Computer Simulations

Decision Trees

1. Decision trees (classification trees)
    - Designed to find the combination of variables that accounts for most of the data. (The measured data are nominal.)
        - A. Splits are made that maximize one category in one branch and minimize it in the other.
        - B. All combinations of independent variables are tried.
        - C. Branches are added until adding more doesn't give you a better fit.
        - D. No statistical significance is calculated.

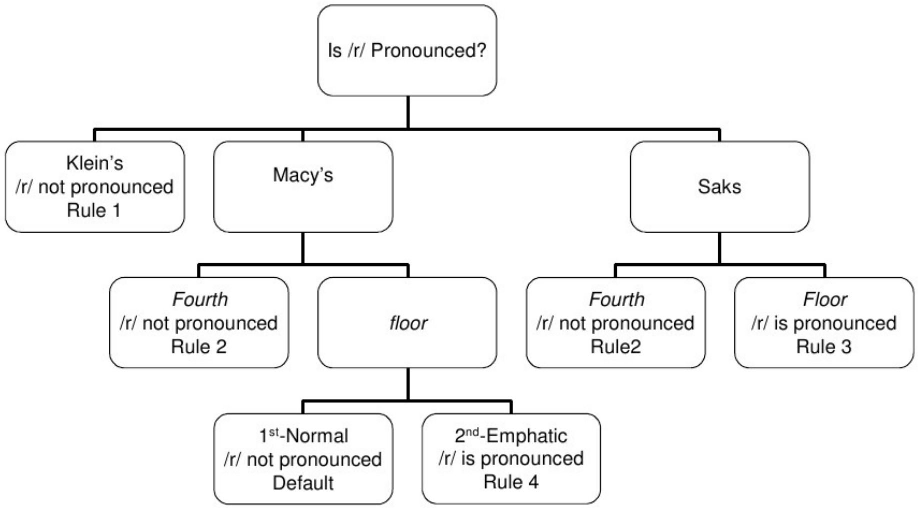
2. Labov's Department Store Study
    - How it was done?
    - What does it show?
    - How do you judge the results?
        - e.g. if there is 25% deletion in Saks and 30% in Macy's?
    - What if there is interaction? (2 or more variables working together)
    - Logistic regression is one method, decision trees are another.

    - For Department store study:
        - Dependent Variable: pronunciation of /r/
        - Independent Variables:
            - Store: Klein's, Macy's, Saks
            - Word: *fourth, floor*
            - Try #: 1st-normal or 2nd-emphatic
        - Question: How do the independent variables effect the dependent variable?

    - Lines of data look like this:

      Kleins, floor, emphatic, no-R.
      Kleins, fourth, emphatic, R.
      Saks, floor, non-emphatic, no-R.

    - Output of decision trees looks like this (when embellished with graphics):



• Rules:

1. Clerks in Klein's do not pronounce /r/ (195/216, 90.3% correct).
2. *Fourth* is pronounced without an /r/ (192/270, 71.1% correct).
3. Clerks in Saks pronounce the /r/ in *floor* (52/82, 63.4% correct).
4. Clerks in Macy's pronounce /r/ in *floor* as an emphatic second response (31/51, 60.8% correct).

• Notice that interactions are found automatically (e.g. store by word; store by word by try #).

2. Oprah Winfrey's pronunciation of [aj] as [aj] or [a] (monophthongization) (Mendoza-Denton, Hay, Jannedy)

• What social factors affect it? The researchers included:

A. The person Oprah was talking about
B. The race of that person
C. The gender of the person
D. Class of word (*I* is a pronoun, *light* is a noun (or verb))
E. Frequency of the word
G. What sound precedes [aj]

• Lines of data look like this: ("So, ***I*** talked to Tina the other day.")
Tina, black, female, *I*, 5443.7, [ow]

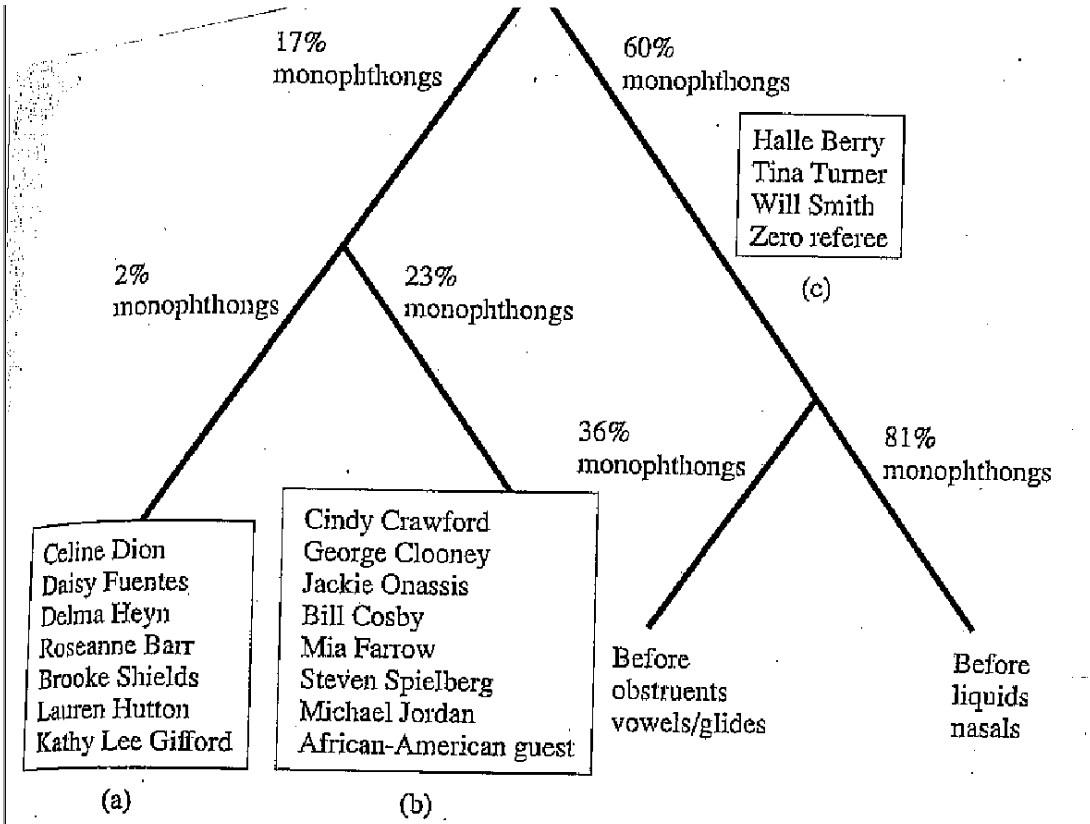


Figure 4.8
CART classification tree for monophthongization

- The decision tree only found a good fit with two variables:
  - A. The person Oprah was talking about
  - G. The kind of sound that precedes [aj]
- What are the "rules" this tree gives?
- Decision trees are good at making sense of messy data.

Choose the comparative that sounds better.

Adjective		
noble	nobler	more noble
common	commoner	more common
careful	carefuler	more careful
hard	harder	more hard
green	greener	more green
accurate	accurater	more accurate
angry	angrier	more angry
mellow	mellower	more mellow
unhappy	unhappier	more unhappy

Analogical Modeling (Skousen)

- Not a diagnostic/statistical program to make sense of data, but a model of mental categorization

- Assumptions
  - A. We store exemplars of past experience in our head.
  - B. We use those exemplars in processing.
  - C. Grammar emerges from relationships between exemplars, not from making rule generalizations from the data.

1. English adjective comparison (Elzinga)

- Some adjectives take the suffix *-er* while others are preceded by *more.*
  *more important, *importanter*, *more fake, *faker*

- How do English speakers know which one to use? People may "follow" these rules (did you follow them in the test words above?):
  - A. Add *-er* if the adjective is monosyllabic.
  - B. Use *more* if the adjective is disyllabic and if it's stressed on the first syllable and ends in *-y, le, ow.*
  - C. Otherwise use *more.*

- Is this a description of English or a rule English speakers follow?
  - A. Of 485 test cases, 93.6% are correctly predicted by these rules
  - B. Problems:
    1. How can people follow rules they can't formulate?
    2. How do you account for variation?
    3. Requires lots of computation, little storage (problem?)
    4. Rules are different for every task.
    5. If you leave a variable out, you can't make good predictions (e.g. *a, an*).

- If there is no rule how do people know the comparative of new words? (*Plocous, steech*)
  Analogy-Find words that are similar to *plocous* and see what their comparatives are and apply that to *plocous.*

- 485 comparatives simulate what is in English speaker's head (adj. from corpus, category from Google).
  - A. Each adjective marked to indicate whether it takes *-er* or *more.*
  - B. Information about each adjective converted to variables:
    1. Number of syllables
    2. Where the stress falls
    3. What consonants and vowels make up the word (ordered)

- 92.6% correct in leave-one-out-simulation (compare to 93.6% for rules–significant?)
    - Outcome is probabilistic (97% *-er*, 3% *more*)
    - Winner take all strategy used. Accounts for variation.
    - Lots of storage, little computation
    - Computation is same for all categorization tasks
    - If you leave variables out, not a lot changes (most of Elzinga's hovered around 90%)

- Just because a computer does it analogically doesn't mean people do.
    - 300 new adjectives chosen.
    - People and computer had to choose *-er* or *more.*
    - Correlation (majority rules) $p < .0005$
    - Problem: No comparison with rule outcome

Give the past tense of the following verbs:

Present Tense	Past Tense
queed	
nace	
bize	
lum	
fleep	
shilk	
gleed	
scoil	
flet	
tesh	
spling	
chake	
gude	
gare	
nold	
chind	

2. English past tense (review dual-route model)

• Variables without outcomes used as a model of the mental lexicon:

Verb	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
transform	0	t	r	æ	n	s	0	-	0	f	o	r	m	0	F	m
distribute	s	t	r	I	0	-	-	0	b	j	u	t	0	-	N	t
tally	-	0	t	æ	0	-	-	-	0	l	i	0	-	-	N	i

Variable 1: outcome
Variable 2: Final stress on last syllable or No stress on last syllable
Variables 3-16: Phonological makeup of verb

• Many different outcomes (*walk/walked, teach/taught, ring/rang, stick/stuck*). 2179 verbs

• 58 nonce words in study (taken from Albright and Hayes)
What is past tense of *spling?* What do people say? What does the computer model predict?

Example:
1. *splinged* 51.4%
2. *splung* 32.4%
3. *splang* 10.8%

• Correlation between computer predictions and people's predictions (.877, $p < .005$)

Is that a good correlation? What does it mean?

• Issues: type or token frequency and variable alignment

	Token Frequency	Type Frequency
All 2179 verbs	.814	.903
Most frequent 1090 verbs	.777	.877
Least frequent 1089 verbs	.884	.892
Middle frequency 1090 verbs	.899	.896

Success Rates on Prasada and Pinker's Data for Simulations Using Different Frequency Measures.

	Token Frequency	Type Frequency
All 2179 verbs	.896	.996
Most frequent 1090 verbs	.938	.982
Least frequent 1089 verbs	.987	.989
Middle frequency 1090 verbs	.987	.993

Correlations with Prasada and Pinker's Nonce Verbs.

Verb	12	11	10	9	8	7	6	5	4	3	2	1
transform	t	r	æ	n	s	s	f	o	r	m	F	m
distribute	t	r	I	b	j	b	j	u	t	0	N	t
tally	0	t	æ	l	i	æ	l	i	0	-	N	i

Examples of the No-Syllable-Boundary Alignment.

Verb	8	7	6	5	4	3	2	1
transform	tr	æ	ns	f	o	rm	F	m
distribute	str	I	0	bj	u	t	N	t
tally	t	æ	0	l	i	0	N	i

Examples of the Lumped-Consonant-Cluster Alignment.

Verb	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
transform	0	t	r	æ	n	s	0	-	0	f	o	r	m	0	F	m
distribute	s	t	r	I	0	-	-	0	b	j	u	t	0	-	N	t
tally	-	0	t	æ	0	-	-	-	0	l	i	0	-	-	N	i

Examples of the Separated-Consonant-Cluster Alignment

Verb	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1
transform	t	r	-	-	t	r	æ	n	s	-	-	n	s	f	-	-	-	-	f	o	r	m	-	-	r	m	F	m
distribute	s	t	r	s	t	r	I	0	-	-	-	-	0	b	j	-	-	b	j	u	t	-	-	-	-	t	N	t
tally	t	-	-	-	-	t	æ	0	-	-	-	-	0	l	-	-	-	-	l	i	0	-	-	-	-	0	N	i

Examples of Dual-alignment