Quantum Computing of Analogical Modeling of Language

Royal Skousen

26 March 2007

Analogical Modeling (AM)

exemplar or instance-based

- not a rule-based system
- not a neural network

predictions are procedural, not declarative

predicting an outcome in terms of a given context not a nearest neighbor approach

- includes nearest neighbors
- but also non-neighbors under conditions of homogeneity

no training stage

- collecting exemplars
- no setting of paramenters
- no prior determination of variable significance
- predictions made “on the fly”

simple measure of uncertainty

determined by the number of disagreements

between pairs of exemplars within a given space

- simple decision procedure to determine homogeneity
  - minimize the number of disagreements
  - no increase in uncertainty permitted
    - no loss of information

two types of homogeneous spaces

(a) deterministic behavior
no disagreement occurs
exemplars can be anywhere within the space

(b) non-deterministic behavior
disagreement occurs, but is minimized
exemplars are restricted to a single subspace

probability of an exemplar being chosen

(a) proximity

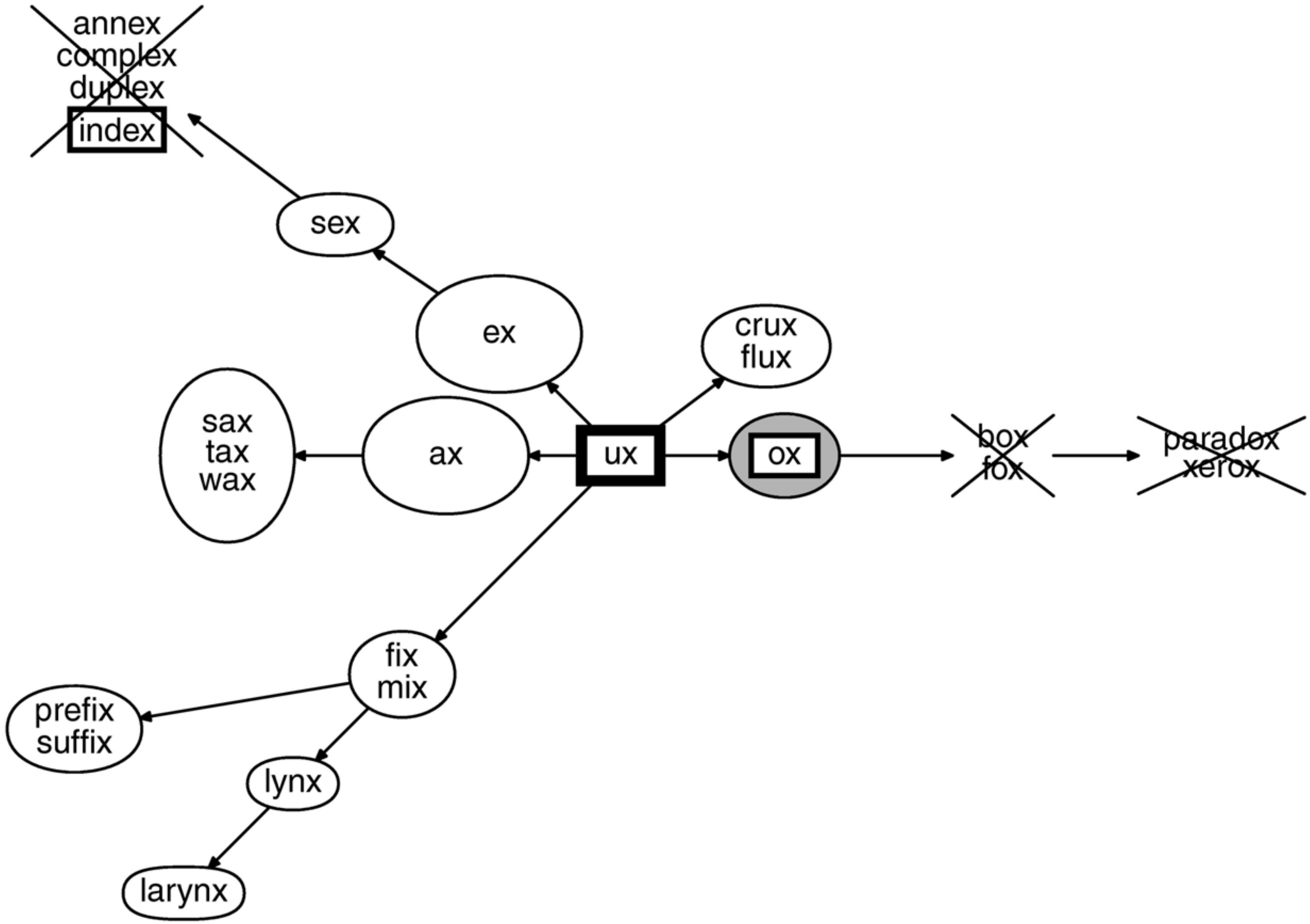
closeness to the given context

(b) gang effect

surrounding exemplars behave alike

(c) heterogeneity

intervening behavior is different


annex
complex
duplex
index
sex
ex
crux
flux
sax
tax
wax
ax
ux
ox
box
fox
paradox
xerox
fix
mix
prefix
suffix
lynx
larynx

a classical categorical "rule"
the indefinite article *a* versus *an* in English
the "crucial" variable:
    consonant versus vowel of the following segment
a boy versus *an apple*

constructing a dataset
- 251 examples
  - 211 of *a* versus 40 of *an*
  - no exceptions to the rule
- 9 variables:
  - phonemes, basic syllable and phrasal structure

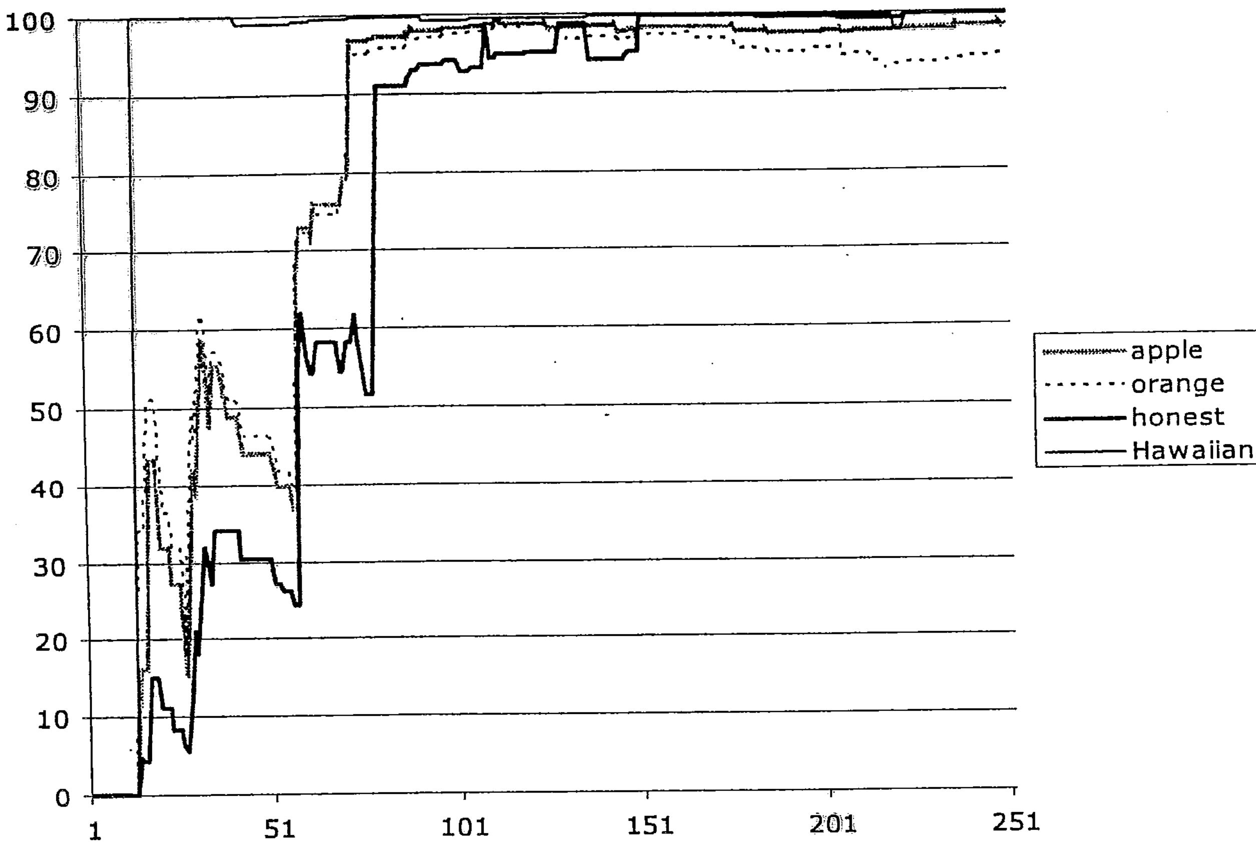
predicting 4 cases where *a* is expected

- virtually 100 percent *a*
- no leakage towards *an*
  - as in “an boy”

predicting 4 cases where *an* is expected

- leakage towards *a*
- small numbers of exemplars: high leakage
- large numbers of exemplars: narrows, but never vanishes

Sapir: “all grammars leak”



100
90
80
70
60
50
40
30
20
10
0
1
51
101
151
201
251
apple
orange
honest
Hawaiian

more than nearest neighbors

example of a gang effect overwhelming the nearest neighbors

Finnish past-tense formation

sorta- ‘to oppress’

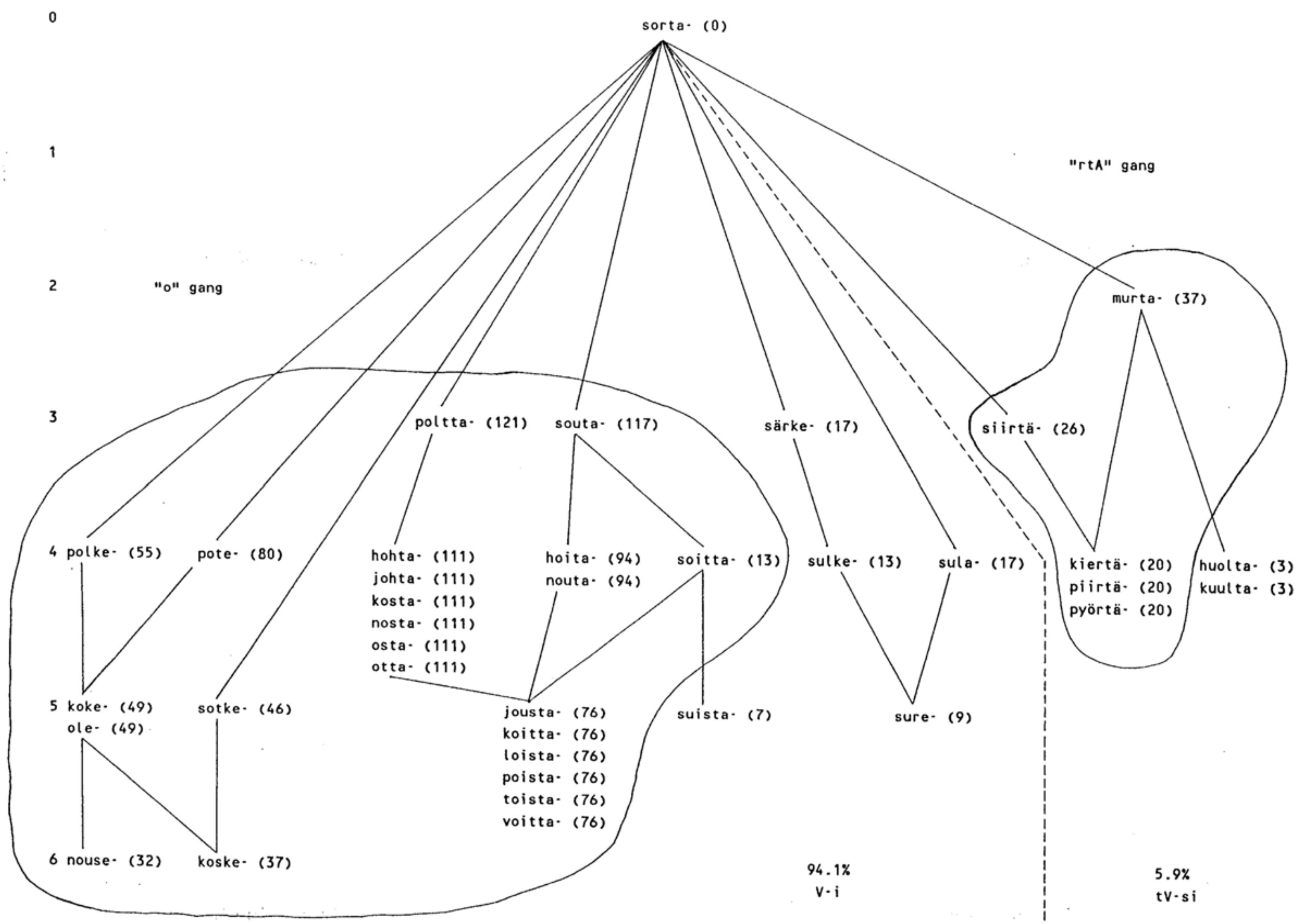
two possible past-tense forms:

tV-si	*sorsi*
V-i	*sorti*

classical rule approaches predict *sorsi*, the historical form yet *sorsi* is what actually occurs:

	sorti	*sorsi*
NSSK	11	1
SKP	2	0

Analogical set for the past tense of the verb *sorta*-



given context
supracontexts (generalized forms of the given context)
each homogeneous supracontext and its exemplars = “a true rule”
each heterogeneous supracontext and its exemplars = “a false rule”

all possible “true rules”
chances of using a particular “true rule”

(a) equally probable rules

(b) proportional to the frequency of the rule
number of exemplars

✓ (c) proportional to the square of the frequency of the rule
number of exemplars squared

uses the same quadratic measure of uncertainty
predicts appropriate degree of fuzziness at boundaries

differences

- no partitioning of the contextual space
  - as in classical rule approaches
- “rules” created “on the fly”, do not permanently exist
- non-deterministic “rules”:
  - no objective probabilities, only exemplars

Quantum Analogical Modeling (QAM)

- given context and its supracontexts
- rules for the given context:
  - rule = supracontext + its exemplars
- three aspects

(1) all possible rules exist in a superposition

reduces the exponential explosion to linearity

both in time and memory (number of qubits)

the one essential quantum concept in QAM

(2) the system evolves:

(a) the amplitude of every heterogeneous rule equals zero

(b) the amplitude of each homogeneous rule equals its relative frequency of occurrence restricted to classical reversible operators

- (a) apply to every rule in the superposition
    - no conditional application to only some of the rules
- (b) only states for the qubits: 0 and 1
    - qubits can be copied
    - ability to replicate any classical reversible operator

fundamental operators

for a qubit vector

NOT(**X**)
CNOT(**A**, **X**)
CCNOT(**A**, **B**, **X**)

for an individual qubit

NOT(X_i)
CNOT(A_j, X_i)
CCNOT(A_j, B_k, X_i)

for a superpositioned qubit vector

$\text{NOT}(\mathbb{X})$
$\text{CNOT}(\mathbb{A}, \mathbb{X})$
$\text{CCNOT}(\mathbb{A}, \mathbb{B}, \mathbb{X})$

for an individual qubit in a superpositioned qubit vector

$\text{NOT}(\mathbb{X}_i)$
$\text{CNOT}(\mathbb{A}_j, \mathbb{X}_i)$
$\text{CCNOT}(\mathbb{A}_j, \mathbb{B}_k, \mathbb{X}_i)$

(3) observation reduces the superposition to a single rule

probability of selection = amplitude squared

classical quantum mechanical approach

two-stage observation of the evolved system

(a) randomly select a single rule

a supracontext and its outcomes

probablility of selection = frequency squared

(b) randomly select an outcome

a exemplar approach to observation

- directional pointers connecting each pair of exemplars
- for each supracontext
  - includes a pointer from each exemplar to itself
  - if n exemplars, there are n^2 directional pointers
- (a) use pointers to determine uncertainty
- (b) use pointers to observe the system
  - observation leads to the squaring of the amplitude
    - the number of exemplars

as the system evolves:

(a) empty supracontexts

no pointers exist at any time

(b) heterogeneous supracontexts (unobservable)

pointers between exemplars become unaccessible

(c) homogeneous supracontexts

number of pointers equals the square of the frequency

measurement: randomly select any accessible pointer
from the pointers in existing (homogeneous) supracontexts
gives the same result as the traditional two-stage measurement in QM

QAM: a general quantum computational algorithm

predicts behavior, including language

supracontexts classified according to the property of heterogeneity

numerous other properties possible

but are they useful?

Nielsen and Chuang (2000:39)

> A significant caveat is that even though a quantum computer can simulate many quantum systems far more efficiently than a classical computer, this does not mean that the fast simulation will allow the desired information about the quantum system to be obtained. When measured, a *kn* qubit simulation will collapse into a definite state, giving only *kn* bits of information; the c^n bits of ‘hidden information’ in the wavefunction is not entirely accessible.

arxiv.org, quantum physics
“Quantum Analogical Modeling: A General Quantum Computing Algorithm for Predicting Language Behavior”
18 October 2005
http://arxiv.org/abs/quant-ph/0510146