 1 : You 've got two empty halves of coconut and you 're bangin ' 'em together . ARTHUR : So ? We have ridden since the snows of winter covered this land , through the kingdom of Mercea , through -- SOLDIER # 1 : Where 'd

Shortening text....

OMITTING 50 MOST FREQUENT:

FREQUENT:1:who  
FREQUENT:1:coconut  
FREQUENT:1:will  
FREQUENT:1:castle  
FREQUENT:1:search  
FREQUENT:1:servant  
FREQUENT:1:must  
FREQUENT:1:snows  
FREQUENT:1:wind  
FREQUENT:1:King  
FREQUENT:1:you  
FREQUENT:1:Where  
FREQUENT:1:halves  
FREQUENT:1:Yes  
FREQUENT:1:Whoa  
FREQUENT:1:Patsy  
FREQUENT:1:join  
FREQUENT:1:other  
FREQUENT:1:Mercea  
FREQUENT:1:one  
FREQUENT:2:this  
FREQUENT:2:'re  
FREQUENT:2:there  
FREQUENT:2:What  
FREQUENT:2:have  
FREQUENT:2:through

```

FREQUENT:2:You
FREQUENT:2:in
FREQUENT:2:ridden
FREQUENT:2:Camelot
FREQUENT:2:land
FREQUENT:2:my
FREQUENT:2:We
FREQUENT:2:is
FREQUENT:3:[
FREQUENT:3:]
FREQUENT:3:I
FREQUENT:4:and
FREQUENT:5:?
FREQUENT:5:.
FREQUENT:6:clap
FREQUENT:6:#
FREQUENT:6:ARTHUR
FREQUENT:6:!
FREQUENT:6:SOLDIER
FREQUENT:7:1
FREQUENT:8:,
FREQUENT:8:the
FREQUENT:10:of
FREQUENT:13::
SCENE KING Halt Who goes It Arthur son Uther Pendragon from Britons
defeater Saxons sovereign all England Pull am ... trusty length
breadth knights me court at speak with your lord master Ridden on a
horse using coconuts 've got two empty bangin ' 'em together So since
winter covered kingdom --

```

18. Write code to print out an index for a lexicon, allowing someone to look up words according to their meanings (or pronunciations; whatever properties are contained in lexical entries).

**Answer:**

19. Write a list comprehension that sorts a list of WordNet synsets for proximity to a given synset. For example, given the synsets `minke_whale.n.01`, `orca.n.01`, `novel.n.01`, and `tortoise.n.01`, sort them according to their `shortest_path_distance()` from `right_whale.n.01`.

**Answer:**

```

[Synset('lesser_rorqual.n.01'),
 Synset('killer_whale.n.01'),
 Synset('novel.n.01'),

```

```

    Synset('tortoise.n.01')]
Sorting...
[Synset('lesser_rorqual.n.01'),
 Synset('killer_whale.n.01'),
 Synset('tortoise.n.01'),
 Synset('novel.n.01')]

```

20. Write a function that takes a list of words (containing duplicates) and returns a list of words (with no duplicates) sorted by decreasing frequency. E.g. if the input list contained 10 instances of the word table and 9 instances of the word chair, then table would appear before chair in the output list.

**Answer:**

```

['White',
 'guy',
 ':',
 'So',
 ',',
 'do',
 'you',
 'have',
 'any',
 'plans',
 'for',
 'this',
 'evening',
 '?',
 'Asian',
 'girl',
 ':']
":":11501
".":10711
",":8722
"I":6625
"?":4177
"you":3932
"#":3580
"the":3506
"to":3239
"a":3084
"!":3042
"s":2582
"n't":2299
"1":1987

```

```
"and":1981
"that":1814
"it":1790
```

25. Read about string edit distance and the Levenshtein Algorithm. Try the implementation provided in `nltk.edit_distance()`. In what way is this using dynamic programming? Does it use the bottom-up or top-down approach? [See also <http://norvig.com/spell-correct.html>]

**Answer:**

```
DISTANCE:conventional:glass: 12
DISTANCE:,:acidity: 7
DISTANCE:OK:from: 4
DISTANCE:93+94+95:Bare: 8
DISTANCE:..to: 2
DISTANCE:..): 1
DISTANCE:proper:on: 5
DISTANCE:sweet:blackcurrant: 11
DISTANCE:but:influence: 8
DISTANCE:..Not: 3
DISTANCE:drying:Might: 6
DISTANCE:pleasant:acidity: 8
DISTANCE:):refined: 7
DISTANCE:two:juiciness: 9
DISTANCE:of:..: 2
DISTANCE:Lovely:the: 5
DISTANCE:like:always: 5
```

Dynamic programming is built off the notion of finding optimal subproblems, where some of these subproblems overlap with each other or the current problem, in order to find the most optimal solution to the current problem. In the case of edit distance, how we edit earlier in the word may affect what the most optimal solution is at the current stage of the solution. In the case of Levenshtein distance, it is a bottom up approach for some algorithms, because a decision is made on what to calculate next.

26. Obtain some raw text, in the form of a single, long string. Use Python's `textwrap` module to break it up into multiple lines. Now write code to add extra spaces between words, in order to justify the output. Each line must have the same width, and spaces must be approximately evenly distributed across each line. No line can begin or end with a space.

**Answer:**

```
UNWRAPPED
#1: You're using coconuts!
```

ARTHUR: What?  
 SOLDIER #1: You've got two empty halves of coconut and you're bangin' 'em together.  
 ARTHUR: So? We have ridden since the snows of winter covered this land, through the  
 SOLDIER #1: Where'd you get the coconuts?  
 ARTHUR: We found t  
 WRAPPED  
 #1: You're using coconuts! ARTHUR:  
 What? SOLDIER #1: You've got two empty  
 halves of coconut and you're bangin' 'em  
 together. ARTHUR: So? We have ridden  
 since the snows of winter covered this  
 land, through the kingdom of Mercea,  
 through-- SOLDIER #1: Where'd you get  
 the coconuts? ARTHUR: We found t

27. Develop a simple extractive summarization tool, that prints the sentences of a document which contain the highest total word frequency. Use `FreqDist()` to count word frequencies, and use `sum` to sum the frequencies of the words in each sentence. Rank the sentences according to their score. Finally, print the `n` highest-scoring sentences in document order. Carefully review the design of your program, especially your approach to this double sorting. Make sure the program is written as clearly as possible.

**Answer:**

PROCESSING:1945-Truman.txt

1468:Nothing is more essential to the future peace of the world than continued cooperation of the nations which had to muster the force necessary to defeat the conspiracy of the Axis powers to dominate the world.

1396:Mr. Speaker, Mr. President, Members of the Congress: It is with a heavy heart that I stand before you, my friends and colleagues, in the Congress of the United States.

-----  
 PROCESSING:1946-Truman.txt

34906:Other increases are for civil aeronautics promotion, the business and manufacturing censuses, and other expanded business services of the Department of Commerce which have been referred to above; the Forest and Soil Conservation Services and other activities of the Department of Agriculture; certain conservation activities of the Department of the Interior; and the collection of internal revenue in the Treasury Department.

30682:I anticipate that net expenditures of the Export- Import Bank and expenditures arising from the British credit and the Bretton Woods Agreements will amount to 2, 614 million dollars, including the noncash item of 950 million



dollars for the Fund, in the fiscal year 1946, and 2, 754 million dollars in the fiscal year 1947.

---

PROCESSING:1947-Truman.txt

6599:Whatever differences there may have been between us and the Soviet Union, however, should not be allowed to obscure the fact that the basic interests of both nations lie in the early making of a peace under which the peoples of all countries may return, as free men and women, to the essential tasks of production and reconstruction.

6427:Fourth, the balancing of the budget in the next fiscal year and the achieving of a substantial surplus to be applied to the reduction of the public debt, Fifth, protection of a fair level of return to farmers in post- war agriculture.

---

PROCESSING:1948-Truman.txt

4904:On this occasion, above all others, the Congress and the President should concentrate their attention, not upon party but upon the country; not upon things which divide us but upon those which bind us together- the enduring principles of our American system, and our common aspirations for the future welfare and security of the people of the United States.

4090:We have learned that the loss of freedom in any area of the world means a loss of freedom to ourselves- that the loss of independence by any nation adds directly to the insecurity of the United States and all free nations.

---

PROCESSING:1949-Truman.txt

3605:Eighth, to authorize an immediate study of the adequacy of production facilities for materials in critically short supply, such as steel; and, if found necessary, to authorize Government loans for the expansion of production facilities to relieve such shortages, and to authorize the construction of such facilities directly, if action by private industry fails to meet our needs.

3476:I explained to you at that time that the task before me was one of the greatest in the history of the world, and that it was necessary to have the complete cooperation of the Congress and the people of the United States.

---

PROCESSING:1950-Truman.txt

4643:These proposals are for the enactment of Federal statutes which will protect all our people in the exercise of their democratic rights and their search for economic opportunity, grant statehood to Alaska and Hawaii, provide a

greater measure of self- government for our island possessions, and accord home rule to the District of Columbia.

4593:During the past year, we have also made a good start in providing housing for low- income groups; we have raised minimum wages; we have gone forward with the development of our natural resources; we have given a greater assurance of stability to the farmer; and we have improved the organization and efficiency of our Government.

---

PROCESSING:1951-Truman.txt

3536:If the democracies had stood up against the invasion of Manchuria in 1931, or the attack on Ethiopia in 1935, or the seizure of Austria in 1938, if they had stood together again st aggression on those occasions as the United Nations has done in Korea, the whole history of our time would have been different.

3057:Our men are fighting, alongside their United Nations allies, because they know, as we do, that the aggression in Korea is part of the attempt of the Russian Communist dictatorship to take over the world, step by step.

---

PROCESSING:1953-Eisenhower.txt

10222:The grand labors of this leadership will involve: Application of America' s influence in world affairs with such fortitude and such foresight that it will deter aggression and eventually secure peace; Establishment of a national administration of such integrity and such efficiency that its honor at home will ensure respect abroad; Encouragement of those incentives that inspire creative initiative in our economy, so that its productivity may fortify freedom everywhere; and Dedication to the well- being of all our citizens and to the attainment of equality of opportunity for all, so that our Nation will ever act with the strength of unity in every task to which it is called.

10063:Our immediate task is to chart a fiscal and economic policy that can:( 1) Reduce the planned deficits and then balance the budget, which means, among other things, reducing Federal expenditures to the safe minimum;( 2) Meet the huge costs of our defense;( 3) Properly handle the burden of our inheritance of debt and obligations;( 4) Check the menace of inflation;( 5) Work toward the earliest possible reduction of the tax burden;( 6) Make constructive plans to encourage the initiative of our citizens.

---

PROCESSING:1954-Eisenhower.txt

8396:This program will include: Modernization of the home mortgage insurance program of the Federal Government; Redirection of the present system of loans and grants- in- aid to cities for slum clearance and redevelopment; Extension of

the advantages of insured lending to private credit engaged in this task of rehabilitating obsolete neighborhoods; Insurance of long- term, mortgage loans, with small down payment for low- income families; and, until alternative programs prove more effective, Continuation of the public housing program adopted in the Housing Act of 1949.

6634:Mr. President, Mr. Speaker, Members of the Eighty- third Congress: It is a high honor again to present to the Congress my views on the state of the Union and to recommend measures to advance the security, prosperity, and well- being of the American people.

---

PROCESSING:1955-Eisenhower.txt

8956:To reduce the gaps in these services, I shall propose: New measures to facilitate construction of needed health facilities and help reduce shortages of trained health personnel; Vigorous steps to combat the misery and national loss involved in mental illness; Improved services for crippled children and for maternal and child health; Better consumer protection under our existing pure food and drug laws; and, finally, Strengthened programs to combat the increasingly serious pollution of our rivers and streams and the growing problem of air pollution.

8844:From this deep faith have evolved three main purposes of our Federal Government: First, to maintain justice and freedom among ourselves and to champion them for others so that we may work effectively for enduring peace; Second, to help keep our economy vigorous and expanding, thus sustaining our international strength and assuring better jobs, better living, better opportunities for every citizen; And third, to concern ourselves with the human problems of our people so that every American may have the opportunity to lead a healthy, productive and rewarding life.

---

PROCESSING:1956-Eisenhower.txt

11463:In the last year, the free world has seen major gains for the system of collective security: the accession to the North Atlantic Treaty Organization and Western European Union of the sovereign Federal German Republic; the developing cooperation under the Southeast Asia Collective Defense Treaty; and the formation in the Middle East of the Baghdad Pact among Turkey, Iraq, Iran, Pakistan and the United Kingdom.

10061:With this purpose, we move in a wide variety of ways and through many agencies to remove the pall of fear; to strengthen the ties with our partners and to improve the cooperative cohesion of the free world; to reduce the burden of armaments, and to stimulate and inspire action among all nations for a world of justice and prosperity and peace.

-----  
PROCESSING:1957-Eisenhower.txt

5171:That program included:( 1) creation of a bipartisan commission to investigate asserted violations of civil rights and to make recommendations;( 2) creation of a civil rights division in the Department of Justice in charge of an Assistant Attorney General;( 3) enactment by the Congress of new laws to aid in the enforcement of voting rights; and( 4) amendment of the laws so as to permit the Federal Government to seek from the civil courts preventive relief in civil rights cases.

4396:These will involve our financial and fiscal affairs; our military and civil defenses; the administration of justice; our agricultural economy; our domestic and foreign commerce; the urgently needed increase in our postal rates; the development of our natural resources; our labor laws, including our labor-management relations legislation, and vital aspects of the health, education and welfare of our people.

-----  
PROCESSING:1958-Eisenhower.txt

4865:In the 1959 budget, increased expenditures for missiles, nuclear ships, atomic energy, research and development, science and education, a special contingency fund to deal with possible new technological discoveries, and increases in pay and incentives to obtain and retain competent manpower add up to a total increase over the comparable figures in the 1957 budget of about\ \$ 4 billion.

4297:But it is my conviction, based on close study of all relevant intelligence, that if we make the necessary effort, we will have the missiles, in the needed quantity and in time, to sustain and strengthen the deterrent power of our increasingly efficient bombers.

-----  
PROCESSING:1959-Eisenhower.txt

4329:It is my purpose to intensify efforts during the coming two years in seeking ways to supplement the procedures of the United Nations and other bodies with similar objectives, to the end that the rule of law may replace the rule of force in the affairs of nations.

4015:In this hope, I am requesting the Secretary of the Treasury to prepare appropriate proposals for revising, at the proper time, our tax structure, to remove inequities and to enhance incentives for all Americans to work, to save, and to invest.

-----

PROCESSING:1960-Eisenhower.txt

5732:Among current problems that require solution are: the need to protect the public interest in situations of prolonged labor- management stalemate; the persistent refusal to come to grips with a critical problem in one sector of American agriculture; the continuing threat of inflation, together with the persisting tendency toward fiscal irresponsibility; in certain instances the denial to some of our citizens of equal protection of the law.

4945:And America' s record of material accomplishment in freedom is written not only in the unparalleled prosperity of our own nation, but in the many billions we have devoted to the reconstruction of Free World economies wrecked by World War II and in the effective help of many more billions we have given in saving the independence of many others threatened by outside domination.

-----  
PROCESSING:1961-Kennedy.txt

19429:Finally, our greatest asset in this struggle is the American people -- their willingness to pay the price for these programs -- to understand and accept a long struggle -- to share their resources with other less fortunate people -- to meet the tax levels and close the tax loopholes I have requested -- to exercise self- restraint instead of pushing up wages or prices, or over-producing certain crops, or spreading military secrets, or urging unessential expenditures or improper monopolies or harmful work stoppages -- to serve in the Peace Corps or the Armed Services or the Federal Civil Service or the Congress -- to strive for excellence in their schools, in their cities and in their physical fitness and that of their children -- to take part in Civil Defense -- to pay higher postal rates, and higher payroll taxes and higher teachers' salaries, in order to strengthen our society -- to show friendship to students and visitors from other lands who visit us and go back in many cases to be the future leaders, with an image of America -- and I want that image, and I know you do, to be affirmative and positive -- and, finally, to practice democracy at home, in all States, with all races, to respect each other and to protect the Constitutional rights of all citizens.

10528:Earlier in the year, I outlined to the Congress a new program for aiding emerging nations; and it is my intention to transmit shortly draft legislation to implement this program, to establish a new Act for International Development, and to add to the figures previously requested, in view of the swift pace of critical events, an additional 250 million dollars for a Presidential Contingency Fund, to be used only upon a Presidential determination in each case, with regular and complete reports to the Congress in each case, when there is a sudden and extraordinary drain upon our regular funds which we cannot foresee -- as illustrated by recent events in Southeast Asia -- and it makes necessary the use of this emergency reserve.

-----  
PROCESSING:1962-Kennedy.txt

15316:To expand our growth and job opportunities, I urge on the Congress three measures: First, the Manpower Training and Development Act, to stop the waste of able-bodied men and women who want to work, but whose only skill has been replaced by a machine, or moved with a mill, or shut down with a mine; Second, the Youth Employment Opportunities Act, to help train and place not only the one million young Americans who are both out of school and out of work, but the twenty-six million young Americans entering the labor market in this decade; and Third, the 8 percent tax credit for investment in machinery and equipment, which, combined with planned revisions of depreciation allowances, will spur our modernization, our growth, and our ability to compete abroad.

11231:Our farm surpluses- our balance of trade, as you all know, to Europe, the Common Market, in farm products, is nearly three or four to one in our favor, amounting to one of the best earners of dollars in our balance of payments structure, and without entrance to this Market, without the ability to enter it, our farm surpluses will pile up in the Middle West, tobacco in the South, and other commodities, which have gone through Western Europe for 15 years.

-----  
PROCESSING:1963-Johnson.txt

3893:From this chamber of representative government, let all the world know and none misunderstand that I rededicate this Government to the unswerving support of the United Nations, to the honorable and determined execution of our commitments to our allies, to the maintenance of military strength second to none, to the defense of the strength and the stability of the dollar, to the expansion of our foreign trade, to the reinforcement of our programs of mutual assistance and cooperation in Asia and Africa, and to our Alliance for Progress in this hemisphere.

3004:The dream of conquering the vastness of space -- the dream of partnership across the Atlantic -- and across the Pacific as well -- the dream of a Peace Corps in less developed nations -- the dream of education for all of our children -- the dream of jobs for all who seek them and need them -- the dream of care for our elderly -- the dream of an all-out attack on mental illness -- and above all, the dream of equal rights for all Americans, whatever their race or color -- these and other American dreams have been vitalized by his drive and by his dedication.

-----  
PROCESSING:1963-Kennedy.txt

9559:To achieve this reduction within the limits of a manageable budgetary deficit, I urge: first, that these cuts be phased over 3 calendar years, beginning in 1963 with a cut of some \$ 6 billion at annual rates; second, that

these reductions be coupled with selected structural changes, beginning in 1964, which will broaden the tax base, end unfair or unnecessary preferences, remove or lighten certain hardships, and in the net offset some\\$ 3. 5 billion of the revenue loss; and third, that budgetary receipts at the outset be increased by\\$ 1. 5 billion a year, without any change in tax liabilities, by gradually shifting the tax payments of large corporations to a more current time schedule.

8588:If they come to realize that their ambitions cannot succeed- if they see their" wars of liberation" and subversion will ultimately fail- if they recognize that there is more security in accepting inspection than in permitting new nations to master the black arts of nuclear war- and if they are willing to turn their energies, as we are, to the great unfinished tasks of our own peoples- then, surely, the areas of agreement can be very wide indeed: a clear understanding about Berlin, stability in Southeast Asia, an end to nuclear testing, new checks on surprise or accidental attack, and, ultimately, general and complete disarmament.

---

PROCESSING:1964-Johnson.txt

5306:Let this session of Congress be known as the session which did more for civil rights than the last hundred sessions combined; as the session which enacted the most far- reaching tax cut of our time; as the session which declared all- out war on human poverty and unemployment in these United States; as the session which finally recognized the health needs of all our older citizens; as the session which reformed our tangled transportation and transit policies; as the session which achieved the most effective, efficient foreign aid program ever; and as the session which helped to build more homes, more schools, more libraries, and more hospitals than any single session of Congress in the history of our Republic.

3403:But whatever the cause, our joint Federal- local effort must pursue poverty, pursue it wherever it exists -- in city slums and small towns, in sharecropper shacks or in migrant worker camps, on Indian Reservations, among whites as well as Negroes, among the young as well as the aged, in the boom towns and in the depressed areas.

---

PROCESSING:1965-Johnson-1.txt

7576:Let a just nation throw open to them the city of promise: -- to the elderly, by providing hospital care under social security and by raising benefit payments to those struggling to maintain the dignity of their later years; -- to the poor and the unfortunate, through doubling the war against poverty this year; -- to Negro Americans, through enforcement of the civil rights law and elimination of barriers to the right to vote; -- to those in other lands that are seeking the promise of America, through an immigration law based on the work

a man can do and not where he was born or how he spells his name.

4957:INCREASED PROSPERITY We can help insure continued prosperity through: -- a regional recovery program to assist the development of stricken areas left behind by our national progress; -- further efforts to provide our workers with the skills demanded by modern technology, for the laboring- man is an indispensable force in the American system; -- the extension of the minimum wage to more than 2 million unprotected workers; -- the improvement and the modernization of the unemployment compensation system.

---

PROCESSING:1965-Johnson-2.txt

5311:And so at the request of your beloved Speaker and the Senator from Montana; the majority leader, the Senator from Illinois; the minority leader, Mr. McCulloch, and other Members of both parties, I came here tonight -- not as President Roosevelt came down one time in person to veto a bonus bill, not as President Truman came down one time to urge the passage of a railroad bill -- but I came down here to ask you to share this task with me and to share it with the people that we both work for.

4304:As we meet here in this peaceful, historic chamber tonight, men from the South, some of whom were at Iwo Jima, men from the North who have carried Old Glory to far corners of the world and brought it back without a stain on it, men from the East and from the West, are all fighting together without regard to religion, or color, or region, in Viet- Nam.

---

PROCESSING:1966-Johnson.txt

8452:Therefore, I will submit legislation to revise the present unrealistic restriction on contributions -- to prohibit the endless proliferation of committees, bringing local and State committees under the act -- to attach strong teeth and severe penalties to the requirement of full disclosure of contributions -- and to broaden the participation of the people, through added tax incentives, to stimulate small contributions to the party and to the candidate of their choice.

7030:The people of Vietnam, North and South, seek the same things: the shared needs of man, the needs for food and shelter and education -- the chance to build and work and till the soil, free from the arbitrary horrors of battle -- the desire to walk in the dignity of those who master their own destiny.

---

PROCESSING:1967-Johnson.txt

9399:-- 90 percent of the cost of developing the State and local plans, master plans, to combat crime in their area; -- 60 percent of the cost of training new



tactical units, developing instant communications and special alarm systems, and introducing the latest equipment and techniques so that they can become weapons in the war on crime; -- 50 percent of the cost of building crime laboratories and police academy- type centers so that our citizens can be protected by the best trained and served by the best equipped police to be found anywhere.

7577:For all the disorders that we must deal with, and all the frustrations that concern us, and all the anxieties that we are called upon to resolve, for all the issues we must face with the agony that attends them, let us remember that" those who expect to reap the blessings of freedom must, like men, undergo the fatigues of supporting it ."

---

PROCESSING:1968-Johnson.txt

5884:The unfinished work of the first session must be completed -- the Higher Education Act, the Juvenile Delinquency Act, conservation measures to save the redwoods of California, and to preserve the wonders of our scenic rivers, the Highway Beautification Act -- and all the other measures for a cleaner, and for a better, and for a more beautiful America.

5500:Virtually all of this expenditure increase represents the mandatory cost of our defense efforts,\\$ 3 billion; increased interest, almost\\$ 1 billion; or mandatory payments under laws passed by Congress -- such as those provided in the Social Security Act that you passed in 1967, and to Medicare and Medicaid beneficiaries, veterans, and farmers, of about\\$ 4 1/ 2 billion; and the additional\\$ 1 billion 600 million next year for the pay increases that you passed in military and civilian pay.

---

PROCESSING:1969-Johnson.txt

5913:Urban unrest, poverty, pressures on welfare, education of our people, law enforcement and law and order, the continuing crisis in the Middle East, the conflict in Vietnam, the dangers of nuclear war, the great difficulties of dealing with the Communist powers, all have this much in common: They and their causes -- the causes that gave rise to them -- all of these have existed with us for many years.

5831:I want to tell you that both of us want to see it removed just as soon as circumstances will permit, but the President- elect has told me that he has concluded that until his administration, and this Congress, can examine the appropriation bills, and each item in the budget, and can ascertain that the facts justify permitting the surtax to expire or to be reduced, he, Mr. Nixon, will support my recommendation that the surtax be continued.

---

PROCESSING:1970-Nixon.txt

5649:In the majesty of this great Chamber we hear the echoes of America's history, of debates that rocked the Union and those that repaired it, of the summons to war and the search for peace, of the uniting of the people, the building of a nation.

5442:The State of the Union Address is traditionally an occasion for a lengthy and detailed account by the President of what he has accomplished in the past, what he wants the Congress to do in the future, and, in an election year, to lay the basis for the political issues which might be decisive in the fall.

-----  
PROCESSING:1971-Nixon.txt

7954:With the stimulus and the discipline of a full employment budget, with the commitment of the independent Federal Reserve System to provide fully for the monetary needs of a growing economy, and with a much greater effort on the part of labor and management to make their wage and price decisions in the light of the national interest and their own self-interest then for the worker, the farmer, the consumer, for Americans everywhere we shall gain the goal of a new prosperity: more jobs, more income, more profits, without inflation and without war.

7179:The White House also released the transcripts of three news briefings on the President's State of the Union proposals: the first, on January 25, 1971, by Senator Hugh Scott and Representative Gerald R. Ford following a Republican Congressional leadership meeting with the President; the second, on January 27, by John D. Khrlichman, Assistant to the President for Domestic Affairs; and the third, on February 2, by Secretary of Commerce Maurice H. Stans, Secretary of Labor James D. Hodgson, and Secretary of Agriculture Clifford M. Hardin following a meeting of the Cabinet.

-----  
PROCESSING:1972-Nixon.txt

5506:They include, among others, our programs to improve life for the aging; to combat crime and drug abuse; to improve health services and to ensure that no one will be denied needed health care because of inability to pay; to protect workers' pension rights; to promote equal opportunity for members of minorities, and others who have been left behind; to expand consumer protection; to improve the environment; to revitalize rural America; to help the cities; to launch new initiatives in education; to improve transportation, and to put an end to costly labor tie-ups in transportation.

4556:I am presenting only vital programs which are within the capacity of this Congress to enact, within the capacity of the budget to finance, and which I believe should be above partisanship programs which deal with urgent priorities

for the Nation, which should and must be the subject of bipartisan action by this Congress in the interests of the country in 1972.

---

PROCESSING:1973-Nixon.txt

1911:The role of the Federal Government as we approach our third century of independence should not be to dominate any facet of American life, but rather to aid and encourage people, communities and institutions to deal with as many of the difficulties and challenges facing them as possible, and to help see to it that every American has a full and equal opportunity to realize his or her potential.

1606:In the field of economic affairs, our objectives will be to hold down taxes, to continue controlling inflation, to promote economic growth, to increase productivity, to encourage foreign trade, to keep farm income high, to bolster small business, and to promote better labor- management relations.

---

PROCESSING:1974-Nixon.txt

7819:But my colleagues, this I believe: With the help of God, who has blessed this land so richly, with the cooperation of the Congress, and with the support of the American people, we can and we will make the year 1974 a year of unprecedented progress toward our goal of building a structure of lasting peace in the world and a new prosperity without war in the United States of America.

7473:Because of this, in the 22, 000- word message on the state of the Union that I have just handed to the Speaker of the House and the President of the Senate, I have been able to deal primarily with the problems of peaceAXwith what we can do here at home in America for the American peopleAXrather than with the problems of war.

---

PROCESSING:1975-Ford.txt

5019:Mr. Speaker, Mr. Vice President, Members of the 94th Congress, and distinguished guests: Twenty- six years ago, a freshman Congressman, a young fellow with lots of idealism who was out to change the world, stood before Sam Rayburn in the well of the House and solemnly swore to the same oath that all of you took yesterday -- an unforgettable experience, and I congratulate you all.

3639:Within the next 10 years, my program envisions: 200 major nuclear powerplants; 250 major new coal mines; 150 major coal- fired powerplants; 30 major new[ oil] refineries; 20 major new synthetic fuel plants; the drilling of many thousands of new oil wells; the insulation of 18 million homes; and the manufacturing and the sale of millions of new automobiles, trucks, and buses that use much less fuel.

-----  
PROCESSING:1976-Ford.txt

4361:My proposals, as all of you know, would reduce domestic natural gas shortages; allow production from Federal petroleum reserves; stimulate effective conservation, including revitalization of our railroads and the expansion of our urban transportation systems; develop more and cleaner energy from our vast coal resources; expedite clean and safe nuclear power production; create a new national energy independence authority to stimulate vital energy investment; and accelerate development of technology to capture energy from the Sun and the Earth for this and future generations.

3526:Like our forefathers, we know that if we meet the challenges of our own time with a common sense of purpose and conviction, if we remain true to our Constitution and to our ideals, then we can know that the future will be better than the past.

-----  
PROCESSING:1977-Ford.txt

6578:From the opportunities which fate and my fellow citizens have given me, as a Member of the House, as Vice President and President of the Senate, and as President of all the people, I have come to understand and place the highest value on the checks and balances which our founders imposed on government through the separation of powers among co- equal legislative, executive, and judicial branches.

5944:There are no soldiers marching in the street except in the Inaugural Parade; no public demonstrations except for some of the dancers at the Inaugural Ball; the opposition party doesn' t go underground, but goes on functioning vigorously in the Congress and in the country; and our vigilant press goes right on probing and publishing our faults and our follies, confirming the wisdom of the framers of the first amendment.

-----  
PROCESSING:1978-Carter.txt

3809:The foundations of this partnership are truth, the courage to face hard decisions, concern for one another and the common good over special interests, and a basic faith and trust in the wisdom and strength and judgment of the American people.

3674:As President, I' ve had to ask you, the Members of Congress, and you, the American people, to come to grips with some of the most difficult and hard questions facing our society.

-----

PROCESSING:1979-Carter.txt

4101:The purpose of SALT, then as now, is not to gain a unilateral advantage for either nation, but to protect the security of both nations, to reverse the costly and dangerous momentum of the nuclear arms race, to preserve a stable balance of nuclear forces, and to demonstrate to a concerned world that we are determined to help preserve the peace.

3688:We' ve demonstrated in this restrained budget that we can build on the gains of the past 2 years to provide additional support to educate disadvantaged children, to care for the elderly, to provide nutrition and legal services for the poor, and to strengthen the economic base of our urban communities and, also, our rural areas.

---

PROCESSING:1980-Carter.txt

8878:We continue to pursue these specific goals: first, to protect the present and long- range interests of the United States; secondly, to preserve the lives of the American hostages and to secure, as quickly as possible, their safe release, if possible, to avoid bloodshed which might further endanger the lives of our fellow citizens; to enlist the help of other nations in condemning this act of violence, which is shocking and violates the moral and the legal standards of a civilized world; and also to convince and to persuade the Iranian leaders that the real danger to their nation lies in the north, in the Soviet Union and from the Soviet troops now in Afghanistan, and that the unwarranted Iranian quarrel with the United States hampers their response to this far greater danger to them.

5398:We move into the 1980' s with confidence and hope and a bright vision of the America we want: an America strong and free, an America at peace, an America with equal rights for all citizens -- and for women, guaranteed in the United States Constitution -- an America with jobs and good health and good education for every citizen, an America with a clean and bountiful life in our cities and on our farms, an America that helps to feed the world, an America secure in filling its own energy needs, an America of justice, tolerance, and compassion.

---

PROCESSING:1981-Reagan.txt

4029:Adding to our troubles is a mass of regulations imposed on the shopkeeper, the farmer, the craftsman, professionals, and major industry that is estimated to add\ \$ 100 billion to the price of the things we buy, and it reduces our ability to produce.

3755:I' m well aware that there are many other desirable and needed tax changes, such as indexing the income tax brackets to protect taxpayers against inflation; the unjust discrimination against married couples if both are working and

earning; tuition tax credits; the unfairness of the inheritance tax, especially to the family- owned farm and the family- owned business; and a number of others.

---

PROCESSING:1982-Reagan.txt

5006:It' s my duty to report to you tonight on the progress that we have made in our relations with other nations, on the foundation we' ve carefully laid for our economic recovery, and finally, on a bold and spirited initiative that I believe can change the face of American government and make it again the servant of the people.

4938:What we do and say here will make all the difference to autoworkers in Detroit, lumberjacks in the Northwest, steelworkers in Steubenville who are in the unemployment lines; to black teenagers in Newark and Chicago; to hard-pressed farmers and small businessmen; and to millions of everyday Americans who harbor the simple wish of a safe and financially secure future for their children.

---

PROCESSING:1983-Reagan.txt

6116:America' s leadership in the world came to us because of our own strength and because of the values which guide us as a society: free elections, a free press, freedom of religious choice, free trade unions, and above all, freedom for the individual and rejection of the arbitrary power of the state.

5560:Nearly half a century ago, in this Chamber, another American President, Franklin Delano Roosevelt, in his second State of the Union message, urged America to look to the future, to meet the challenge of change and the need for leadership that looks forward, not backward.

---

PROCESSING:1984-Reagan.txt

6240:Mr. Speaker, Mr. President, distinguished Members of the Congress, honored guests, and fellow citizens: Once again, in keeping with time- honored tradition, I have come to report to you on the state of the Union, and I' m pleased to report that America is much improved, and there' s good reason to believe that improvement will continue through the days to come.

4334:We can establish a more stable basis for peaceful relations with the Soviet Union; strengthen allied relations across the board; achieve real and equitable reductions in the levels of nuclear arms; reinforce our peacemaking efforts in the Middle East, Central America, and southern Africa; or assist developing countries, particularly our neighbors in the Western Hemisphere; and assist in the development of democratic institutions throughout the world.

-----  
PROCESSING:1985-Reagan.txt

4964:Mr. Speaker, Mr. President, distinguished Members of the Congress, honored guests, and fellow citizens: I come before you to report on the state of our Union, and I'm pleased to report that after 4 years of united effort, the American people have brought forth a nation renewed, stronger, freer, and more secure than before.

4393:New freedom in our lives has planted the rich seeds for future success: For an America of wisdom that honors the family, knowing that if[ as] the family goes, so goes our civilization; For an America of vision that sees tomorrow's dreams in the learning and hard work we do today; For an America of courage whose service men and women, even as we meet, proudly stand watch on the frontiers of freedom; For an America of compassion that opens its heart to those who cry out for help.

-----  
PROCESSING:1986-Reagan.txt

3435:This week I will send you our detailed proposals; tonight let us speak of our responsibility to redefine government's role: not to control, not to demand or command, not to contain us, but to help in times of need and, above all, to create a ladder of opportunity to full employment so that all Americans can climb toward economic power and justice on their own.

3414:Further, after seeing how devastating illness can destroy the financial security of the family, I am directing the Secretary of Health and Human Services, Dr. Otis Bowen, to report to me by year end with recommendations on how the private sector and government can work together to address the problems of affordable insurance for those whose life savings would otherwise be threatened when catastrophic illness strikes.

-----  
PROCESSING:1987-Reagan.txt

4118:Even today, though nearly one in three Soviet families is without running hot water, and the average family spends two hours a day shopping for the basic necessities of life, their government still found the resources to transfer\ \$ 75 billion in weapons to client states in the past five years -- clients like Syria, Vietnam, Cuba, Libya, Angola, Ethiopia, Afghanistan, and Nicaragua.

2901:And the little ones especially so fresh-faced and so eager to know- well, from time to time I've been with them, they will ask about our Constitution, and I hope you Members of Congress will not deem this a breach of protocol if you'll permit me to share these thoughts again with the young people who might be listening or watching this evening.

-----  
PROCESSING:1988-Reagan.txt

7711:Well, these ideas were part of a larger notion -- a vision, if you will, of America herself -- an America not only rich in opportunity for the individual but an America, too, of strong families and vibrant neighborhoods, an America whose divergent but harmonizing communities were a reflection of a deeper community of values -- the value of work, of family, of religion -- and of the love of freedom that God places in each of us and whose defense He has entrusted in a special way to this nation.

5986:It means, too, that the young Americans I spoke of seven years ago, as well as those who might be coming along the Virginia or Maryland shores this night and seeing for the first time the lights of this capital city, the lights that cast their glow on our great halls of government and the monuments to the memory of our great men -- it means those young Americans will find a city of hope in a land that is free.

-----  
PROCESSING:1989-Bush.txt

7910:So, tonight I' m proposing the following initiatives: the beginning of a\ \$ 500 million program to reward America' s best schools, merit schools; the creation of special Presidential awards for the best teachers in every State, because excellence should be rewarded; the establishment of a new program of National Science Scholars, one each year for every Member of the House and Senate, to give this generation of students a special incentive to excel in science and mathematics; the expanded use of magnet schools, which give families and students greater choice; and a new program to encourage alternative certification, which will let talented people from all fields teach in our classrooms.

5732:Mr. Speaker, Mr. President, and distinguished Members of the House and Senate, honored guests, and fellow citizens: Less than 3 weeks ago, I joined you on the West Front of this very building and, looking over the monuments to our proud past, offered you my hand in filling the next page of American history with a story of extended prosperity and continued peace.

-----  
PROCESSING:1990-Bush.txt

12008:And our challenge today is to take this democratic system of ours, a system second to none, and make it better: a better America, where there' s a job for everyone who wants one; where women working outside the home can be confident their children are in safe and loving care and where government works to expand child- care alternatives for parents; where we reconcile the needs of a clean environment and a strong economy; where ' ' Made in the USA' ' is



recognized around the world as the symbol of quality and progress; where every one of us enjoys the same opportunities to live, to work, and to contribute to society and where, for the first time, the American mainstream includes all of our disabled citizens; where everyone has a roof over his head and where the homeless get the help they need to live in dignity; where our schools challenge and support our kids and our teachers and where all of them make the grade; where every street, every city, every school, and every child is drug- free; and finally, where no American is forgotten -- our hearts go out to our hostages who are ceaselessly on our minds and in our efforts.

4476:This administration is determined to encourage the creation of capital, capital of all kinds: physical capital -- everything from our farms and factories to our workshops and production lines, all that is needed to produce and deliver quality goods and quality services; intellectual capital -- the source of ideas that spark tomorrow' s products; and of course our human capital -- the talented work force that we' ll need to compete in the global market.

---

PROCESSING:1991-Bush-1.txt

9605:We' ve prepared a detailed series of proposals that include: a budget that promotes investment in America' s future -- in children, education, infrastructure, space, and high technology; legislation to achieve excellence in education, building on the partnership forged with the 50 Governors at the education summit, enabling parents to choose their children' s schools and helping to make America number one in math and science; a blueprint for a new national highway system, a critical investment in our transportation infrastructure; a research and development agenda that includes record levels of Federal investment, and a permanent tax credit to strengthen private R& D and to create jobs; a comprehensive national energy strategy that calls for energy conservation and efficiency, increased development, and greater use of alternative fuels; a banking reform plan to bring America' s financial system into the 21st century so that our banks remain safe and secure and can continue to make job- creating loans for our factories, our businesses, and home buyers.

4478:The courage and success of the RAF pilots, of the Kuwaiti, Saudi, French, the Canadians, the Italians, the pilots of Qatar and Bahrain -- all are proof that for the first time since World War II, the international community is united.

---

PROCESSING:1991-Bush-2.txt

3129:In a very real sense, this victory belongs to them -- to the privates and the pilots, to the sergeants and the supply officers, to the men and women in the machines and the men and women who made them work.

2567:But I wish to depart from tradition tonight and express to you on behalf of the Congress and the country, and through you to the members of our Armed Forces, our warmest congratulations on the brilliant victory of the Desert Storm Operation.

---

PROCESSING:1992-Bush.txt

4219:I' ll tell you, those of you who say," Oh, no, someone who' s comfortable may benefit from that ," you kind of remind me of the old definition of the Puritan who couldn' t sleep at night, worrying that somehow, someone somewhere was out having a good time.

4130:States throughout the country are beginning to operate with new assumptions that when able-bodied people receive Government assistance, they have responsibilities to the taxpayer: A responsibility to seek work, education, or job training; a responsibility to get their lives in order; a responsibility to hold their families together and refrain from having children out of wedlock; and a responsibility to obey the law.

---

PROCESSING:1993-Clinton.txt

9140:And finally, I propose a program that got a great response from the American people all across this country last year: a program of national service to make college loans available to all Americans and to challenge them at the same time to give something back to their country as teachers or police officers or community service workers; to give them the option to pay the loans back, but at tax time so they can' t beat the bill, but to encourage them instead to pay it back by making their country stronger and making their country better and giving us the benefit of their knowledge.

8401:As we move toward the 4th year, with the explosion in health care costs, as I said, projected to account for 50 percent of the growth of the deficit between now and the year 2000, there must be planned cuts in payments to providers, to doctors, to hospitals, to labs, as a way of controlling health care costs.

---

PROCESSING:1994-Clinton.txt

8742:But we all know, too, we can do all these things, put our economic house in order, expand world trade, target the jobs of the future, guarantee equal opportunity, but if we' re honest, we' ll all admit that this strategy still cannot work unless we also give our people the education, training, and skills they need to seize the opportunities of tomorrow.

8577:We can' t renew our country unless more of us -- I mean, all of us -- are willing to join the churches and the other good citizens, people like all the

ministers I' ve worked with over the years or the priests and the nuns I met at Our Lady of Help in east Los Angeles or my good friend Tony Campollo in Philadelphia, unless we' re willing to work with people like that, people who are saving kids, adopting schools, making streets safer.

---

PROCESSING:1995-Clinton.txt

16263:It has fallen to every generation since then to preserve that idea, the American idea, and to deepen and expand its meaning in new and different times: to Lincoln and to his Congress to preserve the Union and to end slavery; to Theodore Roosevelt and Woodrow Wilson to restrain the abuses and excesses of the industrial revolution and to assert our leadership in the world; to Franklin Roosevelt to fight the failure and pain of the Great Depression and to win our country' s great struggle against fascism; and to all our Presidents since to fight the cold war.

10959:We propose to collapse nearly 70 Federal programs and not give the money to the States but give the money directly to the American people, offer vouchers to them so that they, if they' re laid off or if they' re working for a very low wage, can get a voucher worth\ \$ 2, 600 a year for up to 2 years to go to their local community colleges or wherever else they want to get the skills they need to improve their lives.

---

PROCESSING:1996-Clinton.txt

6859:So all of us, not just as parents but all of us in our other roles -- our media, our schools, our teachers, our communities, our churches and synagogues, our businesses, our governments -- all of us have a responsibility to help our children to make it and to make the most of their lives and their God- given capacities.

6424:On behalf of all Americans, especially those who need their Social Security payments at the beginning of March, I also challenge the Congress to preserve the full faith and credit of the United States, to honor the obligations of this great Nation as we have for 220 years, to rise above partisanship and pass a straightforward extension of the debt limit and show people America keeps its word.

---

PROCESSING:1997-Clinton.txt

9916:So tonight I issue a call to action: action by this Congress, action by our States, by our people, to prepare America for the 21st century; action to keep our economy and our democracy strong and working for all our people; action to strengthen education and harness the forces of technology and science; action to build stronger families and stronger communities and a safer environment; action

to keep America the world' s strongest force for peace, freedom, and prosperity; and above all, action to build a more perfect Union here at home.

8236:We should challenge all Americans in the arts and humanities to join with our fellow citizens to make the year 2000 a national celebration of the American spirit in every community, a celebration of our common culture in the century that has passed and in the new one to come in the new millennium, so that we can remain in the world' s beacon not only of liberty but of creativity, long after the fireworks have faded.

---

PROCESSING:1998-Clinton.txt

11697:I ask all Americans to support our project to restore all our treasures so that the generations of the 21st century can see for themselves the images and the words that are the old and continuing glory of America; an America that has continued to rise through every age, against every challenge, a people of great works and greater possibilities, who have always, always found the wisdom and strength to come together as one nation -- to widen the circle of opportunity, to deepen the meaning of our freedom, to form that more perfect Union.

11097:It is a time to build, to build the America within reach, an America where everybody has a chance to get ahead with hard work; where every citizen can live in a safe community; where families are strong, schools are good, and all our young people can go on to college; an America where scientists find cures for diseases, from diabetes to Alzheimer' s to AIDS; an America where every child can stretch a hand across a keyboard and reach every book ever written, every painting ever painted, every symphony ever composed; where government provides opportunity and citizens honor the responsibility to give something back to their communities; an America which leads the world to new heights of peace and prosperity.

---

PROCESSING:1999-Clinton.txt

7179:But I want all the folks at home listening to this to know that we need every State and local government, every business, large and small, to work with us to make sure that this Y2K computer bug will be remembered as the last headache of the 20th century, not the first crisis of the 21st.

7088:Once again, a Government that is a progressive instrument of the common good, rooted in our oldest values of opportunity, responsibility, and community; devoted to fiscal responsibility; determined to give our people the tools they need to make the most of their own lives in the 21st century; a 21st century Government for 21st century America.

---

PROCESSING:2000-Clinton.txt

14085:There' s a new spirit of service in America, a movement we' ve tried to support with AmeriCorps, expanded Peace Corps, unprecedented new partnerships with businesses, foundations, community groups; partnerships, for example, like the one that enlisted 12, 000 companies which have now moved 650, 000 of our fellow citizens from welfare to work; partnerships to battle drug abuse, AIDS, teach young people to read, save America' s treasures, strengthen the arts, fight teen pregnancy, prevent violence among young people, promote racial healing.

9547:I ask you to pass a national security budget that keeps our military the best trained and best equipped in the world, with heightened readiness and 21st century weapons, which raises salaries for our service men and women, which protects our veterans, which fully funds the diplomacy that keeps our soldiers out of war, which makes good on our commitment to our U. N. dues and arrears.

-----  
PROCESSING:2001-GWBush-1.txt

3680:Our budget provides more than\ 700 million over the next 10 years for a federal compassion capital fund, with a focused and noble mission, to provide a mentor to the more than 100 million children with a parent in prison, and to support other local efforts to fight illiteracy, teen pregnancy, drug addiction and other difficult problems.

3185:Forty years ago, and then 20 years ago, two Presidents, one Democrat, one Republican, John F. Kennedy and Ronald Reagan, advocated tax cuts to, in President Kennedy' s words, get this country moving again.

-----  
PROCESSING:2001-GWBush-2.txt

2685:We will direct every resource at our command -- every means of diplomacy, every tool of intelligence, every instrument of law enforcement, every financial influence, and every necessary weapon of war -- to the disruption and to the defeat of the global terror network.

2465:THE PRESIDENT: Mr. Speaker, Mr. President Pro Tempore, members of Congress, and fellow Americans: In the normal course of events, Presidents come to this chamber to report on the state of the Union.

-----  
PROCESSING:2002-GWBush.txt

4134:In four short months, our nation has comforted the victims, begun to rebuild New York and the Pentagon, rallied a great coalition, captured, arrested, and rid the world of thousands of terrorists, destroyed Afghanistan' s terrorist training camps, saved a people from starvation, and freed a country

from brutal oppression.

3194:I' m a proud member of my party -- yet as we act to win the war, protect our people, and create jobs in America, we must act, first and foremost, not as Republicans, not as Democrats, but as Americans.

---

PROCESSING:2003-GWBush.txt

6898:In the ruins of two towers, at the western wall of the Pentagon, on a field in Pennsylvania, this Nation made a pledge, and we renew that pledge tonight: Whatever the duration of this struggle, and whatever the difficulties, we will not permit the triumph of violence in the affairs of men ... free people will set the course of history.

4873:America is working with the countries of the region- South Korea, Japan, China, and Russia- to find a peaceful solution, and to show the North Korean government that nuclear weapons will bring only isolation, economic stagnation, and continued hardship.

---

PROCESSING:2004-GWBush.txt

7817:This particular criticism is hard to explain to our partners in Britain, Australia, Japan, South Korea, the Philippines, Thailand, Italy, Spain, Poland, Denmark, Hungary, Bulgaria, Ukraine, Romania, the Netherlands --( applause) -- Norway, El Salvador, and the 17 other countries that have committed troops to Iraq.

5139:Since we last met in this chamber, combat forces of the United States, Great Britain, Australia, Poland and other countries enforced the demands of the United Nations, ended the rule of Saddam Hussein, and the people of Iraq are free.

---

PROCESSING:2005-GWBush.txt

4470:In the past four years, we provided tax relief to every person who pays income taxes, overcome a recession, opened up new markets abroad, prosecuted corporate criminals, raised homeownership to its highest level in history, and in the last year alone, the United States has added 2. 3 million new jobs.

4173:We' ve created a new department of government to defend our homeland, focused the FBI on preventing terrorism, begun to reform our intelligence agencies, broken up terror cells across the country, expanded research on defenses against biological and chemical attack, improved border security, and trained more than a half- million first responders.

-----  
PROCESSING:2006-GWBush.txt

4864:Mr. Speaker, Vice President Cheney, members of Congress, members of the Supreme Court and diplomatic corps, distinguished guests, and fellow citizens: Today our nation lost a beloved, graceful, courageous woman who called America to its founding ideals and carried on a noble dream.

4636:And we do not forget the other half -- in places like Syria and Burma, Zimbabwe, North Korea, and Iran -- because the demands of justice, and the peace of this world, require their freedom, as well.

-----

## Jurafsky and Martin

## 2 Regular Expressions and Automata

1. Write regular expressions.

**Answer:**

-----  
1. the set of all alphabetic strings

REGEX :`^[a-zA-Z]+\`

STRING:mad gab bleb blab BLEB BLAB MAD GABBB FUB fub

RESULT:None

\*\*\*\*\*

REGEX :`^[a-zA-Z]+\`

STRING:adsfyqwoeiruyqwoeiruyzlsdkmnvbzxcmvnb

RESULT:adsfyqwoeiruyqwoeiruyzlsdkmnvbzxcmvnb

\*\*\*\*\*

REGEX :`^[a-zA-Z]+\`

STRING:adsfyqwoeiruyqwoeiruyzlsdkmnvbzxc,mvnb

RESULT:None

\*\*\*\*\*  
-----

2. the set of all lower case alphabetic strings ending in 'b'

REGEX :`^[a-zA-Z]*b\`

STRING:Four score and seven years ago our fathers brought forth on this continent, a

RESULT:None

\*\*\*\*\*

REGEX :`^[a-zA-Z]*b\`

STRING:adsfyqwoeiruyqwoeiruyzlsdkmnvbzxcmvnb

RESULT:adsfyqweoiruyqwoeiruyzlsdkmnvbzxcmvnb

\*\*\*\*\*

REGEX :^[a-zA-Z]\*b\$

STRING:adsfyqweoiruyqwoeiruyzlsdkmnvbzxc,mvnb

RESULT:None

\*\*\*\*\*

-----  
3. the set of all strings with two consecutive repeated words (e.g. "Humbert Humbert" and "the the" but not "the bug" or "the big bug")

REGEX :(\w+)\W\1

STRING:Four score and seven years ago our fathers brought forth on this continent, a

RESULT:None

\*\*\*\*\*

REGEX :(\w+)\W\1

STRING:banana banana

RESULT:banana banana

\*\*\*\*\*

REGEX :(\w+)\W\1

STRING:Humbert Humbert

RESULT:Humbert Humbert

\*\*\*\*\*

REGEX :(\w+)\W\1

STRING:Humbert banana

RESULT:None

\*\*\*\*\*

REGEX :(\w+)\W\1

STRING:banana banana

RESULT:banana banana

\*\*\*\*\*

REGEX :(\w+)\W\1

STRING:Humbert banana

RESULT:None

\*\*\*\*\*

-----  
4. the set of all strings from the alphabet a,b such that each a is immediately preceded by and immediately followed by a b

REGEX :(bab)+

STRING:mad gab bleb blab BLEB BLAB MAD GABBB FUB fub

RESULT:None

\*\*\*\*\*

REGEX :(bab)+

STRING:adsfyqweoiruyqwoeiruyzlsdkmnvbzxcmvnb

RESULT:None

\*\*\*\*\*

REGEX :(bab)+



```

STRING:aba
RESULT:None
*****
REGEX : (bab)+
STRING:bab
RESULT:bab
*****
REGEX : (bab)+
STRING:bbbbbbbabbbbbbb
RESULT:None
*****
REGEX : (bab)+
STRING:babbabbabbabbabbabbabbabbabbabbabbabbabbabbabbab
RESULT:babbabbabbabbabbabbabbabbabbabbabbabbabbabbabbab
*****

```

-----

5. all strings that start at the beginning of the line with an integer and that end at the end of the line with a word

```

REGEX : (^ [0-9]+.*\w+\$) | (^ [0-9]+.*\w+\n ([0-9]+.*\w+\n)* [0-9]+.*\w+\$)
STRING:1 banana

```

```

RESULT:1 banana

```

```

*****
REGEX : (^ [0-9]+.*\w+\$) | (^ [0-9]+.*\w+\n ([0-9]+.*\w+\n)* [0-9]+.*\w+\$)
STRING:1 banana
RESULT:1 banana

```

```

*****
REGEX : (^ [0-9]+.*\w+\$) | (^ [0-9]+.*\w+\n ([0-9]+.*\w+\n)* [0-9]+.*\w+\$)
STRING:
1 banana

```

```

RESULT:None

```

```

*****
REGEX : (^ [0-9]+.*\w+\$) | (^ [0-9]+.*\w+\n ([0-9]+.*\w+\n)* [0-9]+.*\w+\$)
STRING:1 banana
1 apple

```

```

RESULT:1 banana

```

```

1 apple

```

```

*****
REGEX : (^ [0-9]+.*\w+\$) | (^ [0-9]+.*\w+\n ([0-9]+.*\w+\n)* [0-9]+.*\w+\$)
STRING:1 banana
1 apple

```

```

RESULT:1 banana

```

1 apple

\*\*\*\*\*

REGEX : (^ [0-9]+.\*\w+\\$) | (^ [0-9]+.\*\w+\n ([0-9]+.\*\w+\n)\* [0-9]+.\*\w+\\$)

STRING:1 banana

five apple

RESULT:None

\*\*\*\*\*

REGEX : (^ [0-9]+.\*\w+\\$) | (^ [0-9]+.\*\w+\n ([0-9]+.\*\w+\n)\* [0-9]+.\*\w+\\$)

STRING:one banana

five apple

RESULT:None

\*\*\*\*\*

-----  
6. all strings that have both the word grotto and the word raven in them (but not, e.g. words like grottos that merely contain the word grotto

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:grotto raven

RESULT:grotto raven

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:grottos ravens

RESULT:None

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:sgrottos sravens

RESULT:None

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:raven grotto

RESULT:raven grotto

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:ravens grottos

RESULT:None

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:sravens sgrottos

RESULT:None

\*\*\*\*\*

REGEX :.\*\b(grotto\b.\*raven\b)|(raven\b.\*grotto).\*

STRING:fuckass

RESULT:None

\*\*\*\*\*

-----  
7. write a pattern that places the first word of an English sentence in a register. Deal with punctuation.

REGEX :([A-Z]\w+'{0,1}\w{0,1})\*\1

STRING:The cat eats. The cat meows.

RESULT:The cat eats. The cat meows.

\*\*\*\*\*

REGEX :([A-Z]\w+'{0,1}\w{0,1})\*\1

STRING:The cat eats banana pancakes

RESULT:None

\*\*\*\*\*

REGEX :([A-Z]\w+'{0,1}\w{0,1})\*\1

STRING:Shouldn't the cat eats? Shouldn't the cat meows?

RESULT:Shouldn't the cat eats? Shouldn't the cat meows?

\*\*\*\*\*

REGEX :([A-Z]\w+'{0,1}\w{0,1})\*\1

STRING:11567 cat eats. 11567 cat meows.

RESULT:None

\*\*\*\*\*

## 2. Implement an ELIZA-like program.

**Answer:**

THIS IS ELIZA

User : Men are all alike.

ELIZA: IN WHAT WAY

User : They're always bugging us about something or the other.

ELIZA: CAN YOU THINK OF A SPECIFIC EXAMPLE

User : Well, my boyfriend made me come here.

ELIZA: WELL, YOUR BOYFRIEND MADE YOU COME HERE.

User : He says I'm depressed much of the time.

ELIZA: WHY DO YOU THINK YOU ARE

SAY HELLO TO LUKE

User : Join me.

LUKE: JOIN YOU.

User : I am your father.

LUKE: IT'S NOT TRUE. THAT'S IMPOSSIBLE

User : I am a jedi.

LUKE: YOU ARE A JEDI.

User : I am a sith.

LUKE: YOU ARE A SITH.

User : I am angry.

LUKE: ANGRY LEADS TO SUFFERING.

User : I hated it.  
 LUKE: HATED LEADS TO ANGER. ANGER LEADS TO SUFFERING.  
 User : It scared me.  
 LUKE: SCARED IS THE PATH TO THE DARK SIDE  
 User : I fear death.  
 LUKE: FEAR IS THE PATH TO THE DARK

3. Complete the FSA for English money expressions in Fig. 2.15 as suggested in the text following the figure. You should handle amounts up to \$100,000, and make sure that “cent” and “dollar” have the proper plural endings when appropriate.

**Answer:**

4. Design an FSA that recognizes simple date expressions like *March 15*, *the 22nd of November*, *Christmas*. You should try to include all such “absolute” dates (e.g. not “deictic” ones relative to the current day, like *the day before yesterday*). Each edge of the graph should have a word or set of words on it. You should use some sort of shorthand for classes of words to avoid drawing too many arcs (e.e. furniture→desk, chair, table).

**Answer:**

5. Now extend your date FSA to handle deictic expressions like *yesterday*, *tomorrow*, *a week from tomorrow*, *the day before yesterday*, *Sunday*, *next Monday*, *three weeks from Saturday*.

**Answer:**

6. Write an FSA for time-of-day expressions like *eleven o'clock*, *twelve-thirty*, *midnight*, or *a quarter to ten*, and others.

**Answer:**

8. Write a regular expression for the language accepted by the NFSA in Fig. 2.26.

**Answer:**

$r^*a(ab)^+|(ba)^+s^*$