# Logos: Logic in Computer Science

Michel de Rougemont

*University Paris II & IRIF CNRS*

# Plan

1. LICS Laboratories
   - IRIF
   - Simons Institute: https://simons.berkeley.edu/
2. Recent Research Subjects

3. Statistical model for morphology inspired by the Amis language

   -CS-Linguistics paper

# 1. IRIF

1. Algorithmes
   - Algorithmes et Complexité
     - » Quantum
     - » Approximation
     - » Security
   - Calcul distribué
   - Combinatoire
   - Graphes
2. Automates, vérification
   - Automates
   - Vérification
3. Preuves, Programmes et Systèmes

# 2. Subjects for December 19th

1. Complexity
2. Probabilistic algorithms
3. Property testing and sublinear algorithms
4. Descriptive Complexity
5. Games, algorithmic game theory
6. Streaming Algorithms
7. Non worst-case  analysis
8. Random graphs
9. Machine learning
10. LLMs: Large Language Models

# Recent Trends

1. Randomness
   - Randomized algorithms
   - Randomized verification
     - IP, PCP, MIP 1990s
     - MIP*=r.e. 2020
2. SAT problem

   •NP-complet 1970

   •SAT conferences provide solutions on larger and larger instances
3. Approximations
   - Optimisation
   - Property testing
   - Statistical properties
4. Learning
   - Supervised, unsupervised, Pac-learning, VC dimension
   - Graph learning

# 3. Statistical model for morphology inspired by the Amis language

http://mdr.free.fr/td/amis.pdf

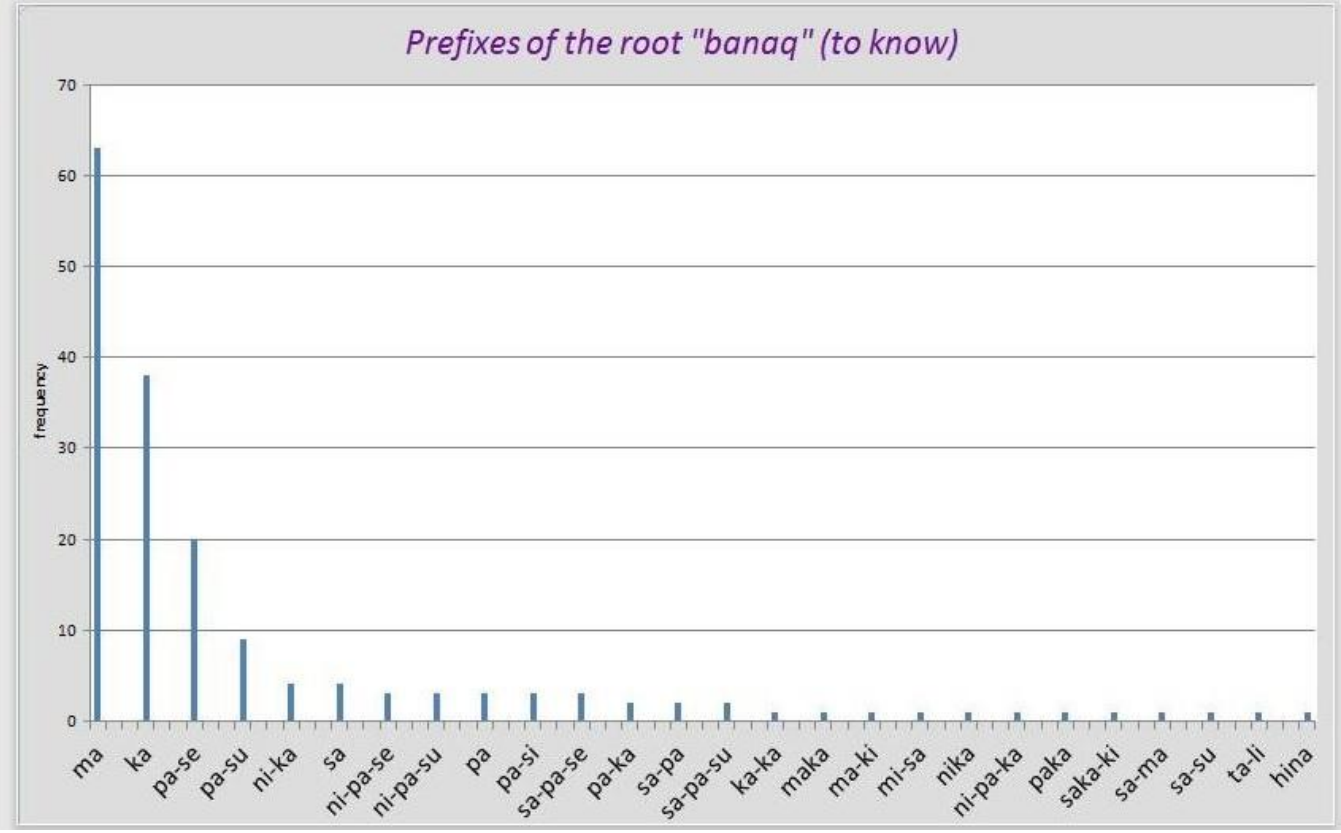
1. Word2Vec: basic construction of modern NLP
   http://nlp.polytechnique.fr/word2vec#french

2. Morphology in languages

- Re-construc-tion
- Auto-ver-sicher-ung
- В-ход-ить
- 船
- 今天

# Amis language: austronesian language

Rich structure of  Prefix-root-suffix



# Correlation Matrices

*1. Mi-padang k-u tumuk t-u suwal n-ira tatakulaq.*
AV-help NOM-ART chief OBL-ART word GEN-that frog[10]
The tumuk supported the words of the frog.

*2. Isu Kungcu, yu ira k-u pa-padang-an,...*
You Princess when exist NOM-ART RED-help-NMZ
You Princess, when (you) had some help,...

*3. Sulinay mi-padang k-u taw,...*
Indeed AV-help NOM-ART people
Indeed when people help,...

*4. Aka-a ka-pawan t-u ni-padang-an n-u taw.*
PROH-IMP NFIN-forget OBL-ART PFV.NMZ-help-NMZ GEN-ART people
Then, you mustn't forget people's help.

$$M_p = \begin{array}{c} \\ k \\ ka \\ n \\ ni \\ mi \\ pa \\ t \end{array} \begin{array}{c} \begin{array}{ccccccc} k & ka & n & ni & mi & pa & t \end{array} \\ \begin{pmatrix} 2 & 0 & 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 2 & 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 1 & 1 & 0 & 2 \end{pmatrix} \end{array}$$

# PCA: Principal Component Analysis

$M = U . U^t \approx$  $= B * C$

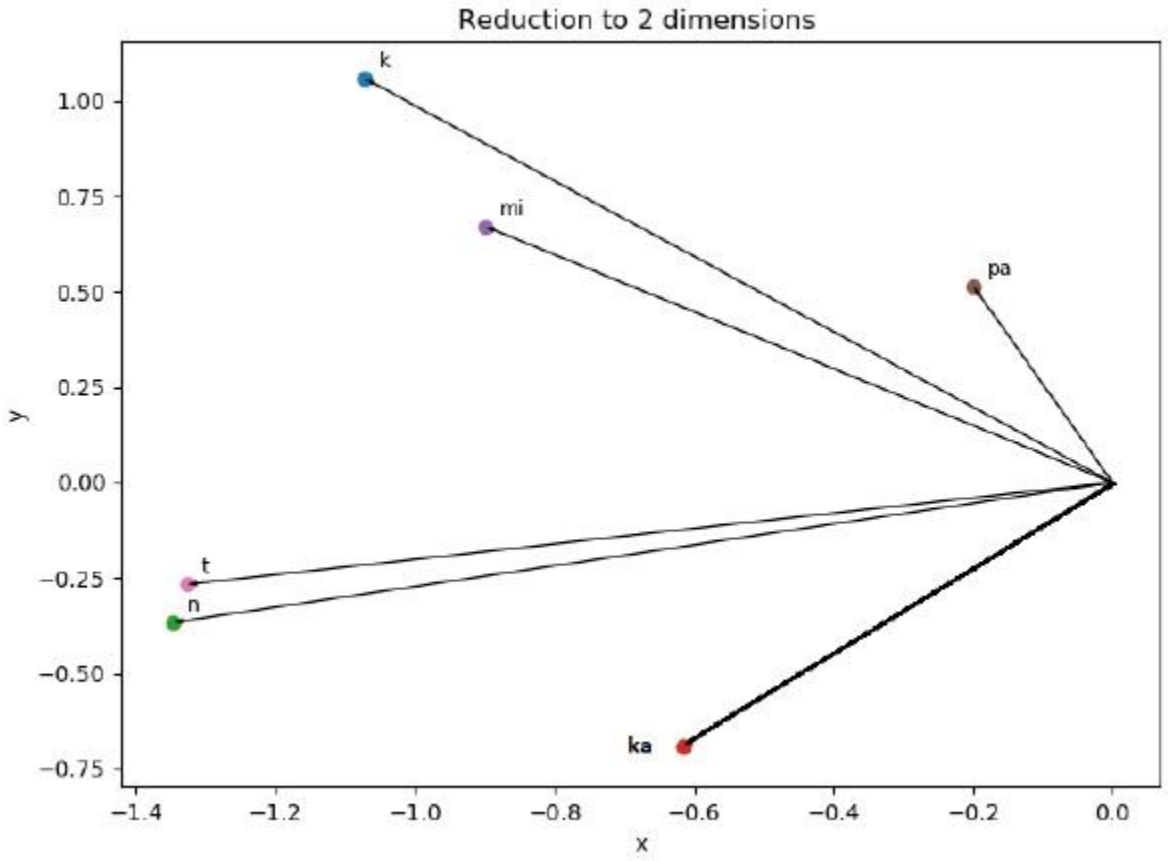
Positive semidefinite matrices

$$M_p = \begin{array}{c} \\ k \\ ka \\ n \\ ni \\ mi \\ pa \\ t \end{array} \begin{array}{c} \begin{array}{ccccccc} k & ka & n & ni & mi & pa & t \end{array} \\ \begin{pmatrix} 2 & 0 & 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 2 & 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 1 & 1 & 0 & 2 \end{pmatrix} \end{array}$$

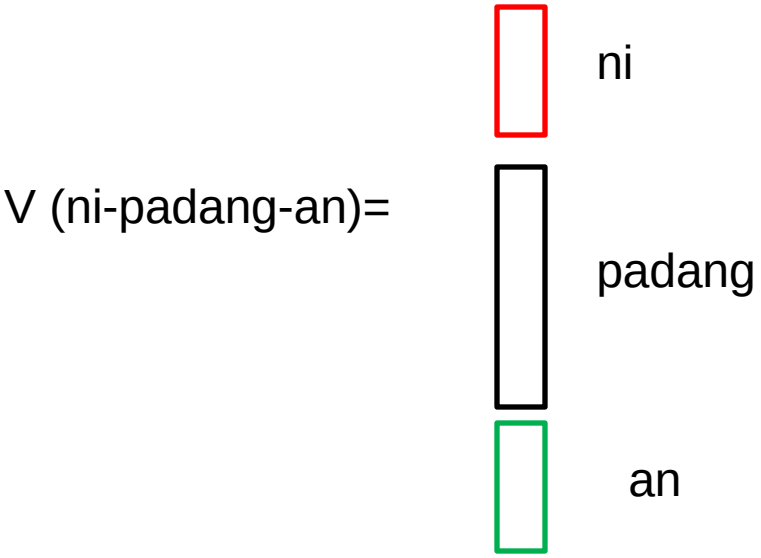
$$B = \begin{array}{c} \\ k \\ ka \\ n \\ ni \\ mi \\ pa \\ t \end{array} \begin{array}{c} \begin{array}{cc} x_1 & x_2 \end{array} \\ \begin{pmatrix} -1.07140232 & 1.05796864] \\ -0.61507482 & -0.69205651 \\ -1.35624875 & -0.37685711 \\ -0.61507482 & -0.69205651 \\ -0.89917084 & 0.67091885 \\ -0.19923754 & 0.51352862 \\ -1.34624875 & -0.36685711 \end{pmatrix} \end{array}$$

$M(i,j) = v_i . v_j$

# Principal Component Analysis on the prefixes



# Structural vector representation




Morphology is included in the embedding

Different from Word2vec

# Conclusion

1. Laboratories
   - IRIF
   - Simons Institute

2. Subjects

3. Statistical model for morphology inspired by the Amis language
   - Word2vec embedding
   - Language morphology

# Testers and Correctors

Distance: Edit, Edit with moves,….

$\varepsilon$−close

1. Words: regular expressions, U=({1,2,…n},<,A,B)

   w=abababcababacacacab

   r= (ab)*(ac)*

   Corrector: w1=abababababacac

   w2=abababababacacacac

   w3=abababababacac

   w4=abababababacacacac

Q(x): A(x) and B(x+1) then CQA(A)={1,3,5,7}