Improved Decipherment of Homophonic Ciphers

Malte Nuhn, Julian Schamper, and Hermann Ney
Human Language Technology and Pattern Recognition, RWTH Aachen University

RWTH AACHEN UNIVERSITY

Problem Definition

Given: **Ciphertext** $f_1^N = f_1 \ldots f_j \ldots f_N$ with $f_j \in V_f$ and **sentence boundaries** $f_0 = f_{N+1} = \$$ together with a **plaintext vocabulary** V_e and a **language model** $p(e)$.

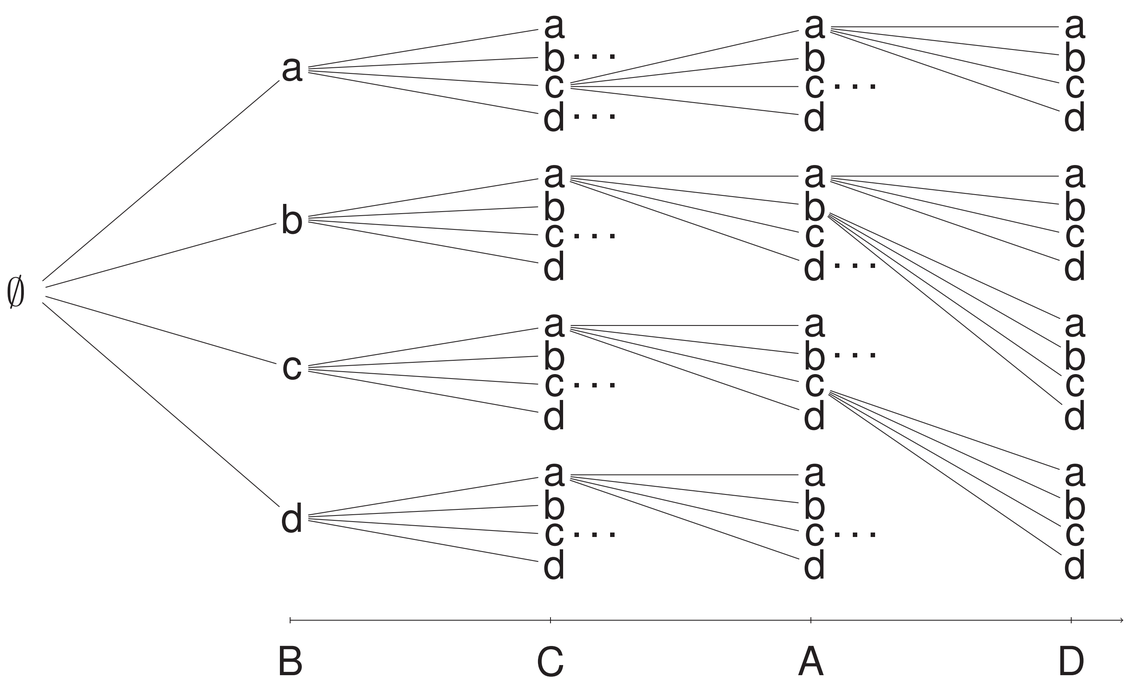
Task: Find a **substitution** (a function $\phi : V_f \to V_e$) that maximizes the probability of the decipherment

$$\hat{\phi} = \arg\max_{\phi} p(\phi(f_1)\phi(f_2)\phi(f_3)...\phi(f_N)) \tag{1}$$

This equation can be further simplified depending on the **structure of the language model** $p(e)$.

Search Space

The space of all possible hypothesis mappings ϕ can be structured as a **tree**. At each level, a decision about how to map a specific cipher symbol to a specific plaintext symbol is made. We call the order in which those decisions are made **extension order**.



$V_f = \{A, B, C, D\}$, $V_e = \{a, b, c, d\}$, extension order (B, C, A, D)

Beam Search Approach

The key idea is to do a **pruned breath first search** through the tree. At each level of the tree, a subset of ϕ's is discarded. The **score estimation function** evaluates **partial hypothesis mappings** ϕ. Only at the leaves of the search tree, the full mapping is available and Equation 3 can be fully evaluated.

Improved Extension Order

Clever ciphers distribute cipher symbols in such a way that only very few **contiguous chunks of deciphered symbols** are available when fixing a small subset of the cipher vocabulary. For example, after fixing the plaintext symbols for *A* and *B* we obtain only four individual unigrams:

```
A C B D A C B D
a . b . a . b .
```

while this cipher would yield one 4-gram:

```
A A B B C C D D
a a b b . . . .
```

For the second case, we can use **more context to evaluate the partial hypothesis**, and can thus perform **more reliable pruning**. Previously, the most frequent remaining cipher symbol was chosen next. We propose to **perform beam search over all possible extension orders**. For each cardinality, we maximize a weighted sum of the number of fixed 1-gram, 2-gram, ... The best partial extension orders are then further expanded until the best full extension order is obtained.

Improved Score Estimation

For n-gram LMs (e.g. $n = 3$), Equation 1 simplifies to

$$\hat{\phi} = \arg\max_{\phi} \left\{ \prod_{j=1}^{N+1} p(\phi(f_j)|\phi(f_{j-2})\phi(f_{j-1})) \right\} \tag{2}$$

$$= \arg\max_{\phi} \left\{ \sum_{f \in V_f} \sum_{f' \in V_f} \sum_{f'' \in V_f} N_{ff'f''} \, S_{\phi(f)\phi(f')\phi(f'')} \right\} \tag{3}$$

$$N_{ff'f''} = \sum_{i=1}^{N+1} \delta(f, f_{i-2})\delta(f', f_{i-1})\delta(f'', f_i)$$

$$S_{ee'e''} = \log p(e''|ee')$$

$N_{ff'f''}$ are ***n*-gram counts** of tokens found in the ciphertext

$S_{ee'e''}$ are the plaintext **log *n*-gram probabilities**

For partial hypotheses, **not all terms in above's sum can be evaluated**, since the mapping $\phi(\cdot)$ might **not yet been defined**. Previously, those undefined terms were simply dropped, yielding an admissible heuristic. Instead, we use **lower order *n*-gram models** as substitute for those undefined terms:

$$p(\, e'' \mid e\ e' \,) \to p_3(e''|ee') \qquad p(\, * \mid e\ e' \,) \to 1$$
$$p(\, e'' \mid *\ e' \,) \to p_2(e''|e') \qquad p(\, * \mid *\ e' \,) \to 1$$
$$p(\, e'' \mid *\ * \,) \to p_1(e'') \qquad p(\, * \mid *\ * \,) \to 1$$
$$p(\, e'' \mid e\ * \,) \to 1 \qquad p(\, * \mid e\ * \,) \to 1$$

Note that this is not an interpolation with lower order *n*-grams. Each symbol is scored only using the maximum amount of context available. Also note: this **new heuristic is not admissible**.

Results

Decipherment of the Zodiac Z-408
We decipher the Zodiac-408 with a beam size < 100 and an LM of order 8 in less than 10s on a single CPU. Using the previously published heuristic, which required a beam size of several million, 48h of CPU time were needed.

Decipherment of part two of the Beale Ciphers
Compared to the Zodiac-408, which has length 408 while having 54 different symbols (7.55 observations per symbol), part two of the Beale Ciphers has length 762 while having 182 different symbols (4.18 observations per symbol). Using a beam size of 10M we achieve a decipherment accuracy of 157 out of 185 symbols correct (symbol error rate $< 5.4\%$). To the best of our knowledge, this is the **first automatic decipherment** of part two of the beale ciphers. We show a **relabeled subset of the cipher**:

$i_{02}\ h_{08}\ a_{03}\ v_{01}\ e_{05}\ d_{09}\ e_{07}\ p_{03}\ o_{07}\ s_{10}\ i_{11}\ t_{03}\ e_{14}\ d_{03}\ i_{03}\ n_{05}\ t_{06}\ h_{01}\ e_{13}\ c_{04}$
$o_{10}\ u_{01}\ n_{01}\ t_{04}\ y_{01}\ o_{12}\ f_{04}\ b_{04}\ e_{15}\ d_{09}\ f_{03}\ o_{04}\ r_{06}\ d_{04}\ a_{07}\ b_{07}\ o_{09}\ u_{03}\ t_{13}\ f_{01}$
$o_{01}\ u_{08}\ r_{05}\ m_{03}\ i_{08}\ l_{09}\ e_{14}\ s_{06}\ f_{01}\ r_{05}\ o_{07}\ m_{04}\ b_{06}\ u_{02}\ f_{04}\ o_{10}\ r_{07}\ d_{01}\ s_{11}\ i_{03}$
$n_{02}\ a_{06}\ n_{03}\ e_{05}\ x_{01}\ c_{03}\ a_{01}\ v_{01}\ a_{03}\ t_{10}\ i_{13}\ o_{03}\ n_{05}\ o_{08}\ r_{06}\ v_{01}\ a_{08}\ u_{03}\ l_{01}\ t_{11}$
$s_{12}\ i_{04}\ x_{01}\ f_{01}\ e_{01}\ e_{03}\ t_{02}\ b_{06}\ e_{07}\ l_{02}\ o_{11}\ w_{06}\ t_{08}\ h_{08}\ e_{15}\ s_{06}\ u_{04}\ r_{06}\ f_{04}\ a_{10}$
$c_{05}\ e_{03}\ o_{01}\ f_{05}\ t_{14}\ h_{05}\ e_{12}\ g_{03}\ r_{02}\ o_{01}\ u_{05}\ n_{02}\ d_{10}\ t_{10}\ h_{07}\ e_{03}\ f_{01}\ o_{08}\ l_{05}\ l_{06}$
$o_{01}\ w_{03}\ i_{09}\ n_{05}\ g_{04}\ a_{10}\ r_{07}\ t_{04}\ i_{02}\ c_{05}\ l_{03}\ e_{03}\ s_{01}\ b_{06}\ e_{15}\ l_{04}\ o_{09}\ n_{01}\ g_{01}\ i_{03}$
$n_{03}\ g_{04}\ j_{01}\ o_{01}\ i_{08}\ n_{08}\ t_{15}\ l_{08}\ y_{01}\ t_{06}\ o_{02}\ t_{01}\ h_{03}\ e_{05}\ p_{01}\ a_{06}\ r_{03}\ t_{02}\ i_{05}\ e_{14}$
$s_{06}\ w_{01}\ h_{04}\ o_{02}\ s_{01}\ e_{06}\ n_{02}\ a_{12}\ m_{01}\ e_{08}\ s_{10}\ a_{09}\ r_{02}\ e_{05}\ g_{01}\ i_{03}\ v_{01}\ e_{01}\ n_{06}\ i_{13}$
$n_{04}\ n_{02}\ u_{05}\ m_{01}\ b_{02}\ e_{13}\ r_{04}\ t_{08}\ h_{03}\ r_{06}\ e_{08}\ e_{09}\ h_{01}\ e_{11}\ r_{02}\ e_{14}\ w_{05}\ i_{04}\ t_{11}\ h_{03}$
$t_{12}\ h_{07}\ e_{10}\ f_{01}\ i_{04}\ r_{03}\ s_{01}\ t_{15}\ d_{09}\ e_{04}\ p_{04}\ o_{12}\ s_{07}\ i_{13}\ t_{01}\ c_{04}\ o_{11}\ n_{06}\ s_{02}\ i_{02}$
$s_{11}\ t_{06}\ c_{04}\ d_{10}\ o_{08}\ f_{02}\ t_{06}\ e_{01}\ n_{05}\ h_{08}\ u_{01}\ n_{06}\ d_{09}\ r_{05}\ e_{13}\ d_{01}\ a_{09}\ n_{04}\ d_{10}\ f_{04}$
$o_{03}\ u_{01}\ r_{01}\ t_{10}\ e_{07}\ e_{13}\ n_{03}\ p_{01}\ o_{10}\ u_{04}\ n_{07}\ d_{04}\ s_{08}\ o_{02}\ f_{06}\ g_{02}\ o_{05}\ l_{01}\ d_{03}\ a_{05}$
$n_{06}\ d_{06}\ t_{01}\ h_{02}\ i_{03}\ r_{04}\ t_{04}\ y_{01}\ e_{01}\ i_{02}\ g_{04}\ h_{08}\ t_{08}\ h_{02}\ u_{03}\ n_{08}\ d_{10}\ r_{02}\ e_{14}\ d_{09}$
$a_{14}\ n_{04}\ d_{09}\ t_{02}\ w_{04}\ e_{05}\ l_{03}\ v_{01}\ e_{05}\ p_{02}\ o_{03}\ u_{08}\ n_{01}\ d_{03}\ s_{06}\ o_{03}\ f_{02}\ s_{10}\ i_{02}\ l_{02}$
$v_{01}\ e_{07}\ r_{05}\ d_{02}\ e_{01}\ p_{04}\ o_{07}\ s_{10}\ i_{12}\ t_{11}\ e_{15}\ d_{10}\ n_{01}\ o_{01}\ v_{01}\ e_{01}\ i_{13}\ g_{01}\ h_{03}\ t_{07}$
$e_{04}\ e_{05}\ n_{03}\ n_{02}\ i_{03}\ n_{05}\ e_{15}\ t_{12}\ e_{05}\ e_{07}\ n_{05}\ t_{17}\ h_{07}\ e_{14}\ s_{11}\ e_{04}\ c_{02}\ o_{08}\ n_{05}\ d_{10}$
$w_{02}\ a_{13}\ s_{09}\ m_{02}\ a_{02}\ d_{03}\ e_{03}\ d_{05}\ e_{15}\ c_{06}\ e_{08}\ i_{08}\ g_{03}\ h_{03}\ t_{09}\ e_{14}\ e_{04}\ n_{04}\ t_{03}\ w_{04}$
$e_{14}\ n_{01}\ t_{10}\ y_{01}\ o_{01}\ n_{04}\ e_{06}\ a_{02}\ n_{03}\ d_{04}\ c_{07}\ o_{07}\ n_{01}\ s_{07}\ i_{03}\ s_{05}\ t_{15}\ e_{11}\ d_{07}\ o_{06}$
$f_{07}\ n_{02}\ i_{08}\ n_{04}\ e_{04}\ t_{07}\ e_{08}\ e_{09}\ n_{01}\ h_{07}\ u_{02}\ n_{07}\ d_{05}\ r_{05}\ e_{05}\ d_{09}\ a_{11}\ n_{05}\ d_{07}\ldots$

Lehrstuhl für Informatik 6 – Computer Science Department – RWTH Aachen University – Aachen, Germany
www: http://www-i6.informatik.rwth-aachen.de
e-mail: <surname>@cs.rwth-aachen.de