Machine reconstruction of the Playfair cipher

Burton Rosenberg

September 8, 2004

Introduction

A Playfair cipher is a digraph substitution, where each letter pair of the plaintext is replaced by a letter pair in the ciphertext. The chart of substitutions is summarized by the placement of letters in a table, called the *Playfair square*, constructed in some customary manner from a keyword. We denote a substitution of plaintext digraph cd with ciphertext digraph ab by $ab - cd$. There are two rules to the substitution, depending on whether cd lie in the same row or column, or if they are simultaneously in different rows and columns. The cases are illustrated:

```
a * c
*   *     a square is formed
d * b

c a * d b   c and d in a common row

c
a
*     c and d in a common column
d
b
```

When the digraph shares a row or column, the shift is circular, e.g. if d were at the bottom of a column then b would be the letter at the top of that same column. We say of the first type of substitution that the letters are *arranged in a square*, of the second, that they are *arranged in a row*, and of the third, that they are *arranged in a column*.

It is not possible to encode a double letter, such as aa, and such digraphs need to be broken up, customarily by inserting the letter `X`, i.e., a`X`a. What happens for messages discussing matters involving the text XXX is not customarily specified.

The obstruction theorems

The Playfair cipher can be expressed algebraically using symbols for in the same row, $=_r$, in the succeeding row, σ_r, in the same column, $=_c$, and in the succeeding row, σ_c. The encipherment rules can be expressed using these symbols. For example, if $a =_c b$ then the letter pair ab enciphers to $\sigma_c(a)\sigma_c(b)$; if $a \neq_c b$ and $a \neq_r b$, then ab enciphers to xy where x and y are the unique solutions of the equations $x =_r a$, $y =_r b$, $x =_c b$, $y =_c a$.

Because encipherment can be described by these relations, and they are invariant to cyclic permutation of rows and columns, we have the fact that the Playfair cipher is invariant to circular permutations in the rows and columns of the Playfair square.

A further important observation is that if $ab - cd$ is a substitution then $ba - dc$ is a substitution.

Given a pair of known or proposed substitutions, they might be combined to give information on the arrangement of the substitutions. In effect, substitutions obstruct the possible arrangements of other substitutions.

Theorem 1 (Obstruction theorem IA) *Let a, b, c, d, e and f be variables representing distinct letters. If $ab - cd$ and $ae - fb$ are both substitutions then $ab - cd$ is arranged in a column and $ae - fb$ is arranged in a square. The theorem also holds if $ae - fb$ is replaced by any of $ea - bf$, $be - fa$, $eb - af$. $ce - fd$, $ec - df$, $de - fc$ or $ed - cf$.*

Theorem 2 (Obstruction theorem IB) *Let a, b, c, d, e and f be variables representing distinct letters. If $ab - cd$ and $ae - bf$ are both substitutions then $ab - cd$ is arranged in a row and $ae - bf$ is arranged in a square. The theorem also holds if $ae - bf$ is replaced by any of $ea - fb$, $be - af$, $eb - fa$, $ce - df$, $ec - fd$, $de - cf$ or $ed - fc$.*

Proof: We show only the proof for the first claim of IA. The variant conditions result from the eight possible applications of the substitution $xy - zt$ into $yx - tz$.

Assume in order to establish a contradiction that $ab - cd$ is arranged in a square. Then c and d are in distinct columns and rows. Therefore $ec - df$ cannot be in a column or a row, and furthermore, if arranged in a square $ec - df$ would have c and d in the same column. So $ab - cd$ cannot be arranged in a square.

If both $ab - cd$ and $ec - df$ are arranged in a row or column, the must be both as a row or column, leading to too many letters in the row or column, so $ec - df$ must be a square.

Having established that $ec - df$ is arranged in a square, c and d are in distinct rows, so $ab - cd$ cannot be arranged in a row. On the other hand, we can arrange $ab - cd$ in a column consistent with $ec - df$ being arranged in a square.

Theorem 3 (Obstruction theorem II) *Let a, b, c and d be variables representing distinct let-*

ters. Then $ab-cd$ and $cd-ab$ are both substitutions if and only if $ab-cd$ is arranged in square. The theorem also holds if $cd-ab$ is replaced by $dc-ba$.

Theorem 4 (Obstruction theorem III) *If the substitution $ab-cd$ contains only three distinct letters, then either $a=d$ or $b=c$ and the arrangement is either in a row or on a column, in which the letters appear consecutively.*

Theorem 5 (Obstruction theorem IV) *Let a, b, c, d, e and f be variables representing distinct letters. If $ab-cd$ and $ae-fc$ are both substitutions then either $ab-cd$ is arranged in a square and $ae-fc$ is arranged in a row or $ae-fc$ is arranged in a square and $ab-cd$ is arranged in a column.*

An example

We look at an example given by Jim Gillogly,

> This message was received by an intercept station in Scotland. The frequency and format indicate that it is a most urgent message from one of our agents who landed a week ago in Norway. His controllers have been unable to read it. Although it clearly uses his backup cipher, the Playfair, the keys assigned to him do not work. We cannot reach him before his normal scheduled transmission in two weeks, so we urgently request that you attempt to decrypt this and let us know the contents. In case it helps, he is carrying materials to assist a previously dropped team in their work regarding the Norsk Hydro facility at Rjukan. His recognition code should appear in the message: It is "beware ice weasels." If he is operating under duress, he should include the phrase "red penguin frenzy." He will use "STOP" between sentences and "END" at the end.

See `http://www.pbs.org/wgbh/nova/decoding/faceoff.html`. This is the ciphertext,

```
VY TE SY ED LU TE RV LF
NV UH DW AR DL CF FB SD
EW NP XK IC FT RE OL KA
LZ YL SL TO BK EV LY AR
MK RB OD NA LD YP LA ET
OL QA DF HS FZ WN AI DS
MU RU OL HR YL LO TW FY
LD IC VL US VS SF ZY LU
NF FX LK TG BC DO BF AL
EW RP FY WL HU LD AR LI
TF LA BF FZ CY FU UF BG
```

We first discover the placement of the crib. With luck, only one of the recognition code or the distress signal will have any potential of matching and only in one particular location — else we will have to proceed under several tentative suppositions as to the true match. We must consider several cases: that the crib is broken on either odd or even letters; that the crib is in the middle of the text surrounded by STOP; that the crib is at the beginning of the text; that the crib is at the end of the text followed by either END or ENDX.

There are some methods for finding the crib. Since their is no sequence xy zs appearing twice in the cipher text, then STOP is broken as S TO P, and so we can look for a pair xy separated by 10 digrams. For instance,

```
ET OL QA DF HS FZ WN AI DS MU RU OL HR
?s to pr ed pe ng ui nf re nz ys to p?
```

We get lucky and this is the only possibility. This gives us a collection of substitution pairs to begin our reconstruction of the Playfair square. The *obstruction theorems* stated formally and proven at the end of this note give the following information: FZ-NG is arranged as a square; MU-NZ is arranged in a column; and AI-NF is arranged in a row. Fitting this together, and placing F in the upper left hand corner of the square (which we can do since a cyclic permutation of rows or columns does not affect the cipher),

```
F I * N A
      M
      *
G * * Z
    * U
```

The obstruction theorems also identify DF-ED as arranged as a row or column, and given the partially filled square, it must be a column with D preceding F in its column position. This determines that there is no row between the v̇erb.M. and Z. We also use substitution WN-UI to place W,

```
F I * N A
      M
G   * Z
E W * U
D
```

The reasoning is now less mechanical. Consider RU-YS. There isn't room in the row or column containing U for this substitution to be arranged as a row or column, so it is arranged as a square. This means that Y is in the same column as U, in the only open row. R is in the same row as Y. Considering QA-PR, if this is arranged in a square then R is in the same row as A, not the same row as Y. So QA-PR is not arranged in a square, it is arranged as a column, and so R is in the same column as A We also use substitution HS-PE to place H,

```
F I * N A
H   * M P
G   * Z Q
E W * U S
D   * Y R
```

The obstruction theorems identify `OL-TO` as arranged in a row or column, and considering the space left in the partially filled square, it must be a arranged in a column. The pair `ET-?S` aligns `T` with `S`. Recognizing the word "FINAL", we place it in the rightmost column,

```
F I N A L
H   M P
G   Z Q
E W U S T
D   Y R O
```

At this point we might notice the alphabetic spiral and attempt to complete the pattern. This leaves with V and either B or C to place in the final column. Of the four possible placements, we recognize the word "victory", and this completes the square,

```
F I N A L
H K M P V
G X Z Q C
E W U S T
D B Y R O
```

Note that the steps taken in this paragraph are not formal, depending instead on pattern recognition within a linguistic context.

Automated reconstruction of squares

The goal is to have an explicit description of the method of reconstructing the Playfair square from known or hypothesized substitutions. The first obstruction theorem is the strongest. First, it is definitive about the arrangement in the square of the involved substitutions. Second, it yields an arrangement of large footprint in the Playfair square. Third, it uses a proof by contradiction, a situation of negative reasoning which complicates automated deduction — hence its statement is expected to be particularly helpful to automated reconstruction of the square.

An automated method can employ the collection of four relations: $a =_r b$ stating that a and b are in the same row; $a =_c b$, a similar relation for columns; $\sigma_r(a, b)$ stating that a is the direct successor to b in a row, i.e. in column $(i + 1) \bmod 5$ if b is in column i; and $\sigma_c(a, b)$, a similar relation for columns. The $=_*$ relations are symmetric, reflexive and transitive. We can also write $a \neq_* b$ to

show that a and b are not in the same row or column. If $x =_r y$ and $x =_c y$ then $x = y$. The σ_* relations have natural properties such as: $a =_* b$ and $\sigma_*(b, c)$ implies $\sigma_*(a, c)$. The σ_* are period five operators. We also employ the notation $=_* (a, \ldots, d)$ when all $a, \ldots, d$ are pairwise star equal. The above proofs could be written line by line formally using this notation. We have refrained from doing so, preferring for readability proofs in prose.