INFOMOTION
we love data

Korrelation von menschenähnlichen Entscheidungsstrukturen und periodisch auftretenden Figuren ohne Flügeln

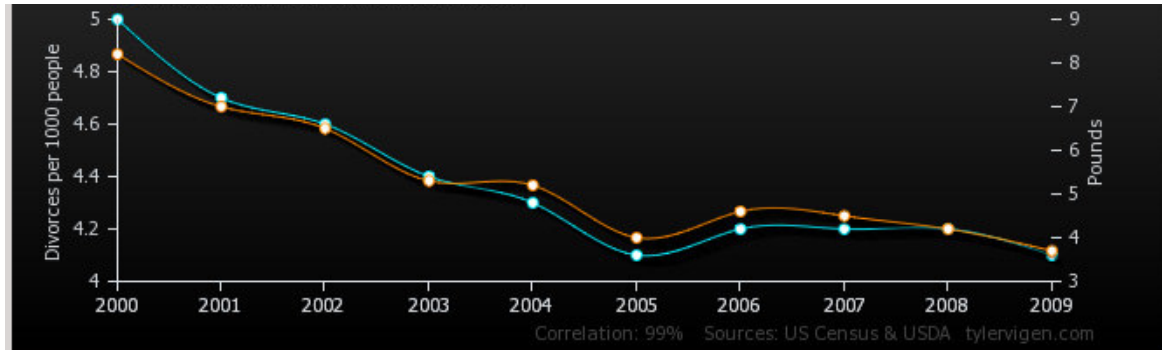
Stuttgart, 12. Dezember 2018

Dr. Nicolai Erbs

Korrelation und Kausalität

KORRELATION

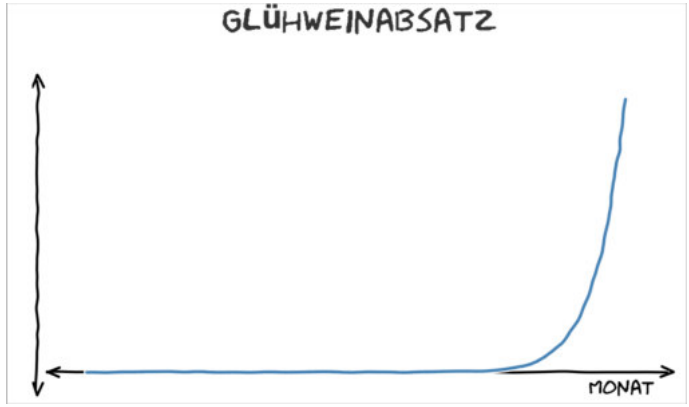
- Zusammenhang zwischen zwei statistischen Variablen
- Zusammenhang kann quantifiziert werden
- Kein zwanghafter kausaler Zusammenhang
- Korrelation ist symmetrisch



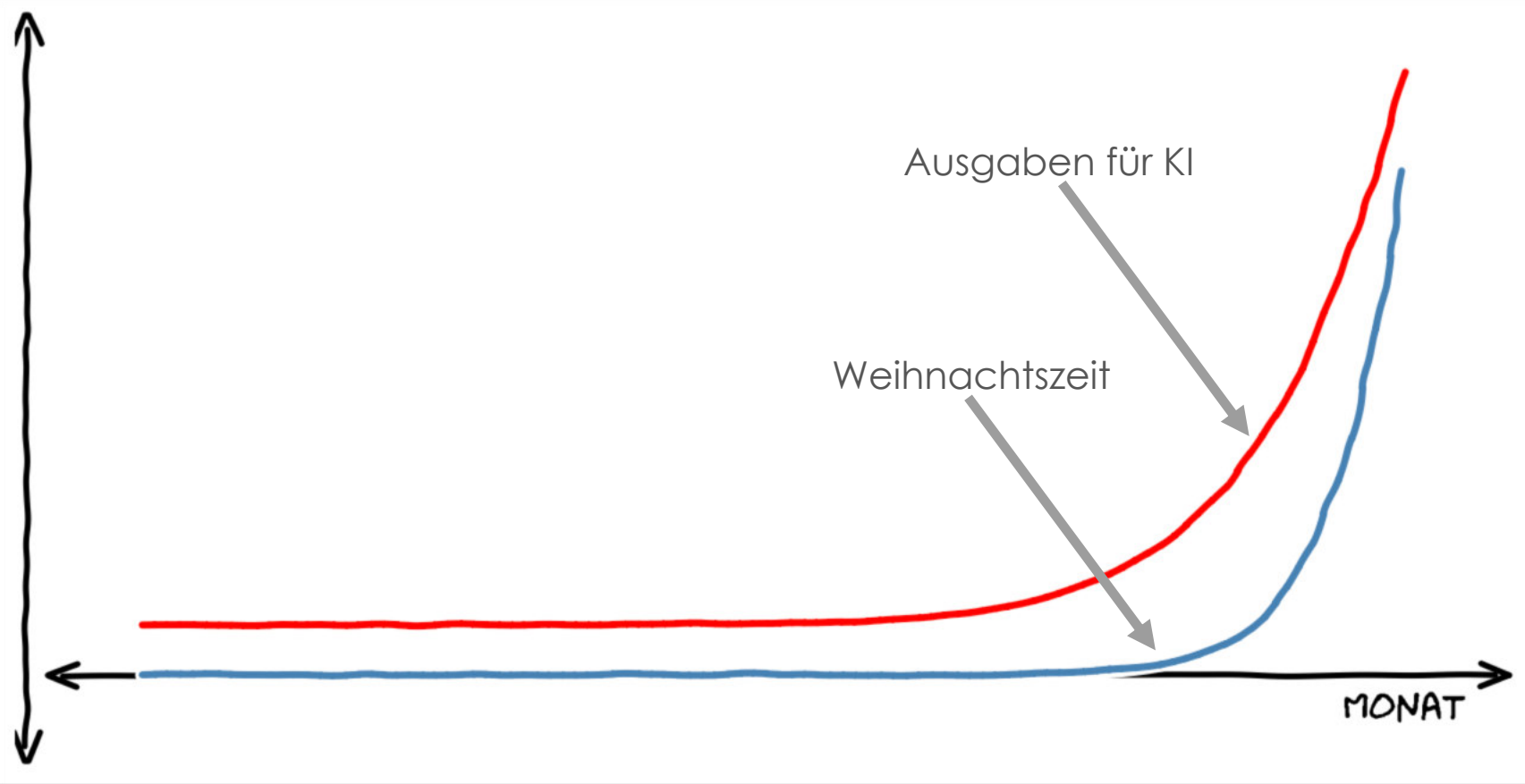
- Scheidungsrate in Maine im Vergleich zu Margarinekonsum

KAUSALITÄT

- Beziehung zwischen Ursache und Wirkung
- Zusammhang kann mono- oder multikausal sein
- Kausalitätsketten möglich



Empirische Wahrnehmung des Ausgabeverhaltens



INFOMOTION
we love data

Korrelation … dungsstrukturen und … ügeln

Den Weihnachtsmann gibt es nur dank Künstlicher Intelligenz!

Nordpol, 24. Dezember 2018
Nicolai „Grinch“ Erbs

Warum sollte es den Weihnachtsmann nicht geben?

1. Der Weihnachtsmann kann nicht alle Wunschlisten lesen!
2. Der Weihnachtsmann kann nicht jedes Geschenk ausliefern!
3. Der Weihnachtsmann kann nicht alle Kinder zuordnen!

Der Weihnachtsmann kann nicht alle Wunschlisten lesen!

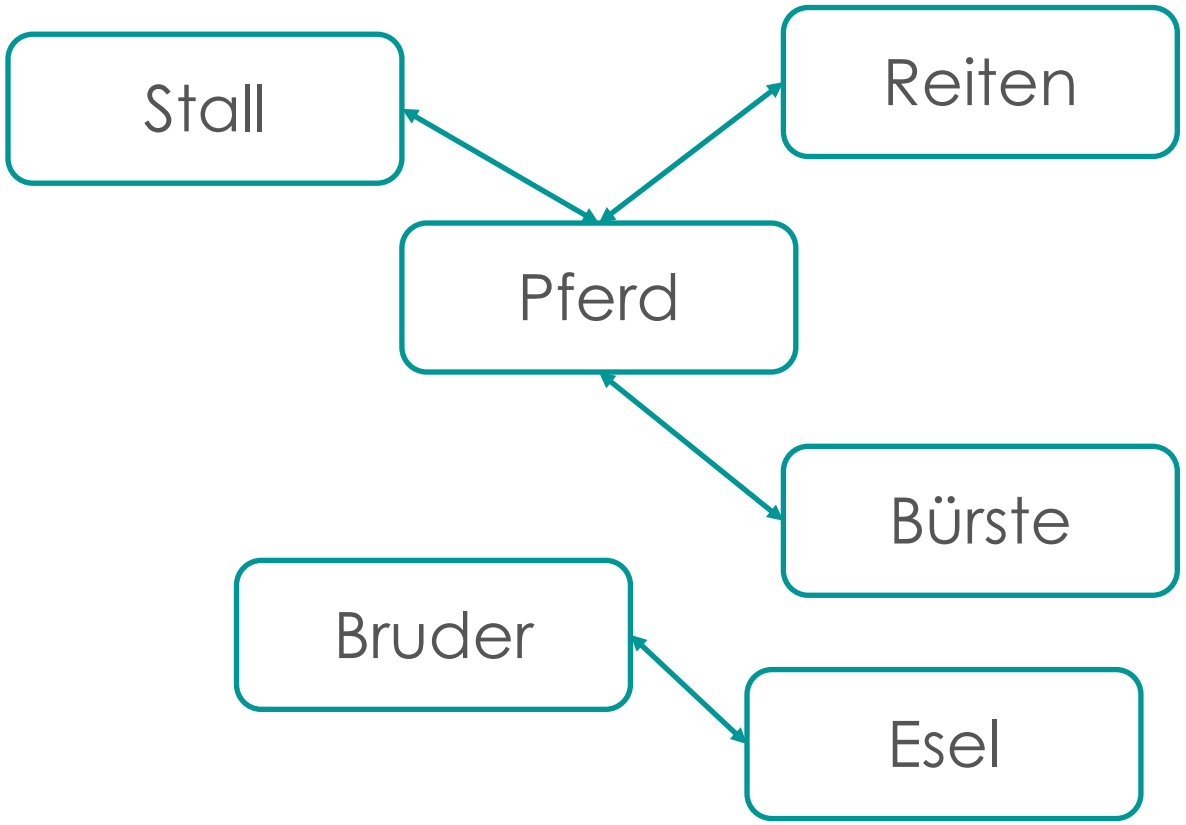
Der Weihnachtsmann kann nicht alle Wunschlisten lesen!

Sprachverarbeitung hilft bei der Analyse

„Lieber Weihnachtsmann,
Ich wünsche mir ein Pferd zum Reiten. Das Pferd braucht einen Stall und ich brauche eine Bürste zum Kämmen. Mein Bruder soll einen Esel haben.“

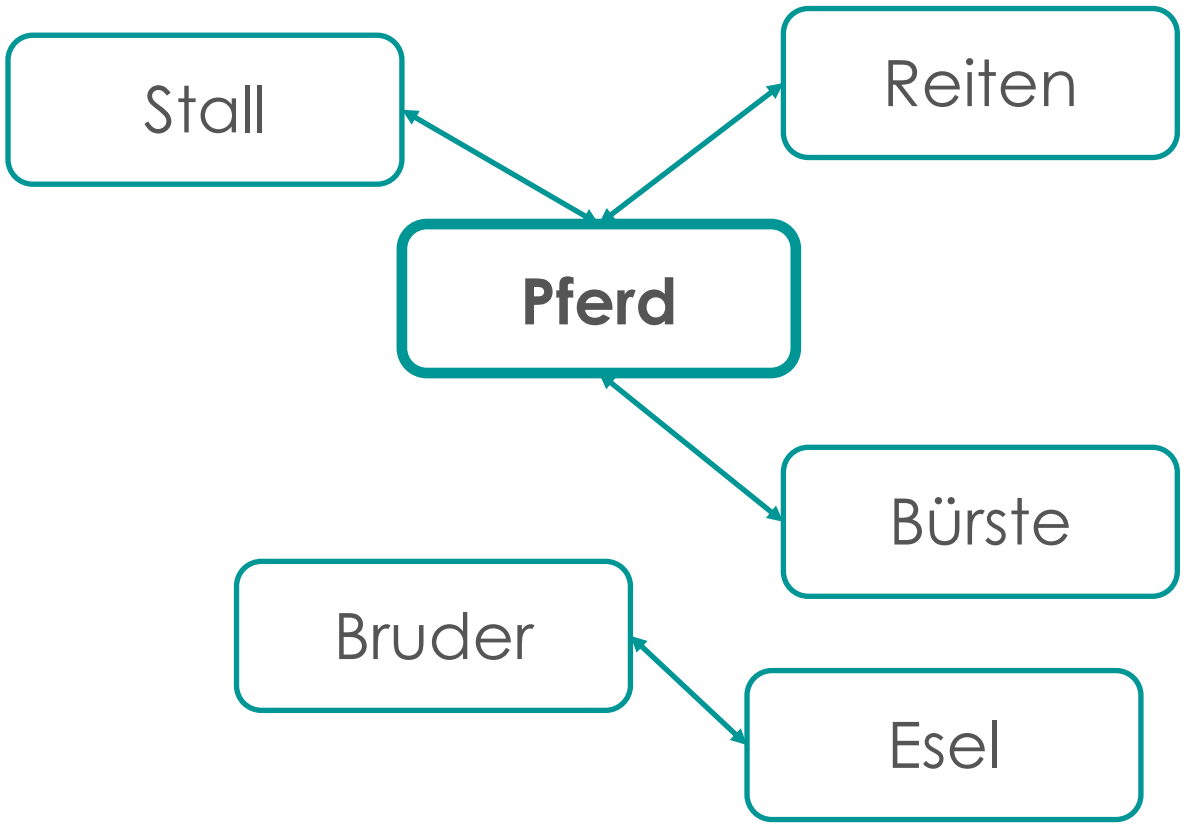
Sprachverarbeitung hilft bei der Analyse

„Lieber Weihnachtsmann,
*Ich wünsche mir ein **Pferd** zum **Reiten**. Das **Pferd** braucht einen **Stall** und ich brauche eine **Bürste**, damit ich **es** kämen kann. Mein **Bruder** soll einen **Esel** haben.“*



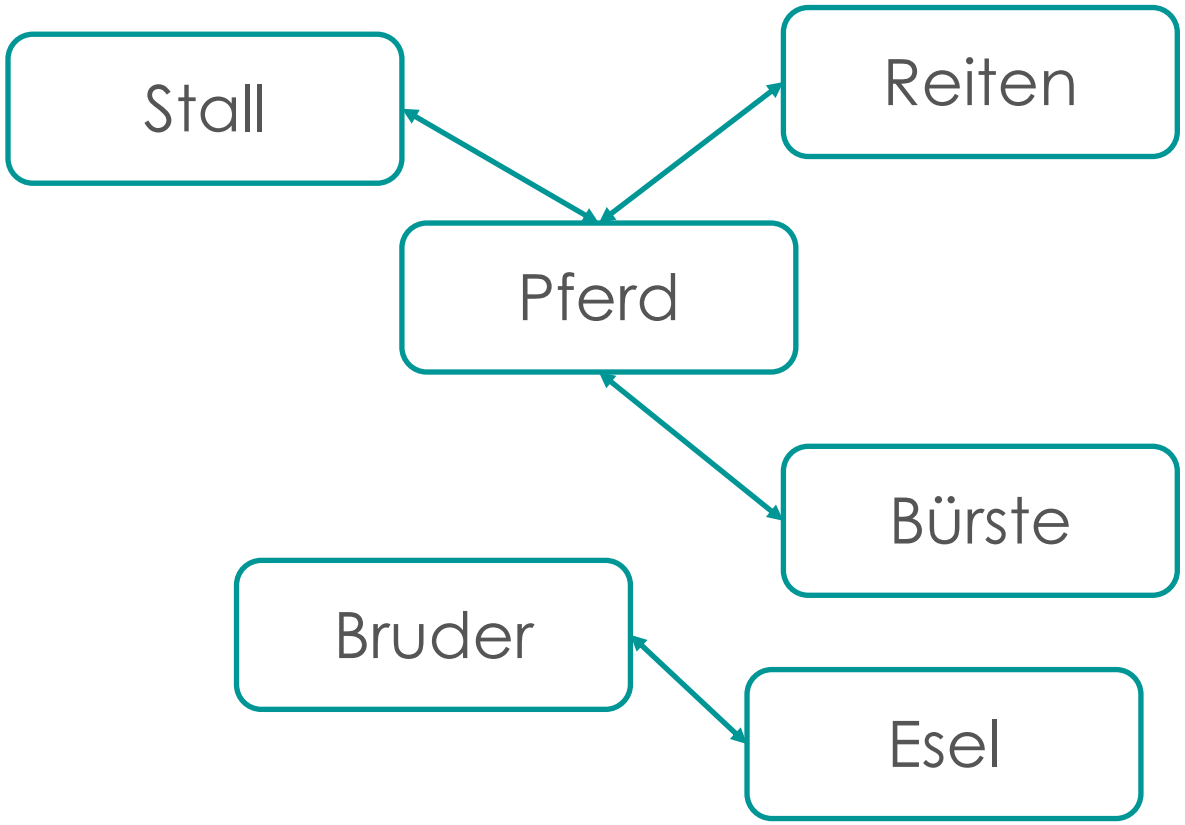
Sprachverarbeitung hilft bei der Analyse

„Lieber Weihnachtsmann,
*Ich wünsche mir ein **Pferd** zum **Reiten**. Das **Pferd** braucht einen **Stall** und ich brauche eine **Bürste**, damit ich **es** kämen kann. Mein **Bruder** soll einen **Esel** haben.“*



Sprachverarbeitung hilft bei der Analyse

„Lieber Weihnachtsmann,
*Ich wünsche mir ein **Pferd** zum **Reiten**. Das **Pferd** braucht einen **Stall** und ich brauche eine **Bürste**, damit ich **es** kämen kann. Mein **Bruder** soll einen **Esel** haben.“*

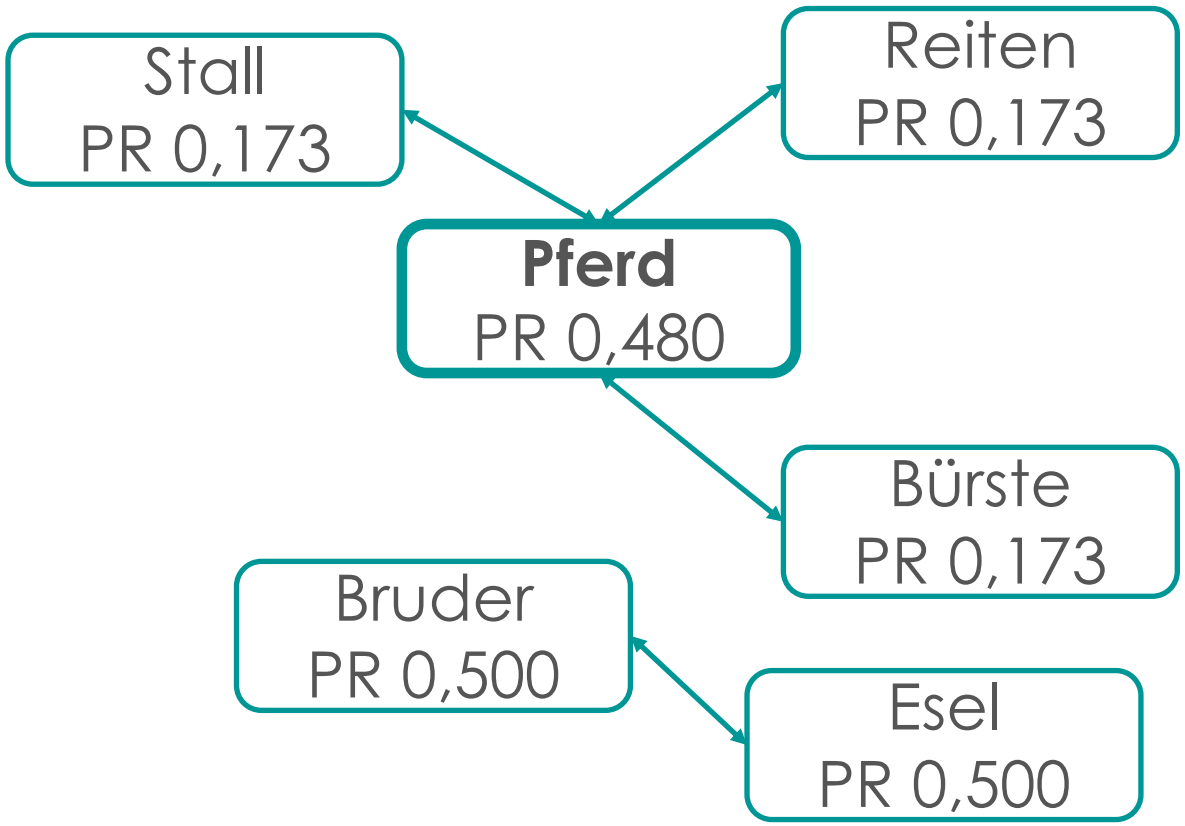


$$R'(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u)$$

Sprachverarbeitung hilft bei der Analyse

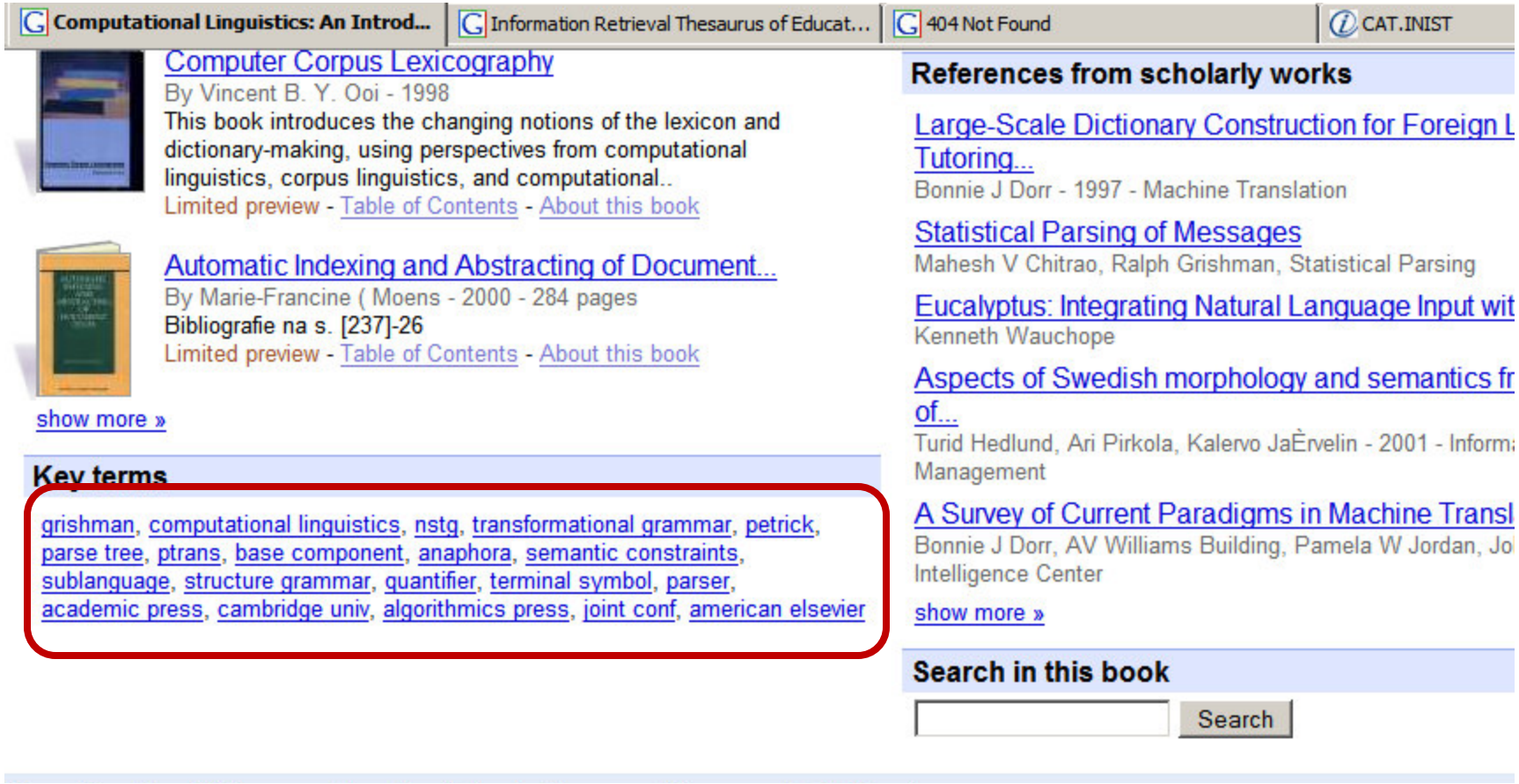
„Lieber Weihnachtsmann,

*Ich wünsche mir ein **Pferd** zum **Reiten**. Das **Pferd** braucht einen **Stall** und ich brauche eine **Bürste**, damit ich **es** kämen kann. Mein **Bruder** soll einen **Esel** haben.“*

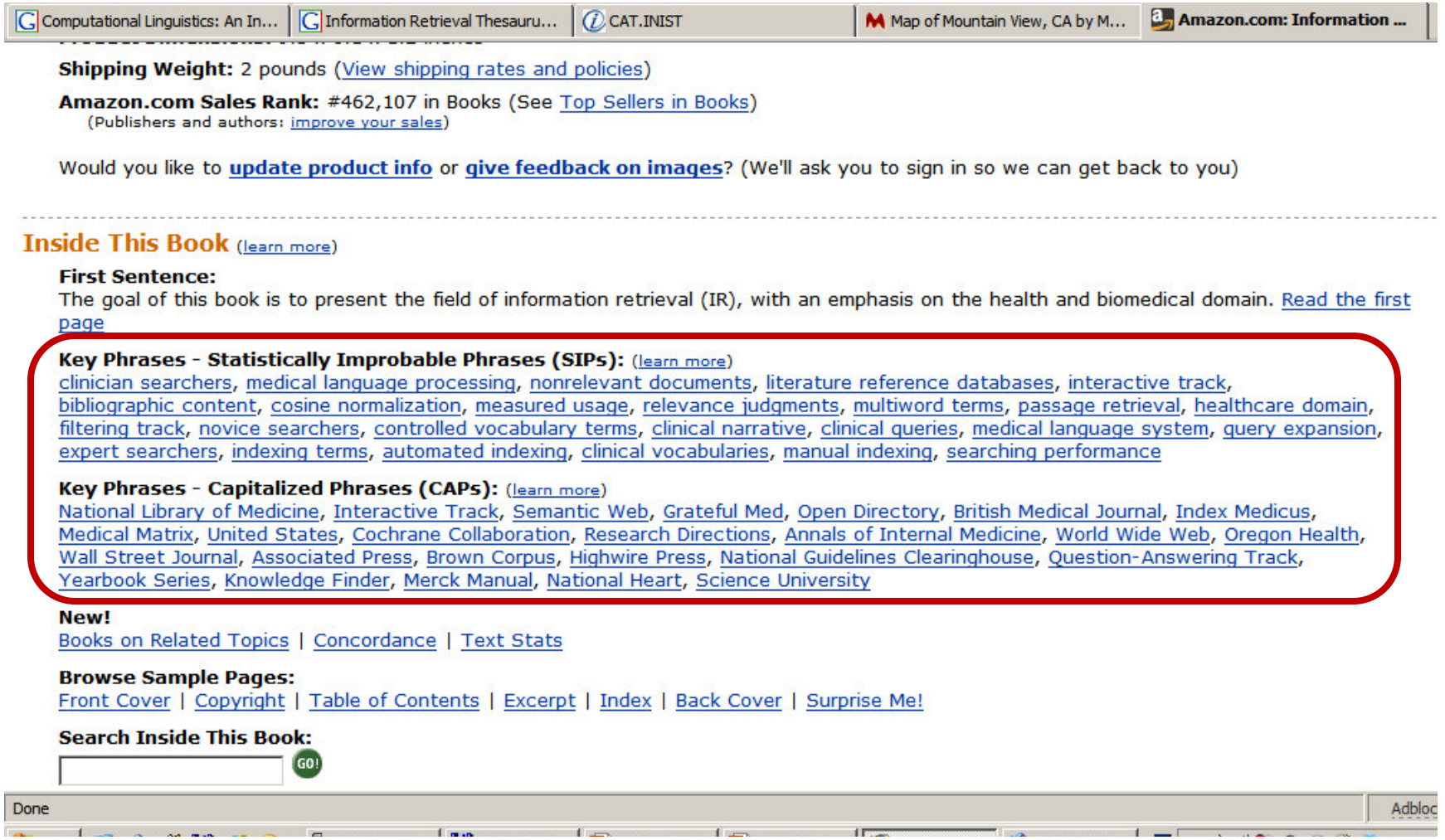


$$R'(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u)$$

Example


Computational Linguistics: An Introd...
Information Retrieval Thesaurus of Educat...
404 Not Found
CAT.INIST
Computer Corpus Lexicography
By Vincent B. Y. Ooi - 1998
This book introduces the changing notions of the lexicon and dictionary-making, using perspectives from computational linguistics, corpus linguistics, and computational..
Limited preview - Table of Contents - About this book
Automatic Indexing and Abstracting of Document...
By Marie-Francine (Moens - 2000 - 284 pages
Bibliografie na s. [237]-26
Limited preview - Table of Contents - About this book
show more »
Key terms
grishman, computational linguistics, nstg, transformational grammar, petrick, parse tree, ptrans, base component, anaphora, semantic constraints, sublanguage, structure grammar, quantifier, terminal symbol, parser, academic press, cambridge univ, algorithmics press, joint conf, american elsevier
References from scholarly works
Large-Scale Dictionary Construction for Foreign L
Tutoring...
Bonnie J Dorr - 1997 - Machine Translation
Statistical Parsing of Messages
Mahesh V Chitrao, Ralph Grishman, Statistical Parsing
Eucalyptus: Integrating Natural Language Input wit
Kenneth Wauchope
Aspects of Swedish morphology and semantics fr
of...
Turid Hedlund, Ari Pirkola, Kalervo JaÈrvelin - 2001 - Inform
Management
A Survey of Current Paradigms in Machine Transl
Bonnie J Dorr, AV Williams Building, Pamela W Jordan, Jo
Intelligence Center
show more »
Search in this book
Search

Example (II)


Computational Linguistics: An In...
Information Retrieval Thesauru...
CAT.INIST
Map of Mountain View, CA by M...
Amazon.com: Information ...
Shipping Weight: 2 pounds (View shipping rates and policies)
Amazon.com Sales Rank: #462,107 in Books (See Top Sellers in Books)
(Publishers and authors: improve your sales)
Would you like to update product info or give feedback on images? (We'll ask you to sign in so we can get back to you)
Inside This Book (learn more)
First Sentence:
The goal of this book is to present the field of information retrieval (IR), with an emphasis on the health and biomedical domain. Read the first page
Key Phrases - Statistically Improbable Phrases (SIPs): (learn more)
clinician searchers, medical language processing, nonrelevant documents, literature reference databases, interactive track, bibliographic content, cosine normalization, measured usage, relevance judgments, multiword terms, passage retrieval, healthcare domain, filtering track, novice searchers, controlled vocabulary terms, clinical narrative, clinical queries, medical language system, query expansion, expert searchers, indexing terms, automated indexing, clinical vocabularies, manual indexing, searching performance
Key Phrases - Capitalized Phrases (CAPs): (learn more)
National Library of Medicine, Interactive Track, Semantic Web, Grateful Med, Open Directory, British Medical Journal, Index Medicus, Medical Matrix, United States, Cochrane Collaboration, Research Directions, Annals of Internal Medicine, World Wide Web, Oregon Health, Wall Street Journal, Associated Press, Brown Corpus, Highwire Press, National Guidelines Clearinghouse, Question-Answering Track, Yearbook Series, Knowledge Finder, Merck Manual, National Heart, Science University
New!
Books on Related Topics | Concordance | Text Stats
Browse Sample Pages:
Front Cover | Copyright | Table of Contents | Excerpt | Index | Back Cover | Surprise Me!
Search Inside This Book:
GO!
Done
Adbloc

Genetische Algorithmen

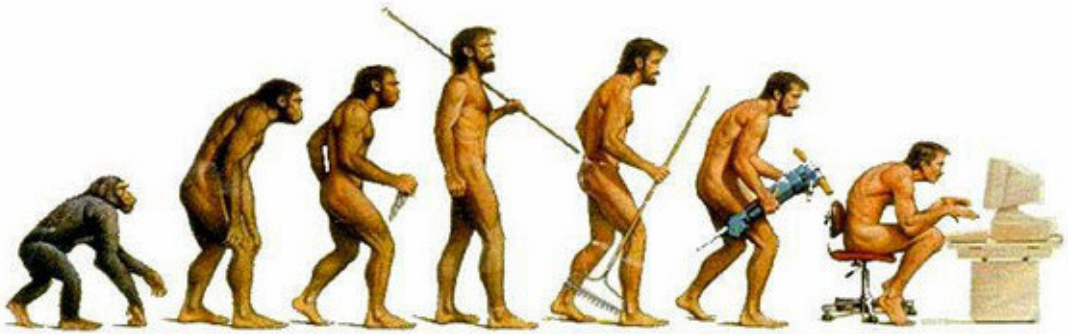
Analogie mit Evolutionstheorie → Simulation der natürlichen Selektion

Betrachte Problemlösung als **Individuum** in einer **Population**

Evaluierungsfunktion ist notwendig, um die Fitness eines Individuums zu bestimmen

Erlaube dem “fittesten” Individuum zu **reproduzieren**

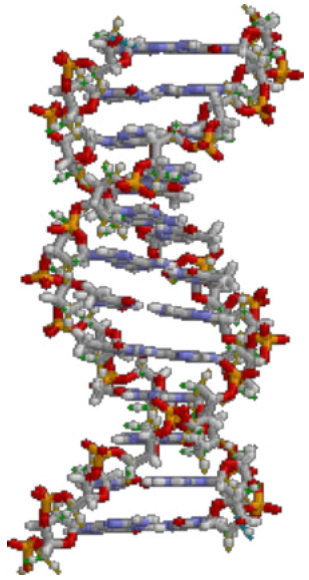
Letztendlich sollte die Population fittere Individuen haben



Konfiguration

Problem muss als **Konfiguration** repräsentiert werden können

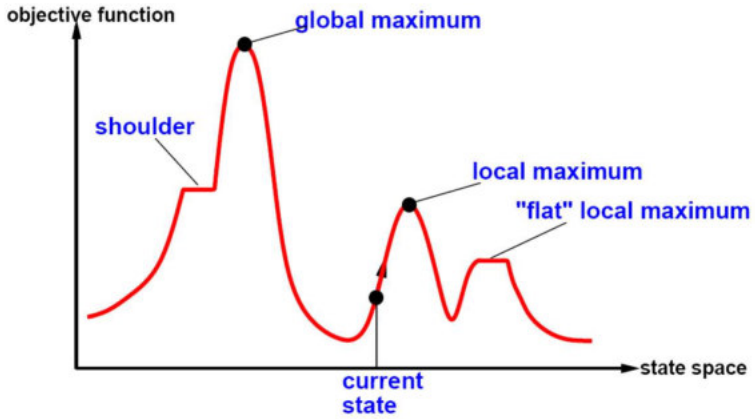
$X =$ | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |



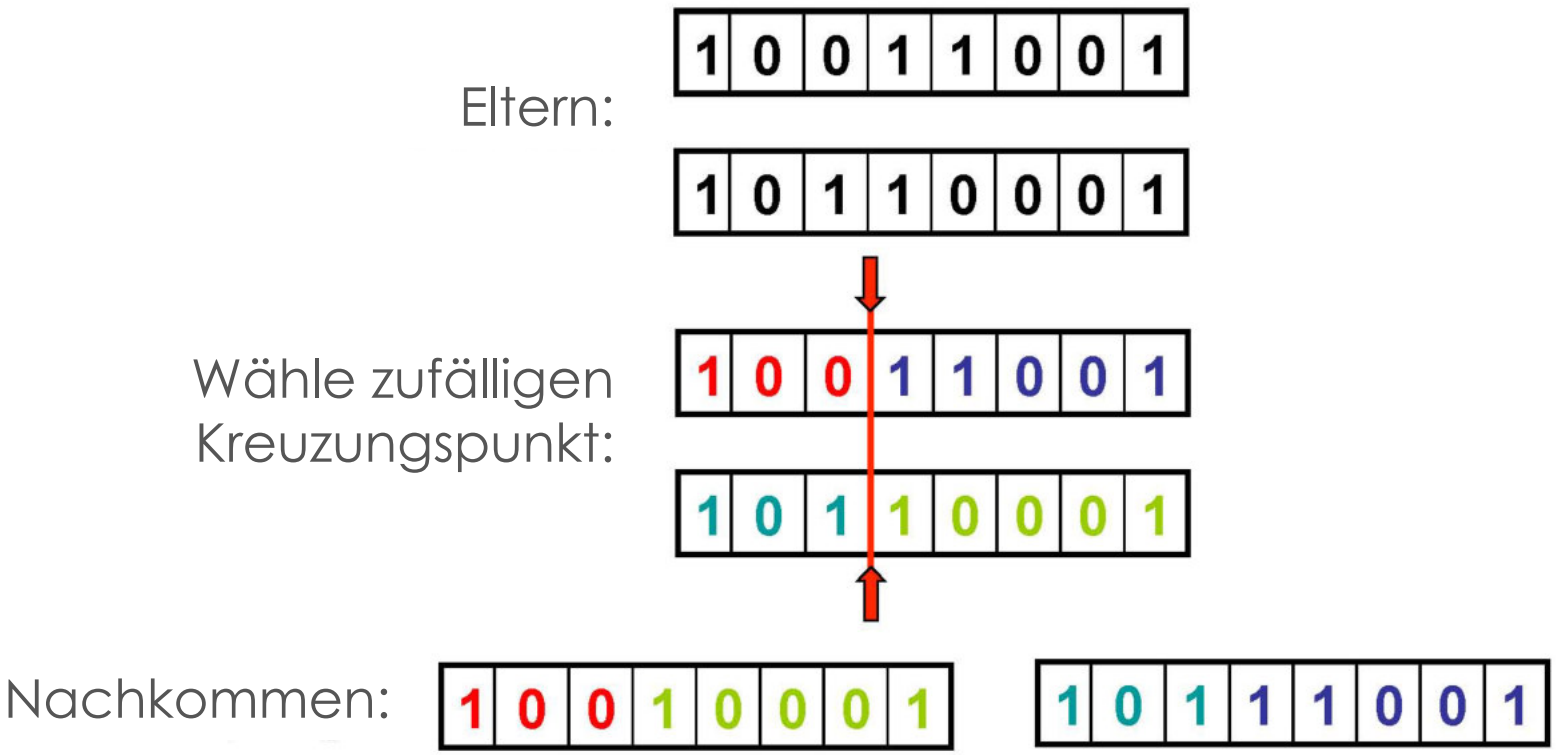
Analogie in Genetik

Starte mit zufälliger Population und wende Operationen an:

- **Kreuzung**
  - Bergsteiger-Komponente
- **Mutation**
  - Zufalls-Komponente



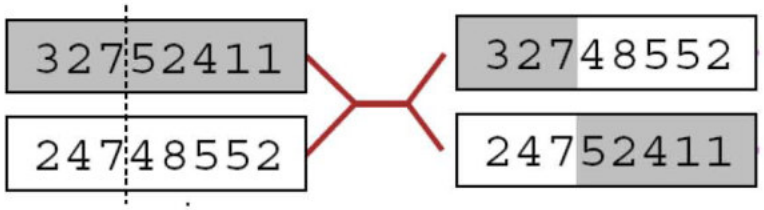
Kreuzung


Eltern:
1 0 0 1 1 0 0 1
1 0 1 1 0 0 0 1
Wähle zufälligen Kreuzungspunkt:
1 0 0 1 1 0 0 1
1 0 1 1 0 0 0 1
Nachkommen:
1 0 0 1 0 0 0 1
1 0 1 1 1 0 0 1

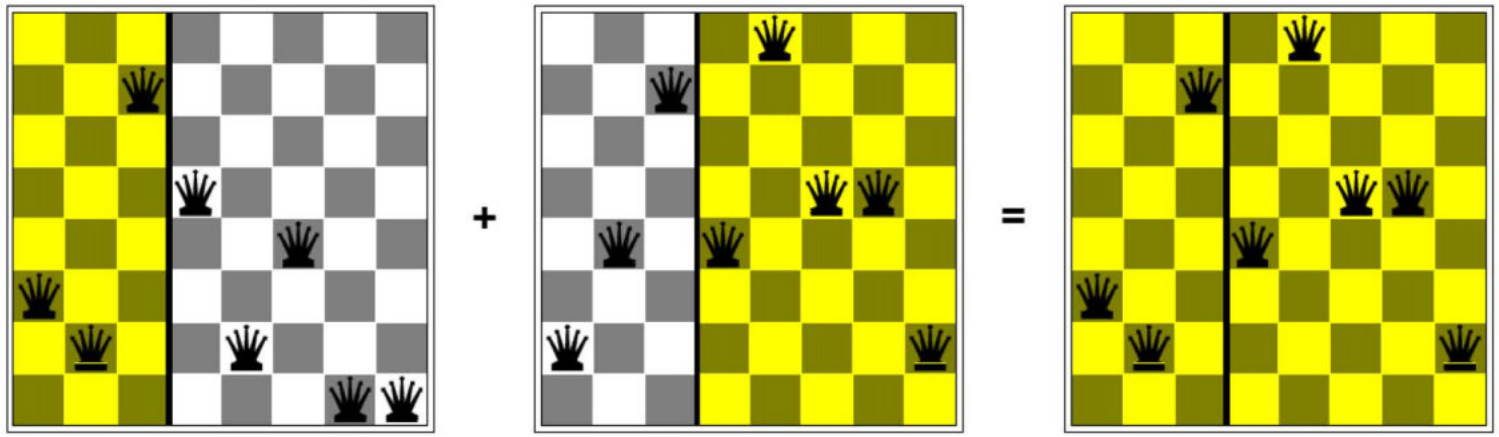
Kreuzung

Modelliert nach Kreuzung der DNA

- Nehme zwei Eltern-Konfigurationen
- Teile sie in zwei Hälften
- Verbinde die Teile

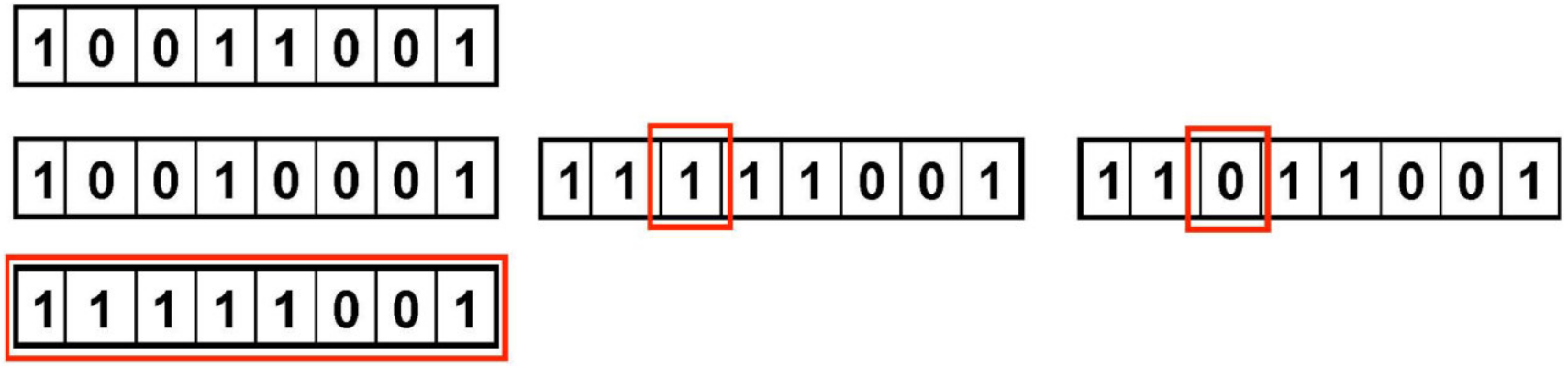


Es ist hilfreiche, wenn Hälften bedeutungsvolle Sub-Konfigurationen sind



Mutation

- Zufällige Wechsel von *n* Elementen in einer Konfiguration
  - Mutation wird als zufällige Änderung der DNA repräsentiert
  - Korrespondiert lose mit Zufallsbewegungen



Nehme ein zufälliges Individuum

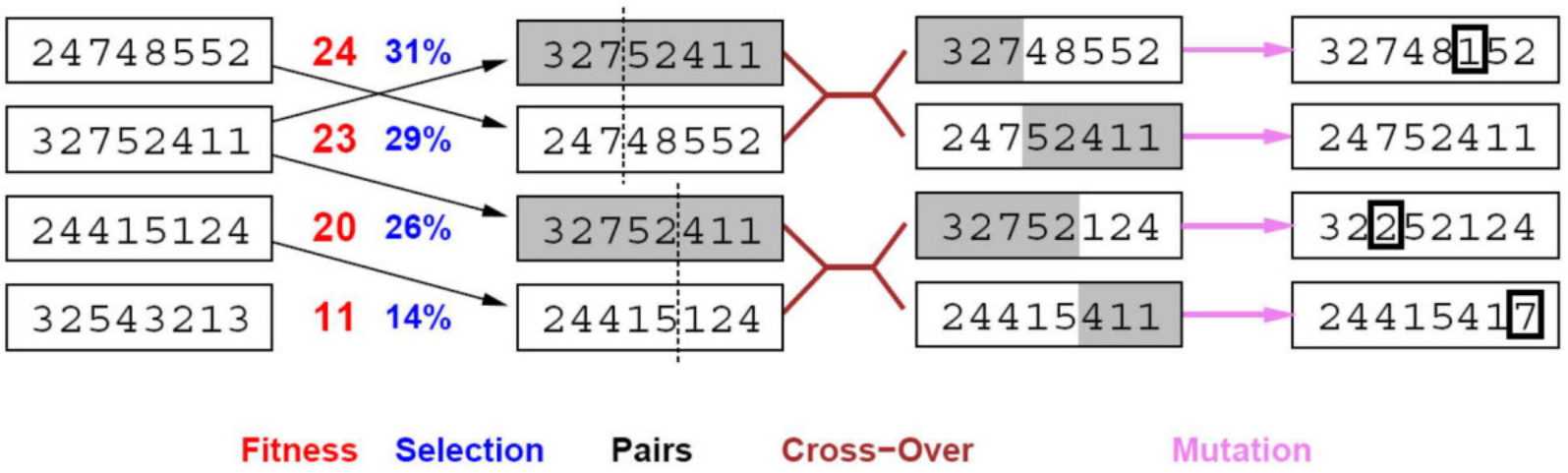
Wähle einen zufälligen Eintrag aus

Verändere den Eintrag

Genetische Algorithmen

Basic Algorithm

- Start with *k* randomly generated states (**population**)
- A state is represented as a string over a finite alphabet
  - often a string of 0s and 1s
- Evaluation function (**fitness function**)
- Produc



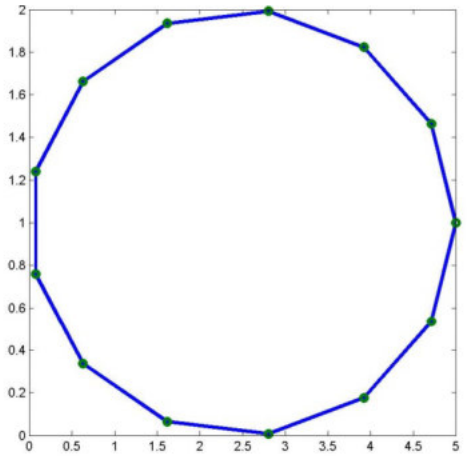
Genetische Algorithmen

```
function GENETIC-ALGORITHM(population, FITNESS_FN)
returns an individual
  inputs: population, a set of individuals
          FITNESS_FN, a function determining the quality of the individual

  repeat
    new_population ← empty set
    loop for i from 1 to SIZE(population) do
      x ← RANDOM-SELECTION(population, FITNESS-FN)
      y ← RANDOM-SELECTION(population, FITNESS-FN)
      child ← REPRODUCE(x, y)
      if (small random probability) then child ← MUTATE(child )
      add child to new_population
    population ← new_population
  until some individual is fit enough or enough time has elapsed
  return the best individual in population, according to FITNESS_FN
```

Wieviel?

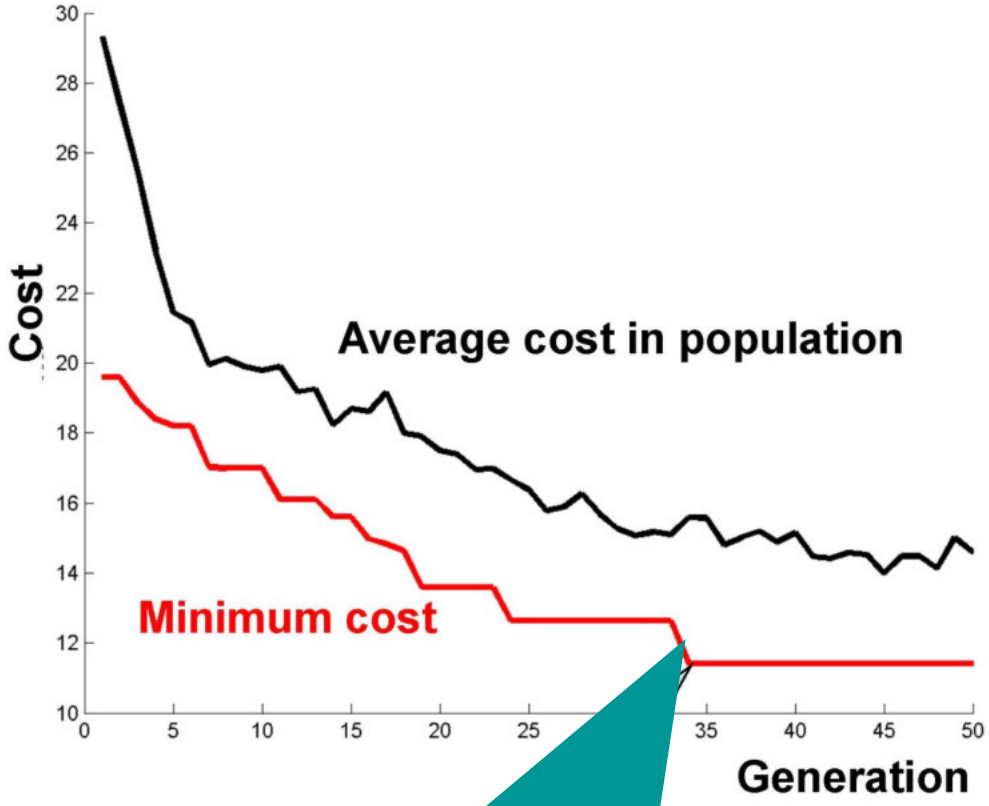
Unser Weihnachtsmann



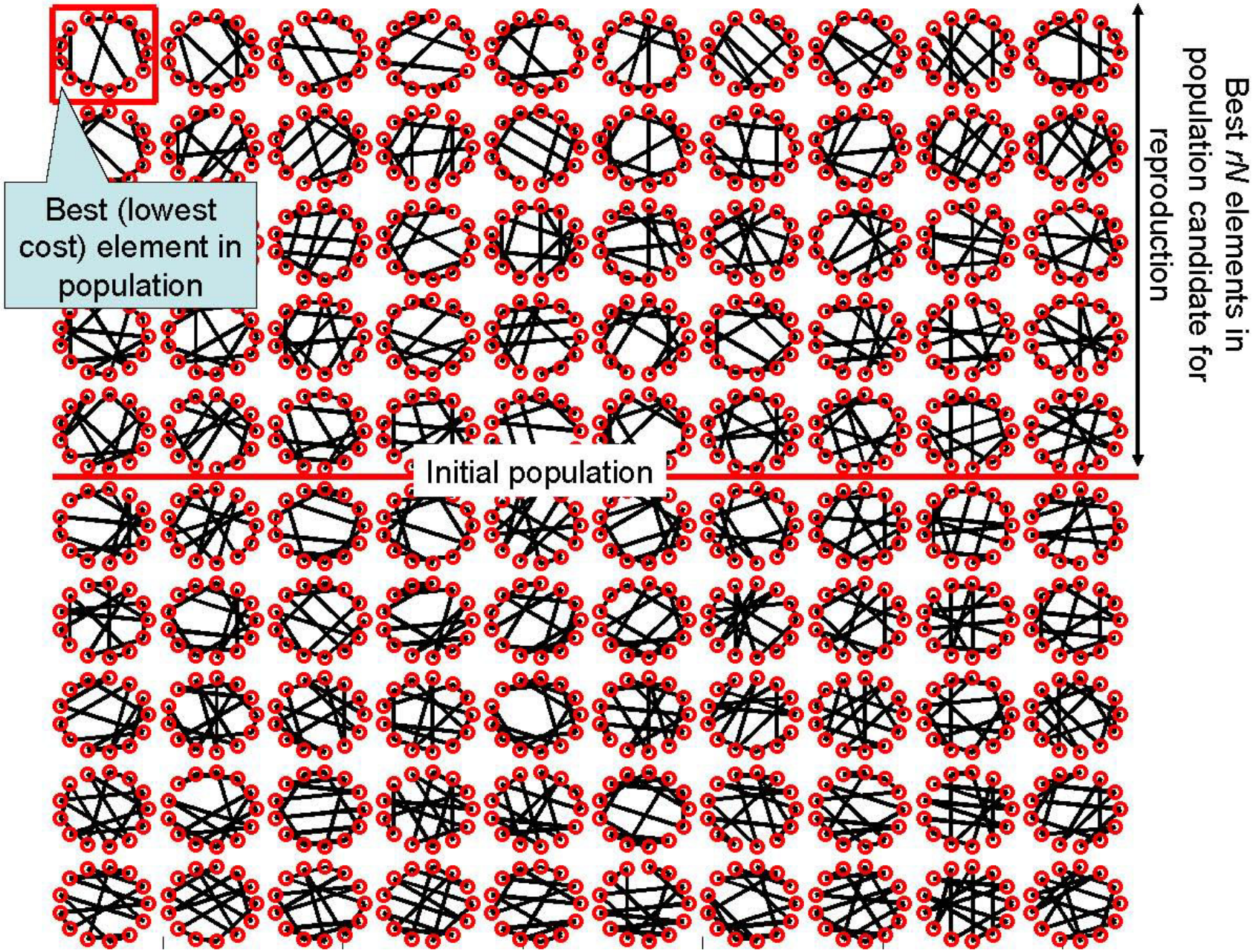
N = 13

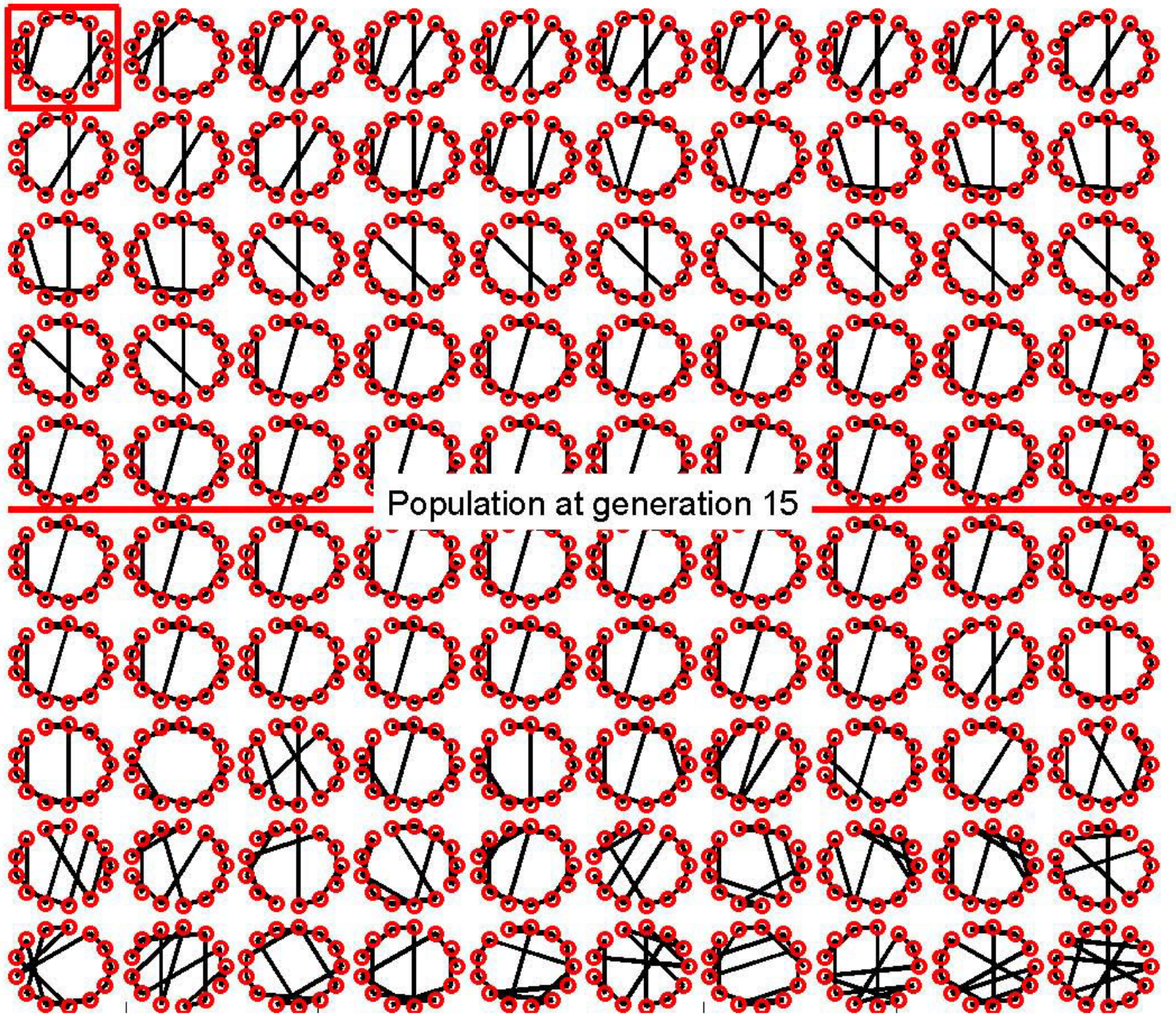
P = 100 Elemente in der Population

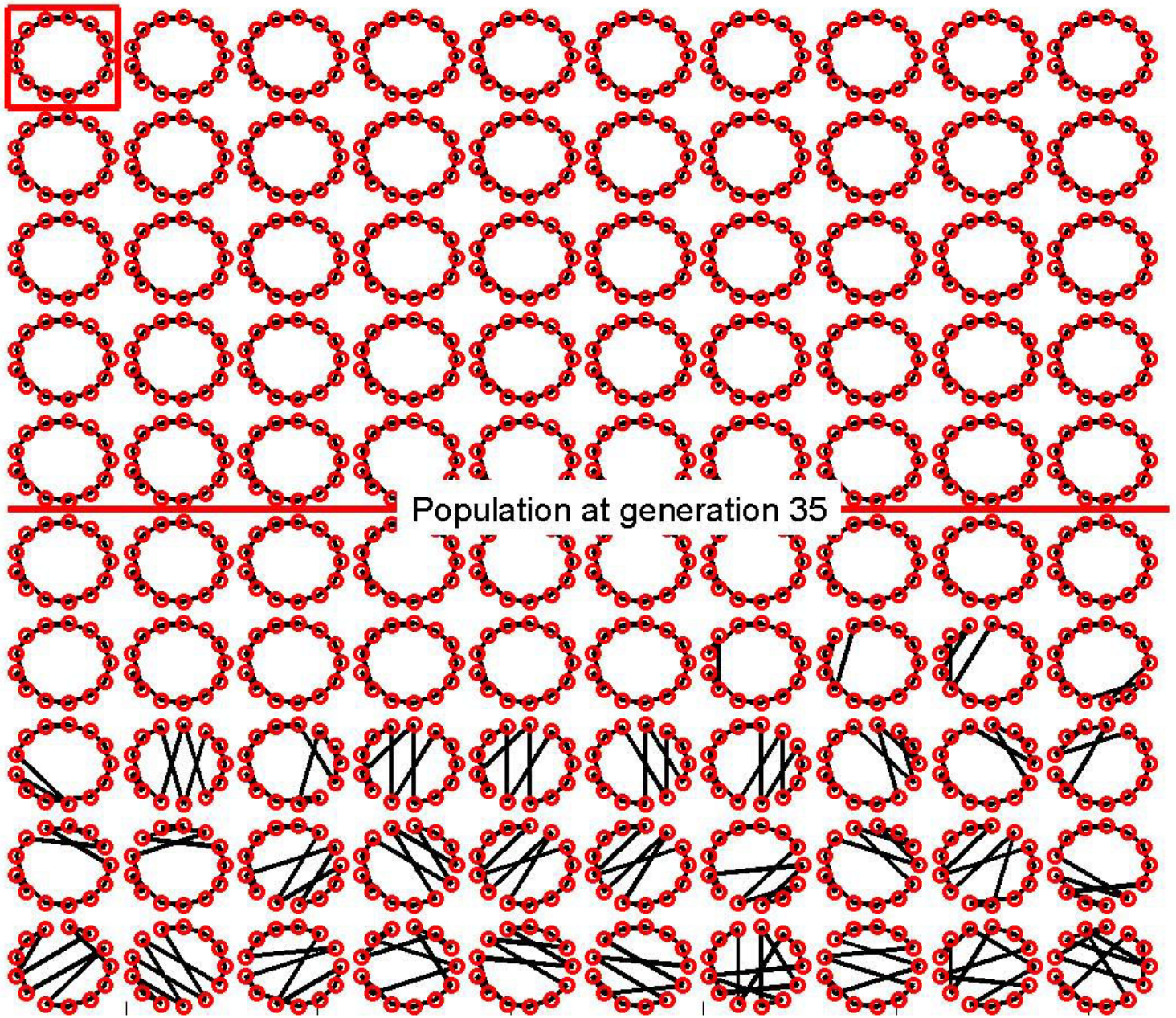
μ = 4% Mutationsrate
r = 50% Reproduktionsrate



Optimale Lösung nach 35 Generationen gefunden


Best (lowest cost) element in population
Initial population
Best rN elements in population candidate for reproduction


Population at generation 15



© INFOMOTION Gm

Der Weihnachtsmann doch nicht alle Kinder erkennen!

Will Ferrell oder Chad Smith?

Will Ferrell oder Chad Smith?

Wie erkennt er nur Will Ferrell?

Quelle: https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

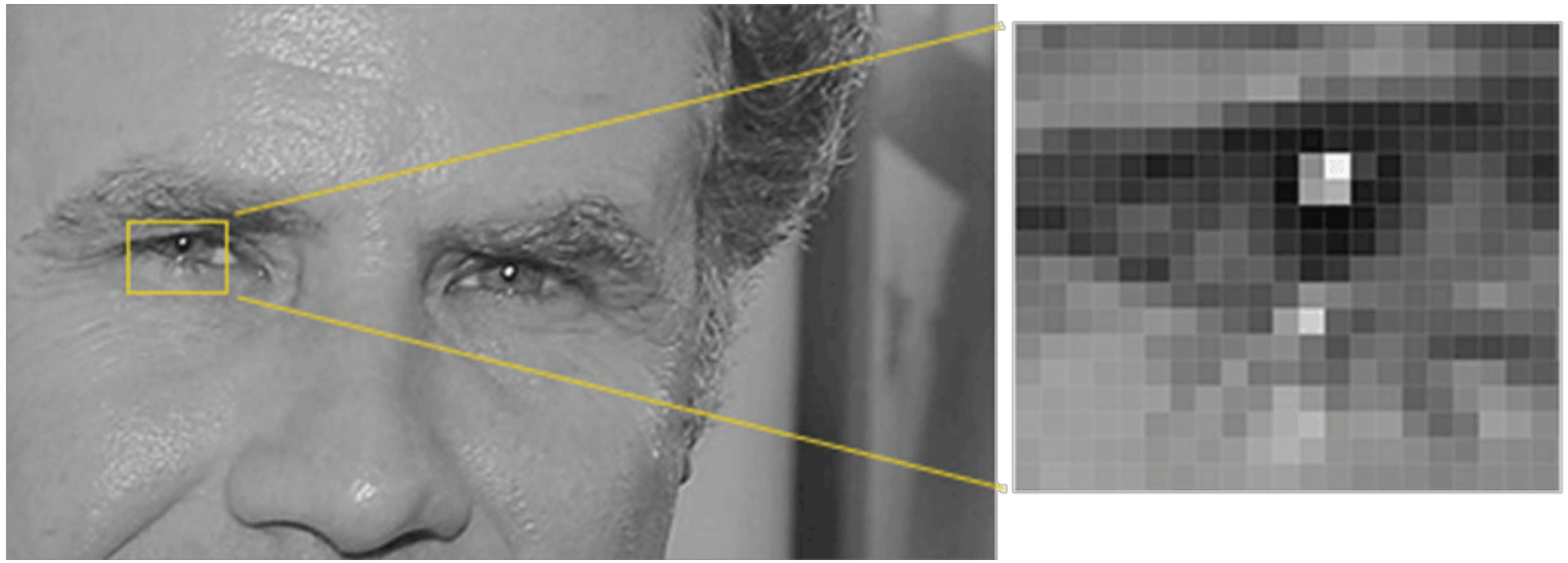
Face Detection

HOG – Detektor

Histogram of
Oriented
Gradients

Face Detection (HOG)

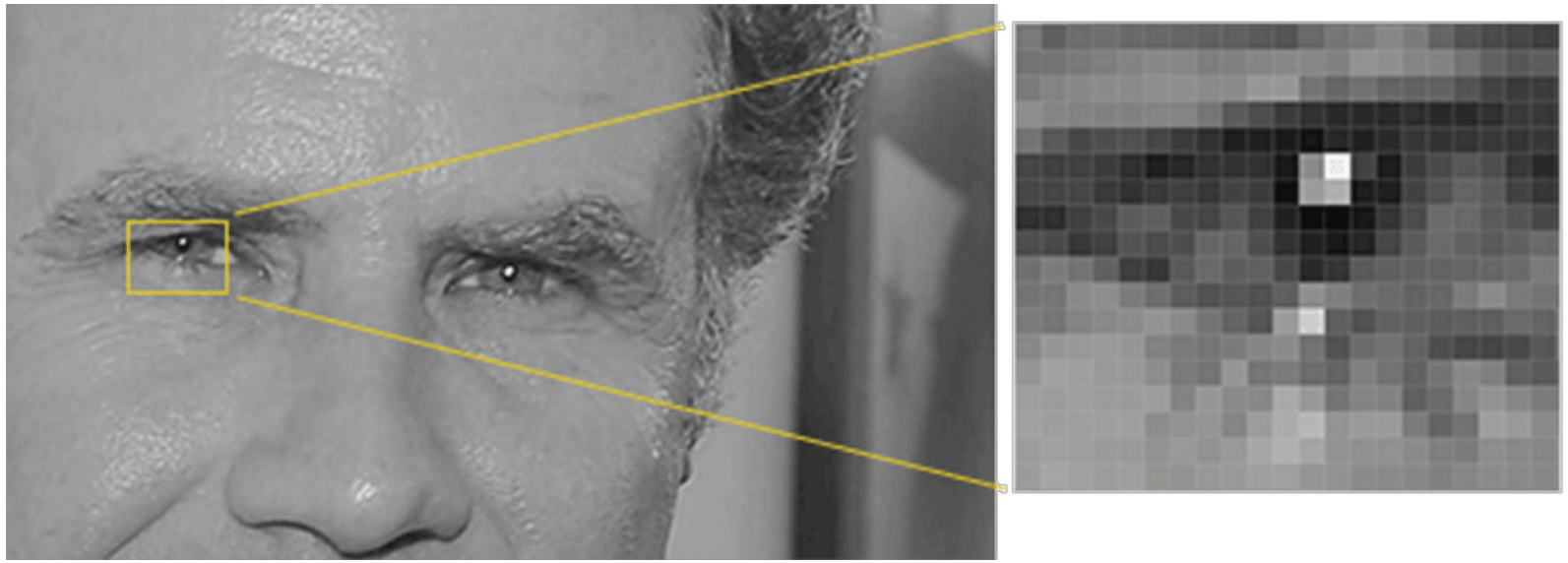
› Nachbarschaft:



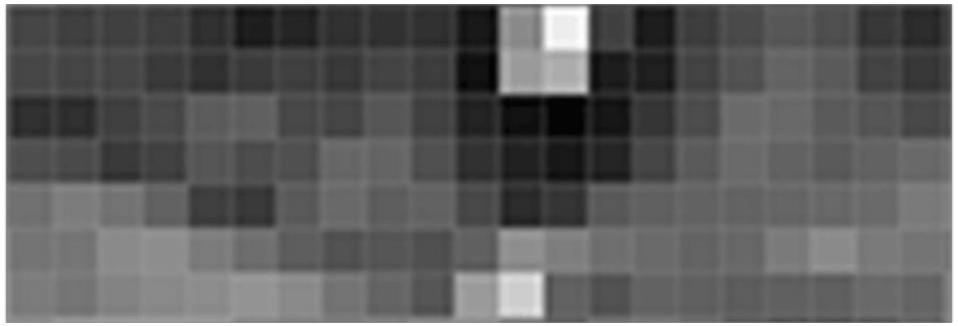
Quelle: https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Face Detection (HOG)

› Nachbarschaft:



› Gradient:



Quelle: https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Face Detection (HOG)

› Ein größerer Bildausschnitt würde dann schon so aussehen:

https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

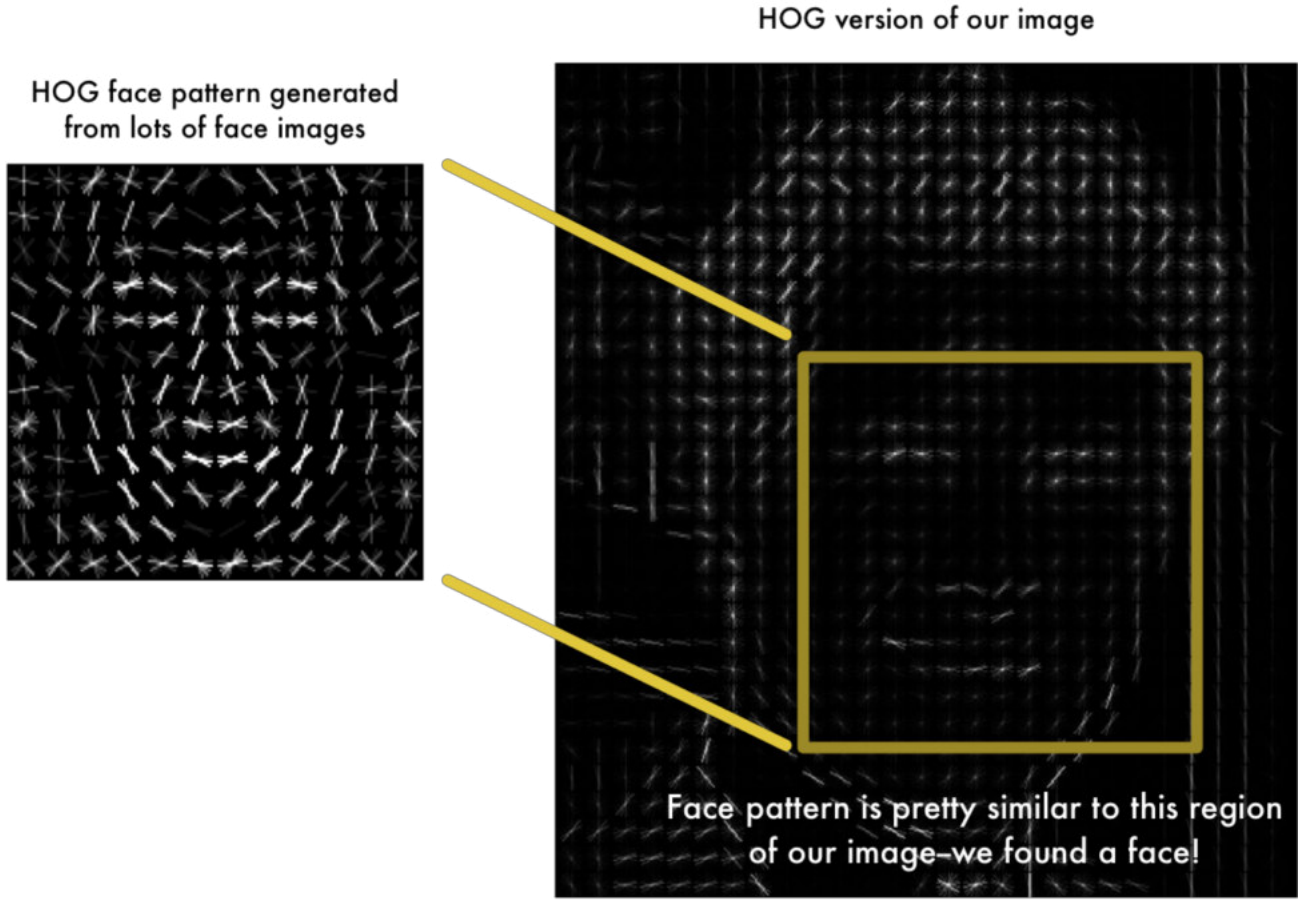
Face Detection (HOG)

› Gesamtbild (Reduziert auf 16x16 Pixel):

https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

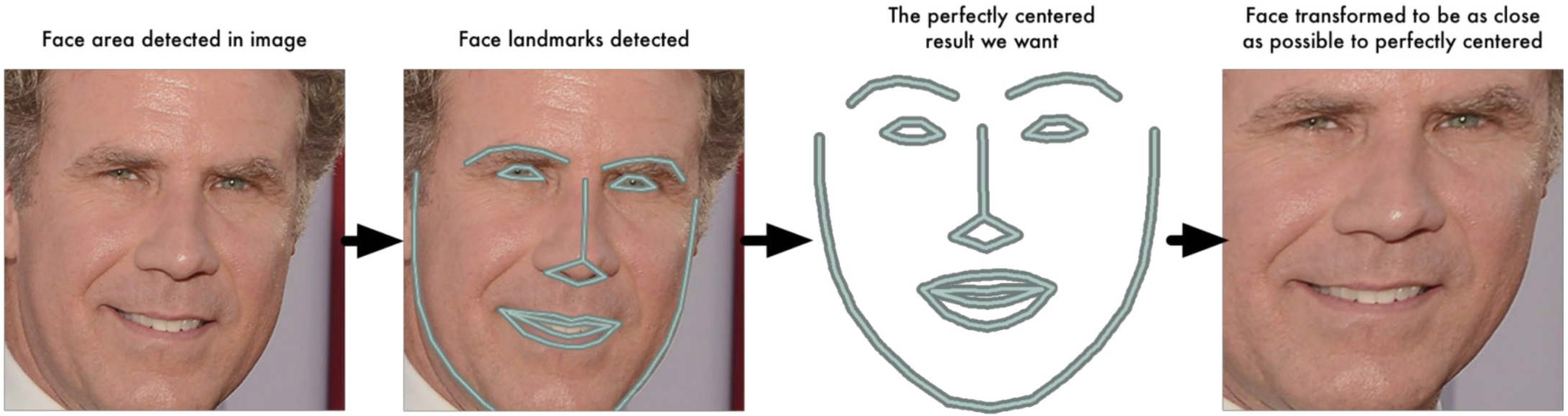
Face Detection (HOG)

› HOG basierte Detektion:



https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Face Recognition (Posing and Projecting Faces)



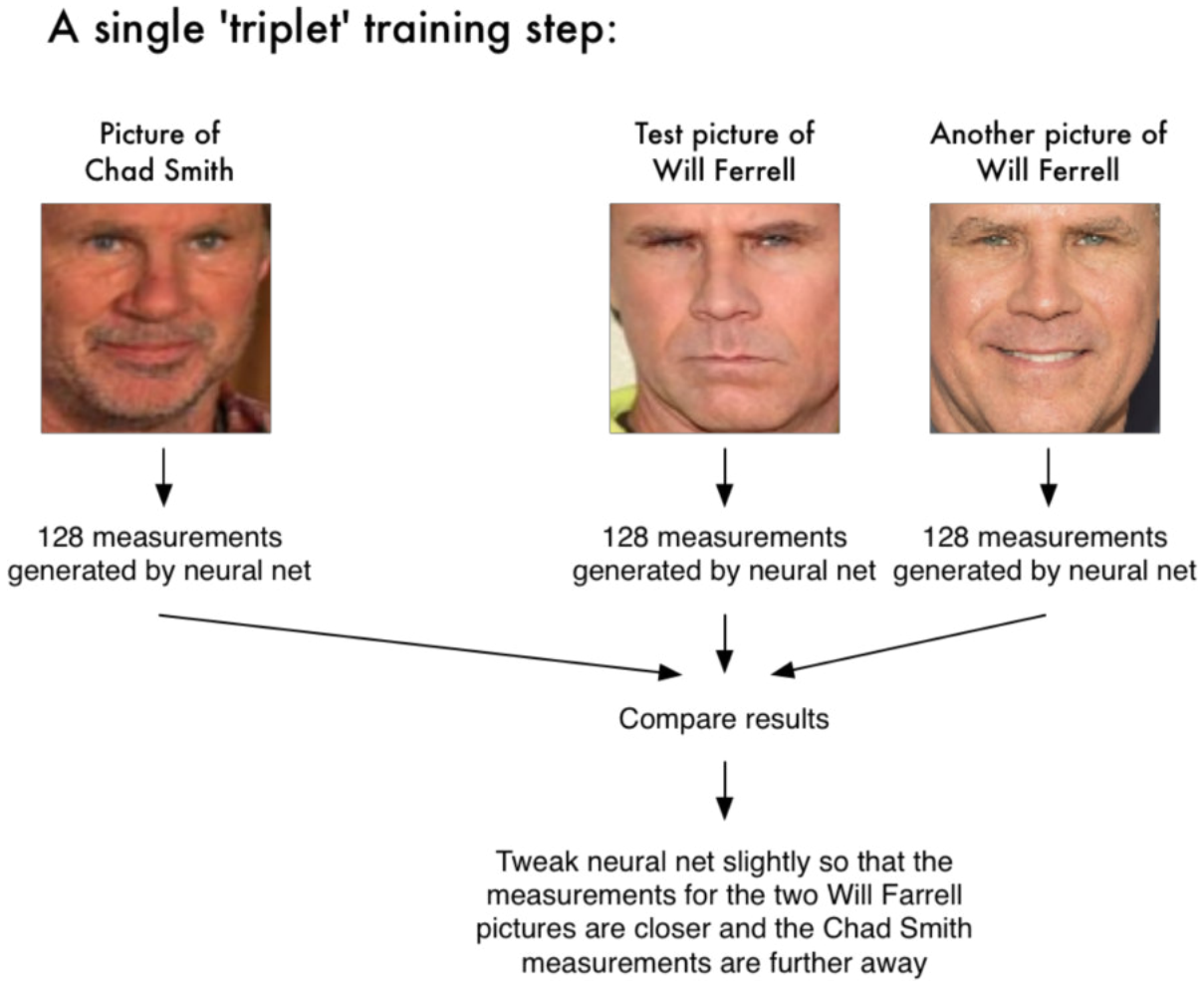
https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Face Encoding

CNN

Convolutional
Neural
Network

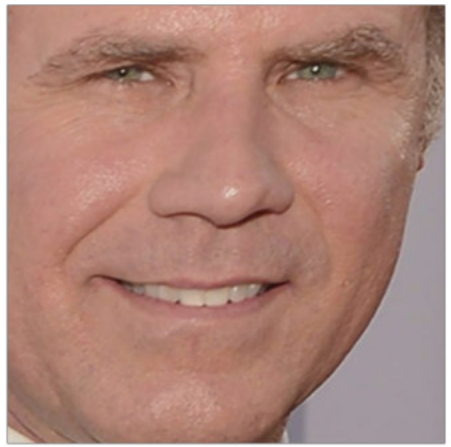
Face Recognition (Encoding Faces/CNN)



https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Face Recognition (Encoding Faces/CNN)

Input Image



128 Measurements Generated from Image

0.097496084868908	0.045223236083984	-0.1281466782093	0.032084941864014
0.12529824674129	0.060309179127216	0.17521631717682	0.020976085215807
0.030809439718723	-0.01981477253139	0.10801389068365	-0.00052163278451189
0.036050599068403	0.065554238855839	0.0731306001544	-0.1318951100111
-0.097486883401871	0.1226262897253	-0.029626874253154	-0.0059557510539889
-0.0066401711665094	0.036750309169292	-0.15958009660244	0.043374512344599
-0.14131525158882	0.14114324748516	-0.031351584941149	-0.053343612700701
-0.048540540039539	-0.061901587992907	-0.15042643249035	0.07819810509681 7
-0.12567175924778	-0.10568545013666	-0.12728653848171	-0.076289616525173
-0.061418771743774	-0.074287034571171	-0.065365232527256	0.12369467318058
0.046741496771574	0.0061761881224811	0.14746543765068	0.056418422609568
-0.12113650143147	-0.21055991947651	0.0041091227903962	0.089727647602558
0.061606746166945	0.11345765739679	0.021352224051952	-0.0085843298584223
0.061989940702915	0.19372203946114	-0.086726233363152	-0.022388197481632
0.10904195904732	0.084853030741215	0.09463594853878	0.020696049556136
-0.019414527341723	0.0064811296761036	0.21180312335491	-0.050584398210049
0.15245945751667	-0.16582328081131	-0.035577941685915	-0.072376452386379
-0.12216668576002	-0.0072777755558491	-0.036901291459799	-0.034365277737379
0.083934605121613	-0.059730969369411	-0.070026844739914	-0.045013956725597
0.087945111095905	0.11478432267904	-0.089621491730213	-0.013955107890069
-0.021407851949334	0.14841195940971	0.078333757817745	-0.17898085713387
-0.018298890441656	0.049525424838066	0.13227833807468	-0.072600327432156
-0.011014151386917	-0.051016297191381	-0.14132921397686	0.0050511928275228
0.0093679334968328	-0.062812767922878	-0.13407498598099	-0.014829395338893
0.058139257133007	0.0048638740554452	-0.039491076022387	-0.043765489012003
-0.024210374802351	-0.11443792283535	0.071997955441475	-0.012062266469002
-0.057223934680223	0.014683869667351	0.05228154733777	0.012774495407939
0.023535015061498	-0.081752359867096	-0.031709920614958	0.069833360612392
-0.0098039731383324	0.037022035568953	0.11009479314089	0.11638788878918
0.020220354199409	0.12788131833076	0.18632389605045	-0.015336792916059
0.0040337680839002	-0.094398014247417	-0.11768248677254	0.10281457751989
0.051597066223621	-0.10034311562777	-0.040977258235216	-0.082041338086128

https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78

Will Ferrell (93%)
Chad Smith (56%)
Jimmy Fallon (83%)

Empirisch beweisen!

You may not realize it, but Artificial Intelligence is all around us.

- Judy Woodruff

Empirisch beweisen!

You may not realize it, but Christmas is all around us.

Quellen

Die Mythen von der wissenschaftlichen Wiederlegung des Weihnachtsmanns

- https://www.familie-ahlers.de/witze/weihnachtsmann.html
- Ruthe
  - http://www.youtube.com/watch?v=ftOdaLxEj74
  - http://www.youtube.com/watch?v=S1C359uhRWs
  - http://www.youtube.com/watch?v=kZknnwl0aAI

Korrelation und Kausalität

- http://www3.math.uni-paderborn.de/~agbiehler/sis/sisonline/struktur/jahrgang21-2001/heft2/Langfassungen/2001-2_Matth.pdf

"They knew what I wanted"

Digital Innovation LAB

INFOMOTION GmbH

NICOLAI ERBS
Senior Consultant bei INFOMOTION, CTO Privollino und Grinch

INFOMOTION GMBH
Ludwigstraße 33-37
60327 Frankfurt

www.infomotion.de

T +49 69 97460-7000
F +49 69 97460-799
M +49 1636375636

nicolai.erbs@infomotion.de

Alle Angaben basieren auf dem derzeitigen Kenntnisstand. Änderungen vorbehalten. Dieses Dokument der INFOMOTION GmbH ist ausschließlich für den Adressaten bzw. Auftraggeber bestimmt. Es bleibt bis zu einer ausdrücklichen Übertragung von Nutzungsrechten Eigentum der INFOMOTION GmbH. Jede Bearbeitung, Verwertung, Vervielfältigung und/oder gewerbsmäßige Verbreitung des Werkes ist nur mit Einverständnis der INFOMOTION GmbH zulässig.