Natural Language Steganography and an “AI-complete” Security Primitive

by Richard Bergmair

Oct-03 – Apr-04

University of Derby in Austria
Technische Universität München

conducted under the supervision of
Stefan Katzenbeisser

References

Natural Language Steganography
and an “AI-complete” Security Primitive

for reference, see:

- Richard Bergmair. Towards linguistic steganography: A systematic investigation of approaches, systems, and issues. April 2004. final year thesis. handed in in partial fulfillment of the degree requirements for “B.Sc. (Hons.) of Computer Studies” to the University of Derby. Available online: `http:// bergmair.cjb.net/ pub/ towlingsteg-rep-inoff-a4.pdf`

References

Natural Language Steganography
and an “AI-complete” Security Primitive

for reference, see:

- Richard Bergmair and Stefan Katzenbeisser. Towards human interactive proofs in the text-domain. In Kan Zhang and Yuliang Zheng, editors, *Proceedings of the 7th Information Security Conference*, volume 3225 of *Lecture Notes in Computer Science*, pages 257–267. Springer Verlag, September 2004. Available online: `http:// bergmair.cjb.net/ pub/ towhiptext-proc.pdf`

Motivation

Cryptosystems are designed to protect our sensitive data from evil adversaries.

Cryptosystems are designed to protect our sensitive data from evil adversaries.

Wrong!

Cryptosystems are designed to protect our sensitive data from evil adversaries.

Wrong!

...well, maybe not.

Cryptosystems are designed to protect our sensitive data from evil adversaries.

Wrong!

...well, maybe not.

..., but then again...

Cryptosystems are designed to protect our sensitive data from **evil** adversaries.

What is “**evil**”?

What is “**evil**”?

What is “**evil**”?

“Hacker ethics” is about
a **good** individual in a **bad** society.

What is “**evil**”?

“Hacker ethics” is about
a **good** individual in a **bad** society.

“Witch hunt ethics” is about
a **bad** individual in a **good** society.

Motivation

Cryptosystems are designed to protect our sensitive data from evil adversaries

Motivation

Cryptosystems are designed to protect our sensitive data from evil adversaries

like

- the evil spy intercepting sensitive communication

Cryptosystems are designed to protect our sensitive data from evil adversaries

like

- the evil spy intercepting sensitive communication
- the criminal fraudster replaying banking transactions

Motivation

Cryptosystems are designed to protect our sensitive data from evil adversaries

like

- the evil spy intercepting sensitive communication
- the criminal fraudster replaying banking transactions
- the nosy neighbor reading your mail

Cryptosystems are designed to protect our sensitive data from evil adversaries

like

- the evil spy intercepting sensitive communication
- the criminal fraudster replaying banking transactions
- the nosy neighbor reading your mail

good individual / bad society?

Motivation

Stegosystems are designed to hide our sensitive data from evil adversaries

Stegosystems are designed to hide our sensitive data from evil adversaries

like

- the evil government censor infringing on our right to freedom of opinion and expression.

Stegosystems are designed to hide our sensitive data from evil adversaries

like

- the evil government censor infringing on our right to freedom of opinion and expression.
- the greedy employer limiting our access to computers and anything which might teach us something about the way the world really works.

Stegosystems are designed to hide our sensitive data from evil adversaries

like

- the evil government censor infringing on our right to freedom of opinion and expression.
- the greedy employer limiting our access to computers and anything which might teach us something about the way the world really works.

good individual / bad society!

Motivation

- Evil spies,
- criminal fraudsters, and
- nosy neighbors

do not control your communication channel!

Motivation

- Evil spies,
- criminal fraudsters, and
- nosy neighbors

do not control your communication channel!

- Evil governments and
- greedy employers

do!

A shift in perspectives:

A shift in perspectives:

Alice and Bob do not control their communication channel.

A shift in perspectives:

Alice and Bob do not control their communication channel.

Wendy the warden does!

Motivation

What happens if Eve the eavesdropper intercepts a secure cryptogram?

What happens if Eve the eavesdropper intercepts a secure cryptogram?

Nothing!

What happens if Eve the eavesdropper intercepts a secure cryptogram?

Nothing!

- the evil spy won’t know the sensitive information

What happens if Eve the eavesdropper intercepts a secure cryptogram?

Nothing!

- the evil spy won't know the sensitive information
- the criminal fraudster cannot read the banking transaction

What happens if Eve the eavesdropper intercepts a secure cryptogram?

Nothing!

- the evil spy won't know the sensitive information
- the criminal fraudster cannot read the banking transaction
- the nosy neighbor won't see the contents of your mail

What happens if Wendy the warden intercepts a secure cryptogram?

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

What will witch-hunt ethics assert about the presupposedly bad individual in the good society, who seeks to conceal the content of his communication?

Motivation

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

What will witch-hunt ethics assert about the presupposedly bad individual in the good society, who seeks to conceal the content of his communication?

That Alice and Bob have something evil to hide!

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

- the greedy employer will fire Alice and Bob

Motivation

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

- the greedy employer will fire Alice and Bob
- the evil government will send Alice and Bob to Guantanomo Bay

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

- the greedy employer will fire Alice and Bob
- the evil government will send Alice and Bob to Auschwitz

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

As long as there is a way for Wendy to tell ciphertext from cleartext, Alice and Bob will not live in peace!

What happens if Wendy the warden intercepts a secure cryptogram?

Serious consequences!

As long as there is a way for Wendy to tell ciphertext from cleartext, Alice and Bob will not live in peace!

Solution: Alice and Bob must use steganographic methods, rather than purely cryptographic ones, in order to hide not only the content of a message from the adversary, but its very existence!

Motivation

The difference between a cryptogram and a steganogram, is that a steganogram always appears innocuous to Wendy.

The difference between a cryptogram and a steganogram, is that a steganogram always appears **innocuous** to Wendy.

But what is **innocuous**?

The difference between a cryptogram and a steganogram, is that a steganogram always appears **innocuous** to Wendy.

But what is **innocuous**?

In the simplest case Wendy has a list of innocuous cover symbols.

Introduction

$$C = \{ \textit{Midshires is a nice little city}, \\ \textit{Midshires is a great little city}, \\ \textit{Midshires is a fine little city}, \\ \textit{Midshires is a decent little city}, \\ \textit{Midshires is a wonderful little city}, \\ \textit{Midshires is a nice little town}, \\ \textit{Midshires is a great little town} \}.$$

If $c \in C$, then Wendy knows that c is innocuous.

Introduction

$$M = \{ \textit{I don't like my government!}, \textit{I don't like my internet provider!}, \textit{I don't like my employer!}, \textit{I'm wearing ladies' underwear!} \}.$$

Alice wants to send Bob a message $m \in M$.

Introduction

Instead of enlisting $C = \{$*Midshires is a ...*$\}$:

- We know an innocuous sentence $c = \{$*Midshires is a nice little town*$,\}$
- We have a dictionary, that tells us that the words $\{$*nice*, *great*, *fine*, *decent*, *wonderful*$\}$ and $\{$*city*, *town*$\}$ are **synonymous**, i.e. mean the same.
- We know that, by substituting a word in c by a synonym, we never make an innocuous sentence suspicious, since we do not alter its meaning.

Introduction

Instead of enlisting $C = \{$*Midshires is a ...*$\}$:

$$C \;=\; \textit{Midshires is a} \left\{ \begin{array}{c} \textit{nice} \\ \textit{great} \\ \textit{fine} \\ \textit{decent} \\ \textit{wonderful} \end{array} \right\} \textit{little} \left\{ \begin{array}{c} \textit{city} \\ \textit{town} \end{array} \right\}$$

Instead of enlisting $M = \{$*I don't like ...*$\}$:

- We assume that Alice and Bob will exchange arbitrary bitstrings, so $M = \{0, 1\}^*$

Alice and Bob Fool Wendy

$$\textit{Midshires is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \\ \texttt{01} & \textit{great} \\ \mathbf{\texttt{10}} & \textbf{\textit{fine}} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \right\} \textit{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \mathbf{\texttt{1}} & \textbf{\textit{town}} \end{array} \right\}.$$

- $m_1 = \texttt{101}$ encodes to
  $c_1 =$ *Midshires is a fine little town*
- $m_2 = \texttt{010}$ encodes to
  $c_2 =$ *Midshires is a great little city*
- ...

Alice and Bob Fool Wendy

$$\textit{Midshires is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \\ \texttt{01} & \textit{great} \\ \mathbf{10} & \textbf{\textit{fine}} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \right\} \textit{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \mathbf{1} & \textbf{\textit{town}} \end{array} \right\}.$$

Keith Winstein and Mark Chapman have actually built variants of this system.

Wendy Strikes Back

Statistic characteristics of the secret message “show trough” to the word-choices in the steganogram!

Wendy Strikes Back

$$\text{is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \\ \texttt{01} & \textit{great} \\ \texttt{10} & \textit{fine} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \right\} \text{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \texttt{1} & \textit{town} \end{array} \right\}.$$

- $m_0 = \texttt{101}$
- $m_1 = \texttt{001}$
- $m_2 = \texttt{111}$
- ...

Wendy Strikes Back

$$\text{is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \\ \texttt{01} & \textit{great} \\ \texttt{10} & \textit{fine} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \quad | \quad \right\} \text{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \texttt{1} & \textit{town} \quad | \end{array} \right\}.$$

- $m_1 = \texttt{001}$
- $m_2 = \texttt{111}$
- $m_3 = \texttt{101}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & | \\ \texttt{01} & \textit{great} & \\ \texttt{10} & \textit{fine} & | \\ \texttt{11} & \textit{decent} & \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & \\ \texttt{1} & \textit{town} & \| \end{array} \right\}.$$

- $m_2 = \texttt{111}$
- $m_3 = \texttt{101}$
- $m_4 = \texttt{000}$
- ...

Wendy Strikes Back

is a { `00` *nice* | `01` *great* `10` *fine* | `11` *decent* | `??` *wonderful* } *little* { `0` *city* `1` *town* ||| }.

- $m_3 = \texttt{101}$
- $m_4 = \texttt{000}$
- $m_5 = \texttt{010}$
- ...

Wendy Strikes Back

$$\textit{is a}\left\{\begin{array}{lll} \texttt{00} & \textit{nice} & | \\ \texttt{01} & \textit{great} & \\ \texttt{10} & \textit{fine} & || \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array}\right\} \textit{little}\left\{\begin{array}{lll} \texttt{0} & \textit{city} & \\ \texttt{1} & \textit{town} & |||| \end{array}\right\}.$$

- $m_4 = 000$
- $m_5 = 010$
- $m_6 = 000$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & \| \\ \texttt{01} & \textit{great} & \\ \texttt{10} & \textit{fine} & \| \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & | \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_5 = \texttt{010}$
- $m_6 = \texttt{000}$
- $m_7 = \texttt{010}$
- ...

Wendy Strikes Back

$$\text{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & || \\ \texttt{01} & \textit{great} & | \\ \texttt{10} & \textit{fine} & || \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \text{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & || \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_6 = \texttt{000}$
- $m_7 = \texttt{010}$
- $m_8 = \texttt{100}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & | \\ \texttt{10} & \textit{fine} & || \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_7 = \texttt{010}$
- $m_8 = \texttt{100}$
- $m_9 = \texttt{110}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & || \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & |||| \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_8 = \texttt{100}$
- $m_9 = \texttt{110}$
- $m_{10} = \texttt{111}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & ||| \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||| \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_9 = \texttt{110}$
- $m_{10} = \texttt{111}$
- $m_{11} = \texttt{100}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & ||| \\ \texttt{11} & \textit{decent} & || \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & |||||| \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $m_{10} = \texttt{111}$
- $m_{11} = \texttt{100}$
- $m_{12} = \texttt{111}$
- ...

Wendy Strikes Back

$$\text{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & ||| \\ \texttt{11} & \textit{decent} & ||| \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \text{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & |||||| \\ \texttt{1} & \textit{town} & ||||| \end{array} \right\}.$$

- $m_{11} = \texttt{100}$
- $m_{12} = \texttt{111}$
- $m_{13} = \texttt{011}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & |||| \\ \texttt{11} & \textit{decent} & ||| \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||||| \\ \texttt{1} & \textit{town} & ||||| \end{array} \right\}.$$

- $m_{12} = \texttt{111}$
- $m_{13} = \texttt{011}$
- $m_{14} = \texttt{011}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & |||| \\ \texttt{11} & \textit{decent} & |||| \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||||| \\ \texttt{1} & \textit{town} & |||||| \end{array} \right\}.$$

- $m_{13} = \texttt{011}$
- $m_{14} = \texttt{011}$
- $m_{15} = \texttt{000}$
- ...

Wendy Strikes Back

is a
$$\left\{\begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & ||| \\ \texttt{10} & \textit{fine} & |||| \\ \texttt{11} & \textit{decent} & |||| \\ \texttt{??} & \textit{wonderful} & \end{array}\right\}$$
little
$$\left\{\begin{array}{lll} \texttt{0} & \textit{city} & ||||||| \\ \texttt{1} & \textit{town} & ||||||| \end{array}\right\}.$$

- $m_{14} = \texttt{011}$
- $m_{15} = \texttt{000}$
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & |||| \\ \texttt{10} & \textit{fine} & |||| \\ \texttt{11} & \textit{decent} & |||| \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||||| \\ \texttt{1} & \textit{town} & |||||||| \end{array} \right\}.$$

- $m_{15} = \texttt{000}$
- ...

Wendy Strikes Back

$$\textit{is a}\left\{\begin{array}{lll} \texttt{00} & \textit{nice} & |||| \\ \texttt{01} & \textit{great} & |||| \\ \texttt{10} & \textit{fine} & |||| \\ \texttt{11} & \textit{decent} & |||| \\ \texttt{??} & \textit{wonderful} & \end{array}\right\}\textit{little}\left\{\begin{array}{lll} \texttt{0} & \textit{city} & |||||||| \\ \texttt{1} & \textit{town} & |||||||| \end{array}\right\}.$$

- ...

Wendy Strikes Back

Innocuous covers have statistic characteristics originating from the way native speakers use the language.

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \\ \texttt{01} & \textit{great} \\ \texttt{10} & \textit{fine} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \right\} \textit{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \texttt{1} & \textit{town} \end{array} \right\}.$$

- $c_0 =$ *Midshires is a nice little town*
- $c_1 =$ *Midshires is a nice little city*
- $c_2 =$ *Midshires is a nice little town*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{ll} \texttt{00} & \textit{nice} \quad | \\ \texttt{01} & \textit{great} \\ \texttt{10} & \textit{fine} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful} \end{array} \right\} \textit{little} \left\{ \begin{array}{ll} \texttt{0} & \textit{city} \\ \texttt{1} & \textit{town} \quad | \end{array} \right\}.$$

- $c_1 =$ *Midshires is a nice little city*
- $c_2 =$ *Midshires is a nice little town*
- $c_3 =$ *Midshires is a nice little town*
- ...

Wendy Strikes Back

$$\textit{is a}\left\{\begin{array}{ll}\texttt{00} & \textit{nice} \quad || \\ \texttt{01} & \textit{great} \\ \texttt{10} & \textit{fine} \\ \texttt{11} & \textit{decent} \\ \texttt{??} & \textit{wonderful}\end{array}\right\}\textit{little}\left\{\begin{array}{ll}\texttt{0} & \textit{city} \quad | \\ \texttt{1} & \textit{town} \quad |\end{array}\right\}.$$

- $c_2 =$ *Midshires is a nice little town*
- $c_3 =$ *Midshires is a nice little town*
- $c_4 =$ *Midshires is a nice little city*
- ...

Wendy Strikes Back

$$\text{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||| \\ \texttt{01} & \textit{great} & \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \text{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & | \\ \texttt{1} & \textit{town} & || \end{array} \right\}.$$

- $c_3 =$ *Midshires is a nice little town*
- $c_4 =$ *Midshires is a nice little city*
- $c_5 =$ *Midshires is a great little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & |||| \\ \texttt{01} & \textit{great} & \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & | \\ \texttt{1} & \textit{town} & ||| \end{array} \right\}.$$

- $c_4 =$ *Midshires is a nice little city*
- $c_5 =$ *Midshires is a great little city*
- $c_6 =$ *Midshires is a nice little town*
- ...

Wendy Strikes Back

is a { `00` *nice* ||||| ; `01` *great* ; `10` *fine* ; `11` *decent* ; `??` *wonderful* } little { `0` *city* || ; `1` *town* ||| }.

- $c_5 =$ *Midshires is a great little city*
- $c_6 =$ *Midshires is a nice little town*
- $c_7 =$ *Midshires is a decent little town*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||||| \\ \texttt{01} & \textit{great} & | \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & ||| \end{array} \right\}.$$

- $c_6 =$ *Midshires is a nice little town*
- $c_7 =$ *Midshires is a decent little town*
- $c_8 =$ *Midshires is a great little town*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & |||||| \\ \texttt{01} & \textit{great} & | \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & |||| \end{array} \right\}.$$

- $c_7 =$ *Midshires is a decent little town*
- $c_8 =$ *Midshires is a great little town*
- $c_9 =$ *Midshires is a nice little town*
- ...

Wendy Strikes Back

$$\textit{is a}\left\{\begin{array}{lll} \texttt{00} & \textit{nice} & |||||| \\ \texttt{01} & \textit{great} & | \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array}\right\} \textit{little} \left\{\begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & ||||| \end{array}\right\}.$$

- $c_8 =$ *Midshires is a great little town*
- $c_9 =$ *Midshires is a nice little town*
- $c_{10} =$ *Midshires is a wonderful little town*
- ...

Wendy Strikes Back

$$\textit{is a}\left\{\begin{array}{lll} \texttt{00} & \textit{nice} & |||||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array}\right\}\textit{little}\left\{\begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & |||||| \end{array}\right\}.$$

- $c_9 =$ *Midshires is a nice little town*
- $c_{10} =$ *Midshires is a wonderful little town*
- $c_{11} =$ *Midshires is a great little town*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||||||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & ||||||| \end{array} \right\}.$$

- $c_{10} =$ *Midshires is a wonderful little town*
- $c_{11} =$ *Midshires is a great little town*
- $c_{12} =$ *Midshires is a great little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||||||| \\ \texttt{01} & \textit{great} & || \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & | \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & |||||||| \end{array} \right\}.$$

- $c_{11} =$ *Midshires is a great little town*
- $c_{12} =$ *Midshires is a great little city*
- $c_{13} =$ *Midshires is a fine little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||||||| \\ \texttt{01} & \textit{great} & ||| \\ \texttt{10} & \textit{fine} & \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & | \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||| \\ \texttt{1} & \textit{town} & ||||||||| \end{array} \right\}.$$

- $c_{12} =$ *Midshires is a great little city*
- $c_{13} =$ *Midshires is a fine little city*
- $c_{14} =$ *Midshires is a nice little city*
- ...

Wendy Strikes Back

is a

`00`	*nice*	|||||||
`01`	*great*	||||
`10`	*fine*	
`11`	*decent*	|
`??`	*wonderful*	|

little

`0`	*city*	||||
`1`	*town*	|||||||||

.

- $c_{13} =$ *Midshires is a fine little city*
- $c_{14} =$ *Midshires is a nice little city*
- $c_{15} =$ *Midshires is a fine little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & ||||||| \\ \texttt{01} & \textit{great} & |||| \\ \texttt{10} & \textit{fine} & | \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & | \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||| \\ \texttt{1} & \textit{town} & ||||||||| \end{array} \right\}.$$

- c_{14} = *Midshires is a nice little city*
- c_{15} = *Midshires is a fine little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & |||||||| \\ \texttt{01} & \textit{great} & |||| \\ \texttt{10} & \textit{fine} & | \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & | \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & |||||| \\ \texttt{1} & \textit{town} & ||||||||| \end{array} \right\} .$$

- $c_{15} =$ *Midshires is a fine little city*
- ...

Wendy Strikes Back

$$\textit{is a} \left\{ \begin{array}{lll} \texttt{00} & \textit{nice} & |||||||| \\ \texttt{01} & \textit{great} & |||| \\ \texttt{10} & \textit{fine} & || \\ \texttt{11} & \textit{decent} & | \\ \texttt{??} & \textit{wonderful} & | \end{array} \right\} \textit{little} \left\{ \begin{array}{lll} \texttt{0} & \textit{city} & ||||||| \\ \texttt{1} & \textit{town} & ||||||||| \end{array} \right\}.$$

- ...

Wendy Strikes Back

is a {
00 *nice* ||||
01 *great* ||||
10 *fine* ||||
11 *decent* ||||
?? *wonderful*
} *little* {
0 *city* ||||||||
1 *town* ||||||||
} .

is a {
00 *nice* ||||||||
01 *great* ||||
10 *fine* ||
11 *decent* |
?? *wonderful* |
} *little* {
0 *city* |||||||
1 *town* |||||||||
} .

This weakness is due to
our use of block codes!

Alice, Bob, and Huffman

00101110110010111001011100010100

00 |

01

10

11

Alice, Bob, and Huffman

00**10**111011001011100101110001 0100

00 |

01

10 |

11

Alice, Bob, and Huffman

0010**11**101100101110010111000010100

```
00 |
01
10 |
11 |
```

Alice, Bob, and Huffman

001011**10**1100101110010111000010100

```
00  |
01
10  ||
11  |
```

Alice, Bob, and Huffman

00101110**11**0010111001011100010100

00 |
01
10 ||
11 ||

Alice, Bob, and Huffman

0010111011**00**10111001011100010100

00 ||
01
10 ||
11 ||

Alice, Bob, and Huffman

001011101100**10**1110010111000010100

00	||
01	
10	|||
11	||

Alice, Bob, and Huffman

00101110110010**11**1001011100010100

00 ||
01
10 |||
11 |||

Alice, Bob, and Huffman

0010111011001011**10**0101110001010

00	||
01	
10	||||
11	|||

Alice, Bob, and Huffman

001011101100101110**01**01110001010 0

00	||
01	|
10	||||
11	|||

Alice, Bob, and Huffman

0010111011001011100**01**1100010100

00	||
01	||
10	||||
11	|||

Alice, Bob, and Huffman

0010111011001011100101**11**00010100

00	||
01	||
10	||||
11	||||

Alice, Bob, and Huffman

001011101100101110010111**00**010100

00 |||
01 ||
10 ||||
11 ||||

Alice, Bob, and Huffman

00101110110010111001011100**01**0100

00	|||
01	|||
10	||||
11	||||

Alice, Bob, and Huffman

0010111011001011100101110001**01**00

00 |||

01 ||||

10 ||||

11 ||||

Alice, Bob, and Huffman

001011101100101110010111000101**00**

```
00  ||||
01  ||||
10  ||||
11  ||||
```

Alice, Bob, and Huffman

`0010111011001011100101110001 0100`

00	||||	$p = 1/4$	(**00**, 01, 10, 11)
01	||||	$p = 1/4$	(00, **01**, 10, 11)
10	||||	$p = 1/4$	(00, 01, **10**, 11)
11	||||	$p = 1/4$	(00, 01, 10, **11**)

Alice, Bob, and Huffman

This weakness is due to
our use of block codes!

However, we can overcome it, by using
prefix-free variable length codes.

Alice, Bob, and Huffman

001000110001001110111101011010

```
   0 |
  10
 110
1110
1111
```

Alice, Bob, and Huffman

0**0**10001100010011101111010110 10

0 ||

10

110

1110

1111

Alice, Bob, and Huffman

00**10**0011000100111011110101 1010

```
   0  ||
  10  |
 110
1110
1111
```

Alice, Bob, and Huffman

0010**0**0110001001110111101011010

0 |||

10 |

110

1110

1111

Alice, Bob, and Huffman

00100**0**110001001110111101011010

0	||||
10	|
110	
1110	
1111	

Alice, Bob, and Huffman

`001000`**`110`**`0010011101111010 11010`

0	||||
10	|
110	|
1110	
1111	

Alice, Bob, and Huffman

001000110**0**01001110111101011010

```
   0  |||||
  10  |
 110  |
1110
1111
```

Alice, Bob, and Huffman

0010001100**0**1001110111101011010

0	||||||
10	|
110	|
1110	
1111	

Alice, Bob, and Huffman

00100011000**10**01110111101011010

```
   0  ||||||
  10  ||
 110  |
1110
1111
```

Alice, Bob, and Huffman

0 0 1 0 0 0 1 1 0 0 0 1 0 **0** 1 1 1 0 1 1 1 1 0 1 0 1 1 0 1 0

0								
10			
110		
1110	
1111	

Alice, Bob, and Huffman

`0010001100010`**`1110`**`1111010110 10`

0	|||||||
10	||
110	|
1110	|
1111	

Alice, Bob, and Huffman

0010001100010011101**111**01011010

0	|||||||
10	||
110	|
1110	|
1111	|

Alice, Bob, and Huffman

0 0 1 0 0 0 1 1 0 0 0 1 0 0 1 1 1 0 1 1 1 1 **0** 1 0 1 1 0 1 0

0	||||||||
10	||
110	|
1110	|
1111	|

Alice, Bob, and Huffman

0010001100010011101111 0**10**11010

0 ||||||||

10 |||

110 |

1110 |

1111 |

Alice, Bob, and Huffman

0010001100010011101111010**110**10

0	||||||||
10	|||
110	||
1110	|
1111	|

Alice, Bob, and Huffman

0010001100010011101111010110**10**

0	||||||||
10	||||
110	||
1110	|
1111	|

Alice, Bob, and Huffman

0010001100010011101111010110**10**

0	||||||||	$p = 1/2$ (**0**, 1)
10	||||	$p = 1/4$ (00, 01, **10**, 11)
110	||	$p = 1/8$ (000, 001, ..., **110**, ...)
1110	|	$p = 1/16$ (0000, 0001, ..., **1110**, ...)
1111	|	$p = 1/16$ (0000, 0001, ..., **1111**, ...)

Alice, Bob, and Huffman

`0010001100010011101111010110 10`

0	||||||||	$p = 1/2$ $(\mathbf{0}, 1)$
10	||||	$p = 1/4$ $(00, 01, \mathbf{10}, 11)$
110	||	$p = 1/8$ $(000, 001, ..., \mathbf{110}, ...)$
1110	|	$p = 1/16$ $(0000, 0001, ..., \mathbf{1110}, ...)$
1111	|	$p = 1/16$ $(0000, 0001, ..., \mathbf{1111}, ...)$

The use of prefix free variable length codes in steganography is due to Peter Wayner!

Wendy's Cryptographic Problem

- By word-choice encoding on the basis of a Huffman-code we can provide **mimicry**.

Wendy's Cryptographic Problem

- By word-choice encoding on the basis of a Huffman-code we can provide **mimicry**.
- **Mimicry** turns a secret message $m \in M$ to an innocuous looking cover $c \in C$.

Wendy's Cryptographic Problem

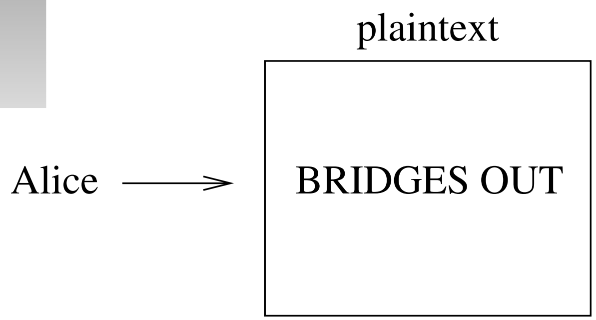
- By word-choice encoding on the basis of a Huffman-code we can provide **mimicry**.
- **Mimicry** turns a secret message $m \in M$ to an innocuous looking cover $c \in C$.
- To know whether our scheme is steganographically secure, we have to ask ourselves the following question:

Wendy's Cryptographic Problem

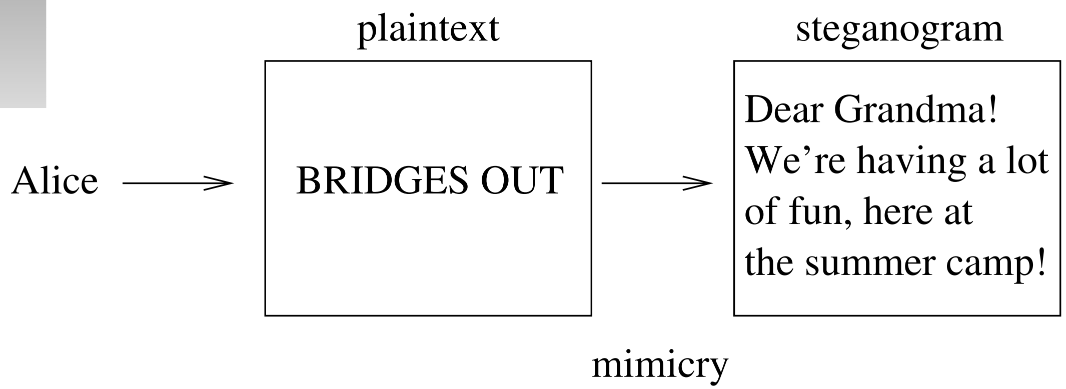
- By word-choice encoding on the basis of a Huffman-code we can provide **mimicry**.
- **Mimicry** turns a secret message $m \in M$ to an innocuous looking cover $c \in C$.
- To know whether our scheme is steganographically secure, we have to ask ourselves the following question:

If it is trivial for Bob to decode a message, then why shouldn't Wendy do the very same thing?

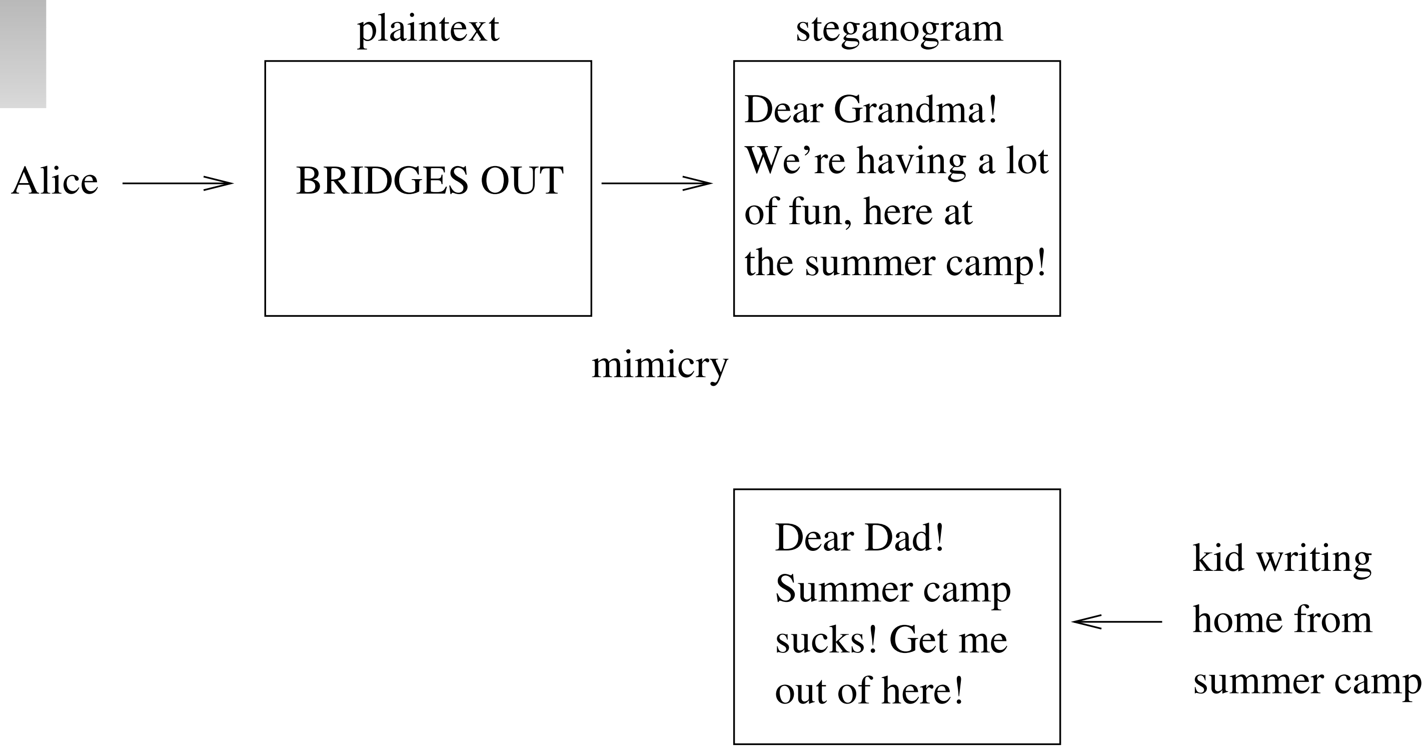
Wendy's Cryptographic Problem



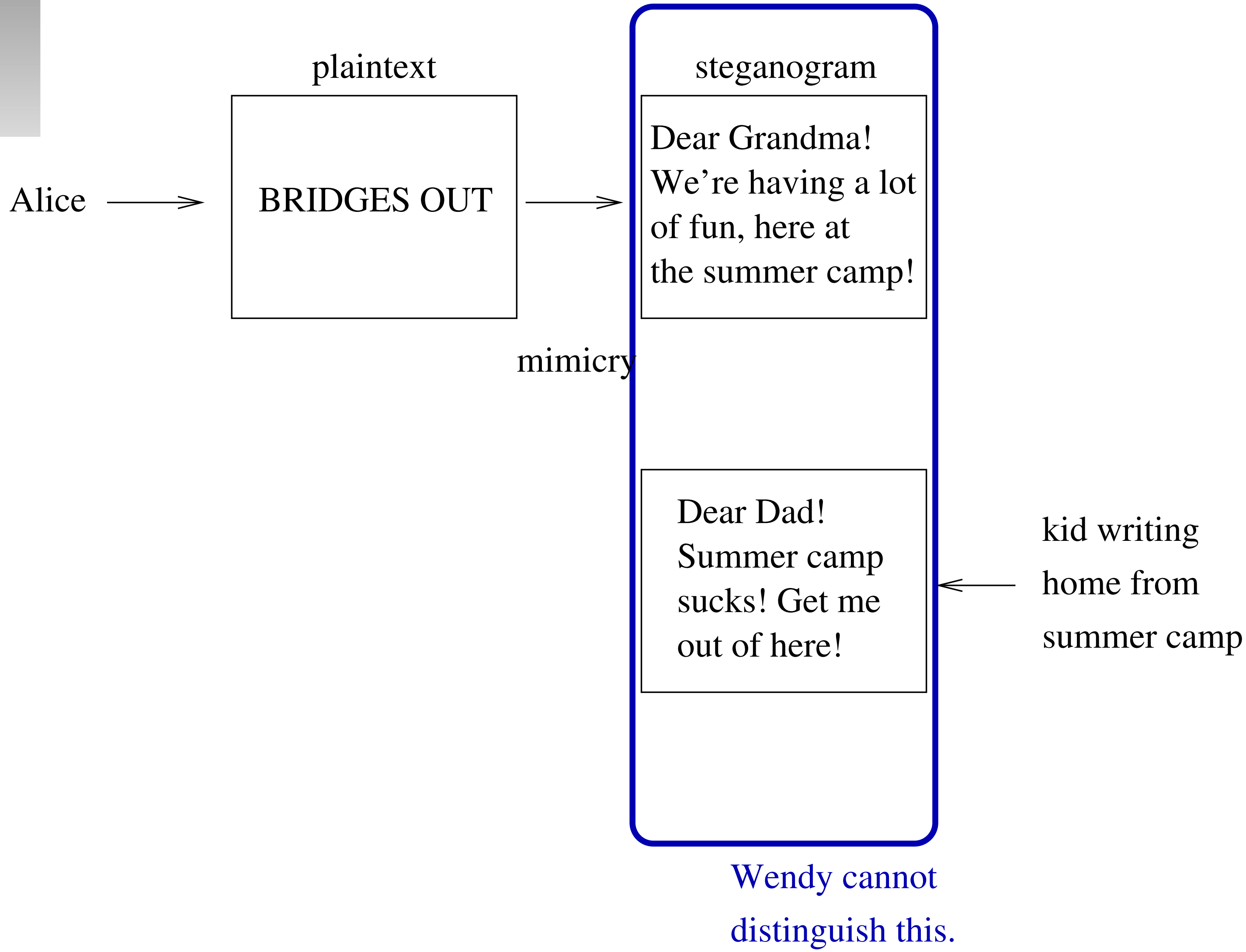
Wendy's Cryptographic Problem



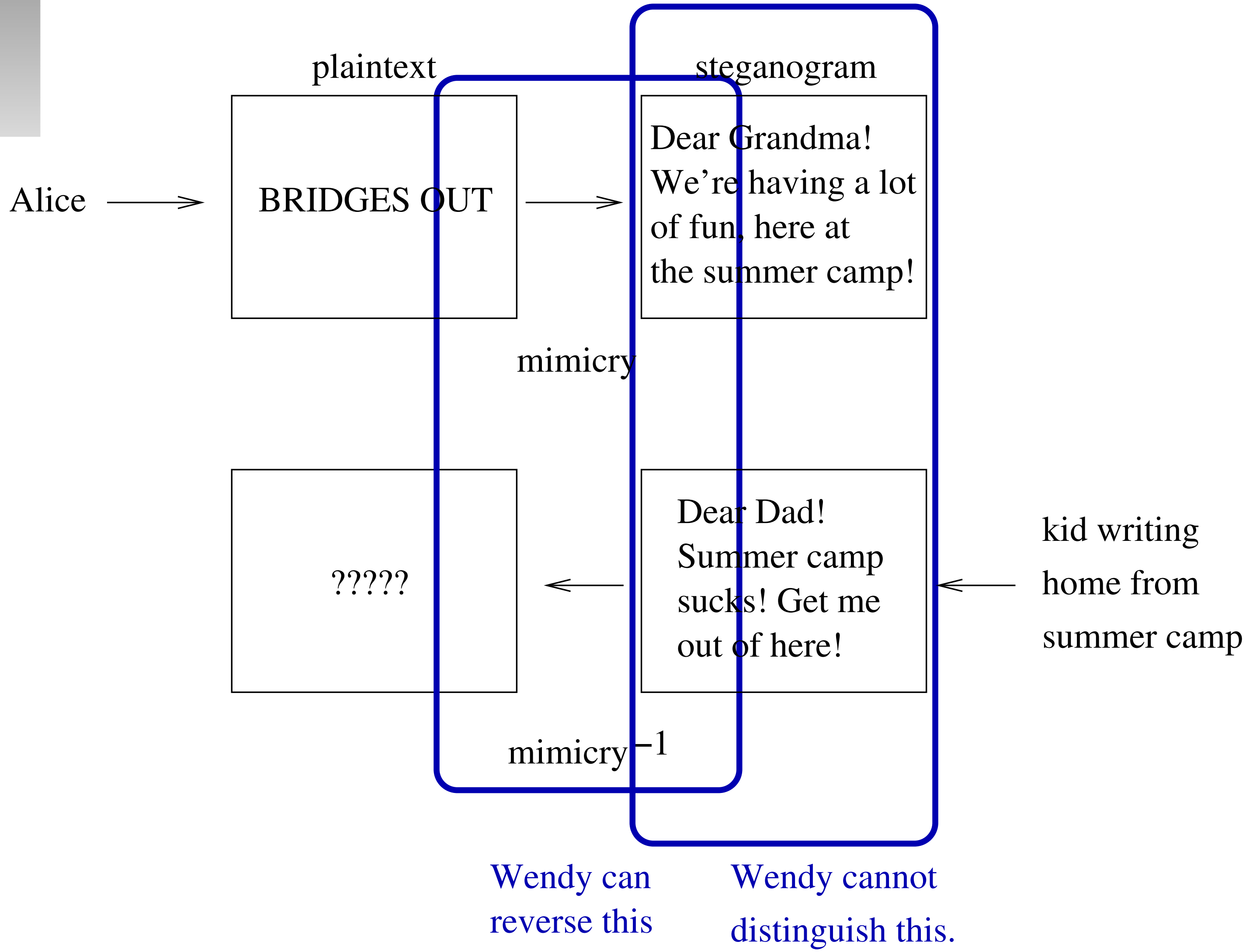
Wendy's Cryptographic Problem



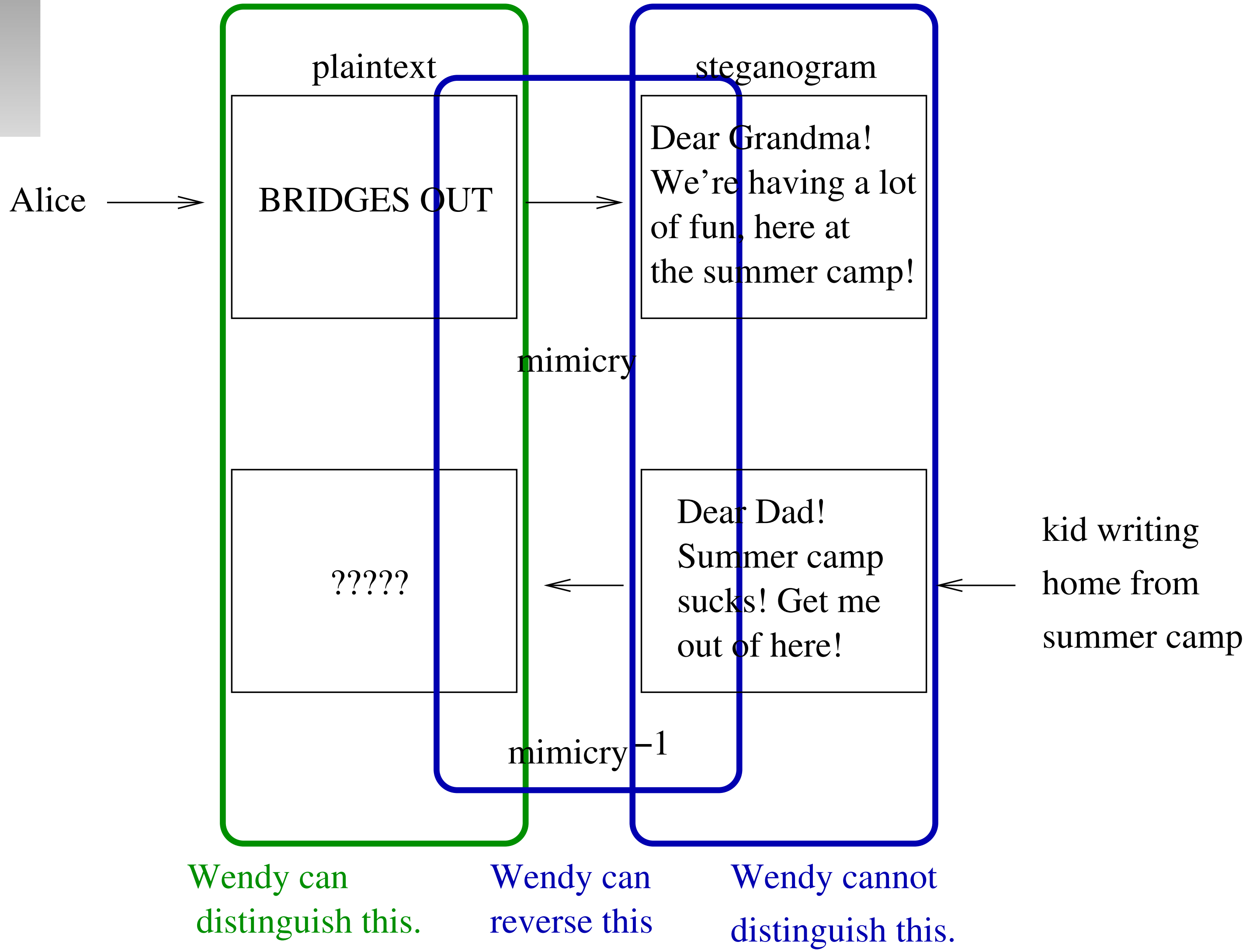
Wendy's Cryptographic Problem



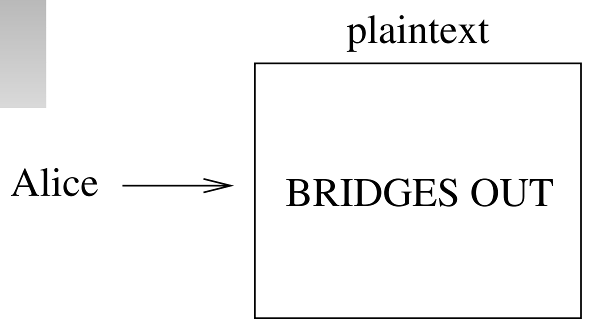
Wendy's Cryptographic Problem



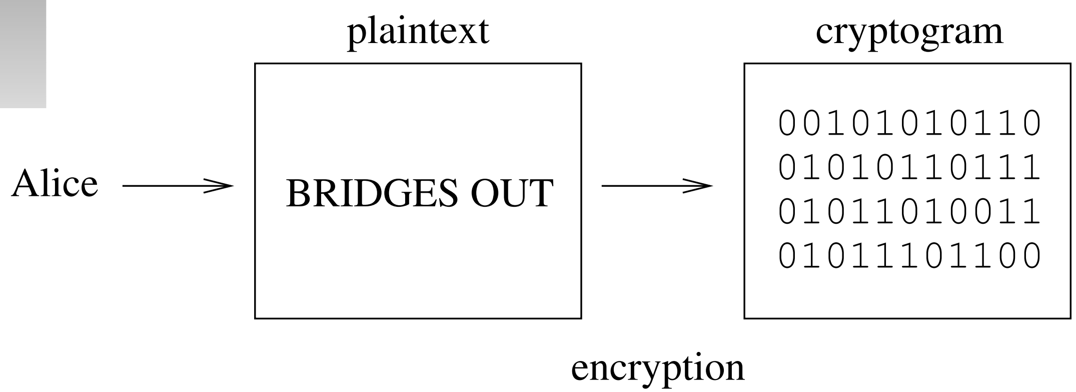
Wendy's Cryptographic Problem



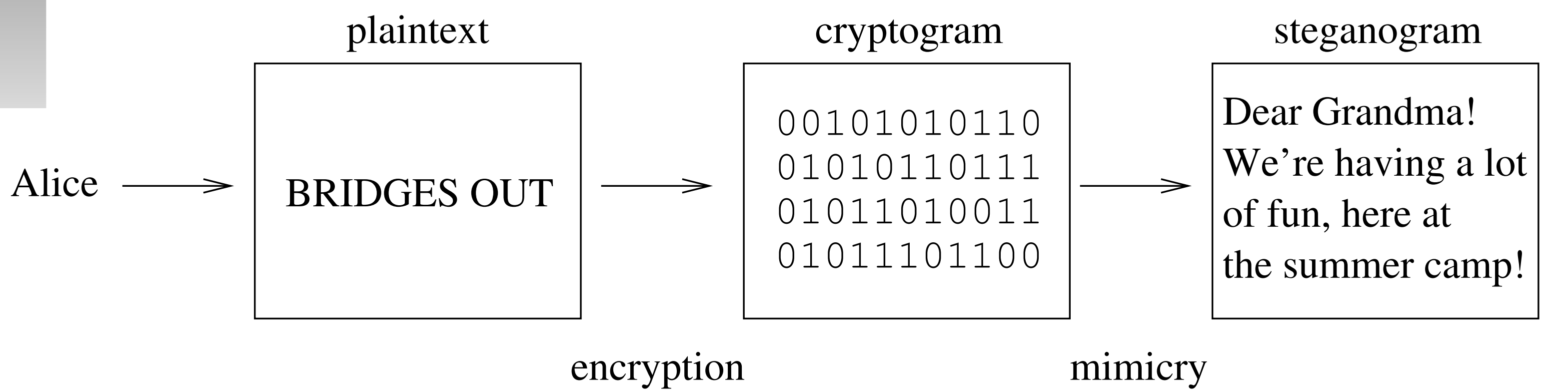
Wendy's Cryptographic Problem



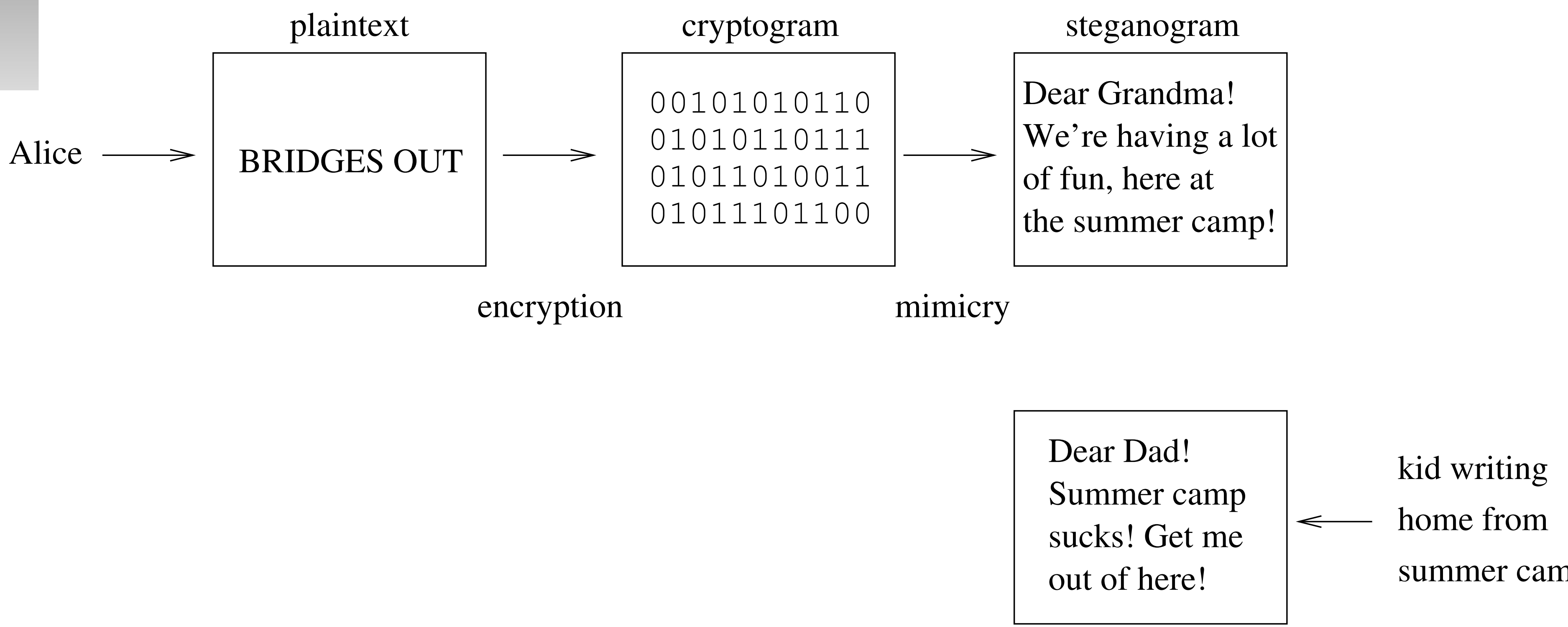
Wendy's Cryptographic Problem



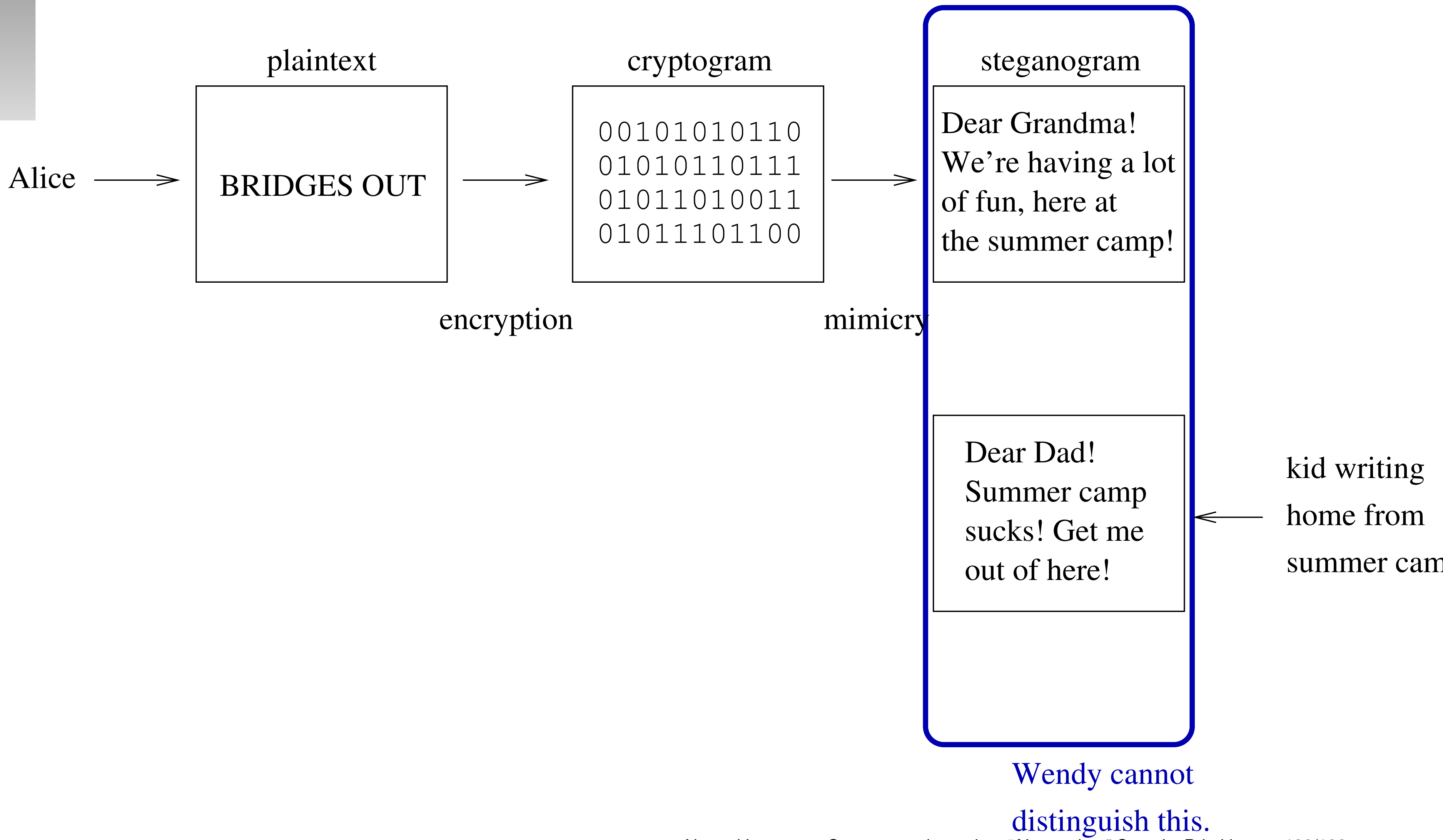
Wendy's Cryptographic Problem



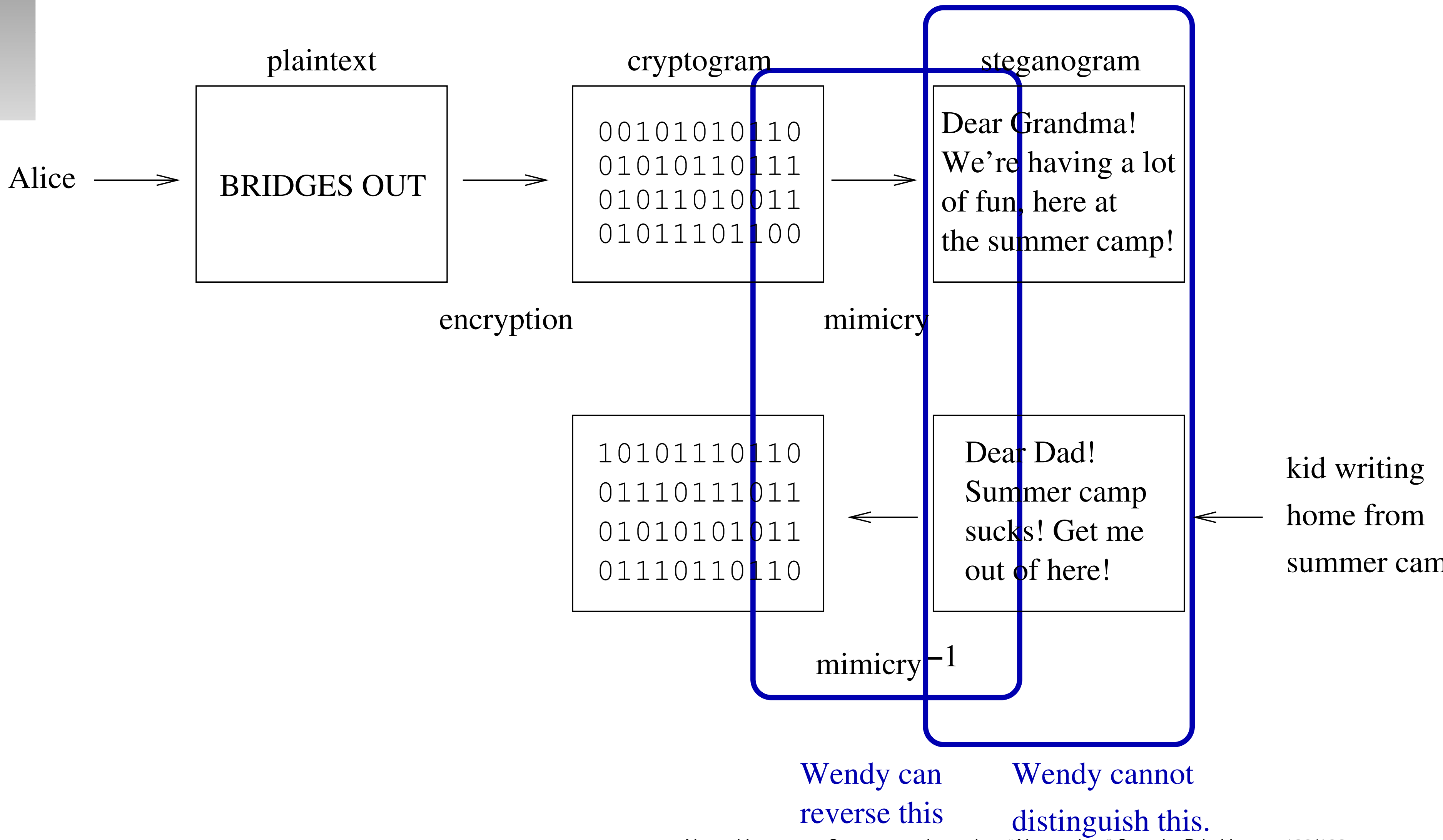
Wendy's Cryptographic Problem



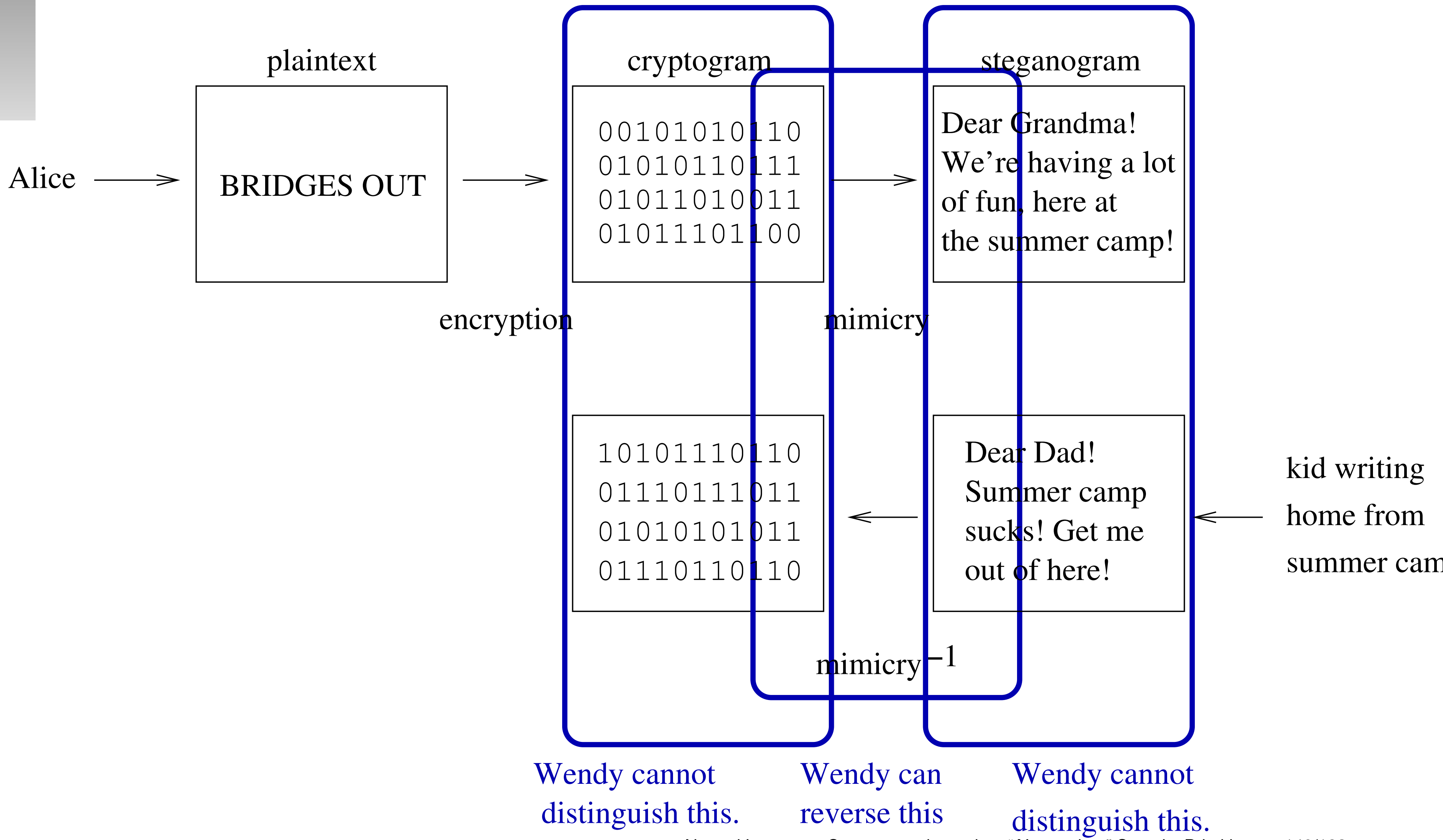
Wendy's Cryptographic Problem



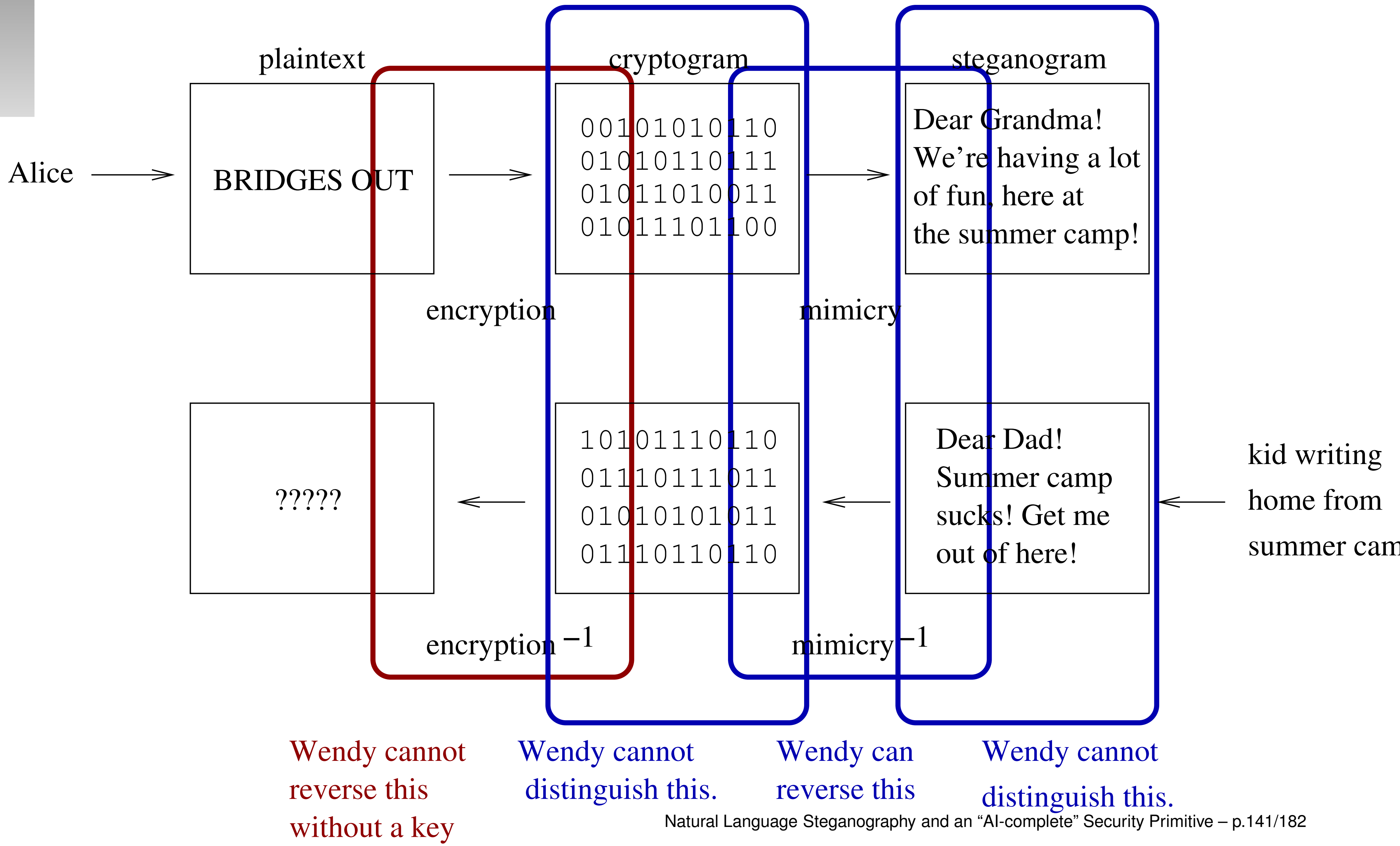
Wendy's Cryptographic Problem



Wendy's Cryptographic Problem



Wendy's Cryptographic Problem



Wendy's Linguistic Problem

And now to something completely different!

And now to something completely different!

AI-complete */A-I k*m-pleet'/* [MIT, Stanford: by analogy with 'NP-complete' *(see NP-)*] adj. Used to describe problems or subproblems in AI, to indicate that the solution presupposes a solution to the 'strong AI problem' (that is, the synthesis of a human-level intelligence). A problem that is AI-complete is, in other words, just too hard[...]

The Jargon Files

Wendy's Linguistic Problem

- Humans can easily solve AI-complete problems.

Wendy's Linguistic Problem

- Humans can easily solve AI-complete problems.
- Computers cannot.

Wendy's Linguistic Problem

- Humans can easily solve AI-complete problems.
- Computers cannot.
- Wendy is a computer.

Wendy's Linguistic Problem

- Humans can easily solve AI-complete problems.
- Computers cannot.
- Wendy is a computer.
- If reversing our steganographic embedding yields an AI-complete problem, Wendy is truly in trouble.

Wendy's Linguistic Problem

- Humans can easily solve AI-complete problems.
- Computers cannot.
- Wendy is a computer.
- If reversing our steganographic embedding yields an AI-complete problem, Wendy is truly in trouble.
- We can construct such a system, by using the linguistic problem of **word-sense disambiguation**.

Wendy's Linguistic Problem

- *It should **move** through several more drafts.*
- *It should **run** through several more drafts.*
- *It should **go** through several more drafts.*

Wendy's Linguistic Problem

- *It should* ***move*** *through several more drafts.*
- *It should* ***run*** *through several more drafts.*
- *It should* ***go*** *through several more drafts.*

- *All articles must* ***move*** *through copy-editing.*
- *All articles must* ***run*** *through copy-editing.*
- *All articles must* ***go*** *through copy-editing.*

Wendy's Linguistic Problem

- *It should **move** through several more drafts.*
- *It should **run** through several more drafts.*
- *It should **go** through several more drafts.*

- *All articles must **move** through copy-editing.*
- *All articles must **run** through copy-editing.*
- *All articles must **go** through copy-editing.*

$$\text{syn}(\text{move}) = \{\text{move}, \text{run}, \text{go}\} \quad ??$$

Wendy's Linguistic Problem

- *That sermon will **move** people.*
- *That sermon will **impress** people.*
- *That sermon will **strike** people.*

Wendy's Linguistic Problem

- *That sermon will **move** people.*
- *That sermon will **impress** people.*
- *That sermon will **strike** people.*

- *Your speech must **move** the audience.*
- *Your speech must **impress** the audience.*
- *Your speech must **strike** the audience.*

Wendy's Linguistic Problem

- *That sermon will **move** people.*
- *That sermon will **impress** people.*
- *That sermon will **strike** people.*

- *Your speech must **move** the audience.*
- *Your speech must **impress** the audience.*
- *Your speech must **strike** the audience.*

$$\text{syn}(\text{move}) = \{\text{move}, \text{impress}, \text{strike}\} \quad ??$$

Wendy's Linguistic Problem

Can we conclude that all these words are **generally** synonymous to move?

$$\text{syn}(\text{move}) = \{\text{move}, \text{run}, \text{go}, \text{impress}, \text{strike}\}$$

Wendy's Linguistic Problem

Can we conclude that all these words are **generally** synonymous to move?

$$\text{syn}(\text{move}) = \{\text{move}, \text{run}, \text{go}, \text{impress}, \text{strike}\}$$

Unfortunately, we can't.

Wendy’s Linguistic Problem

- *It should **move** through several more drafts.*
- *It should **run** through several more drafts.*
- *It should **go** through several more drafts.*

Wendy's Linguistic Problem

- *It should **move** through several more drafts.*
- *It should **run** through several more drafts.*
- *It should **go** through several more drafts.*

BUT

- *Your speech must **move** the audience.*
- **Your speech must **run** the audience.*
- **Your speech must **go** the audience.*

Wendy's Linguistic Problem

- *That sermon will **move** people.*
- *That sermon will **impress** people.*
- *That sermon will **strike** people.*

Wendy's Linguistic Problem

- *That sermon will **move** people.*
- *That sermon will **impress** people.*
- *That sermon will **strike** people.*

BUT

- *All articles must **move** through copy-editing.*
- **All articles must **impress** through copy-editing.*
- **All articles must **strike** through copy-editing.*

Wendy's Linguistic Problem

We cannot include a synset like

$$\text{syn}(\text{move}) = \{\text{move}, \text{run}, \text{go}, \text{impress}, \text{strike}\}$$

in a dictionary!

All we can do is to state that

$$\begin{aligned} \text{syn}(c_1, \text{move}) &= \{\text{move}, \text{run}, \text{go}\} \\ \text{syn}(c_2, \text{move}) &= \{\text{move}, \text{impress}, \text{strike}\} \end{aligned}$$

for some linguistic contexts $c_1 \neq c_2$.

Wendy's Linguistic Problem

Recall the way our encoder works:

- We have an innocuous sentence:
  That sermon will impress people

Wendy's Linguistic Problem

Recall the way our encoder works:

- We have an innocuous sentence:
  That sermon will impress people
- We have a set of words that can be replaced for this {*move*, *impress*, *strike*}

Wendy's Linguistic Problem

Recall the way our encoder works:

- We have an innocuous sentence:
  That sermon will impress people
- We have a set of words that can be replaced for this $\{$*move*, *impress*, *strike*$\}$
- We assign codewords to them like
  move $\rightarrow$ `0`, *impress* $\rightarrow$ `10`, *strike* $\rightarrow$ `11`

Wendy's Linguistic Problem

Recall the way our encoder works:

- We have an innocuous sentence:
  That sermon will impress people
- We have a set of words that can be replaced for this $\{$*move*, *impress*, *strike*$\}$
- We assign codewords to them like
  move $\rightarrow$ `0`, *impress* $\rightarrow$ `10`, *strike* $\rightarrow$ `11`
- To send a secret `0`, we transmit
  That sermon will move people

Wendy's Linguistic Problem

How would Wendy steganalyze this?

- She intercepts
  That sermon will move people

Wendy's Linguistic Problem

How would Wendy steganalyze this?

- She intercepts
  That sermon will move people
- Now she has to find the code for *move*.

Wendy's Linguistic Problem

How would Wendy steganalyze this?

- She intercepts
  That sermon will move people
- Now she has to find the code for *move*.
- However, there will be multiple codes for this:
  - *move* $\rightarrow$ `0`, *impress* $\rightarrow$ `10`, *strike* $\rightarrow$ `11` (correct)
  - *run* $\rightarrow$ `0`, *move* $\rightarrow$ `10`, *go* $\rightarrow$ `11` (incorrect)

Wendy's Linguistic Problem

How would Wendy steganalyze this?

- She intercepts
  That sermon will move people
- Now she has to find the code for *move*.
- However, there will be multiple codes for this:
  - *move* $\rightarrow$ `0`, *impress* $\rightarrow$ `10`, *strike* $\rightarrow$ `11` (correct)
  - *run* $\rightarrow$ `0`, *move* $\rightarrow$ `10`, *go* $\rightarrow$ `11` (incorrect)
- In order to decode this replacement, Wendy has to solve an instance of the AI-complete problem of word-sense ambiguity!

Concluding Remarks

- We showed that steganography can be motivated by the application of hacker ethics to cryptographic system design.

Concluding Remarks

- We showed that steganography can be motivated by the application of hacker ethics to cryptographic system design.
- We showed a simple technique to hide data by replacing words in innocuous text by synonyms.

Concluding Remarks

- We showed that steganography can be motivated by the application of hacker ethics to cryptographic system design.
- We showed a simple technique to hide data by replacing words in innocuous text by synonyms.
- We showed how to detect such steganograms using statistic patterns.

Concluding Remarks

- We showed how to improve the technique, so that detection becomes more difficult.

Concluding Remarks

- We showed how to improve the technique, so that detection becomes more difficult.
- We showed an approach towards making the technique secure against arbitrators who do not have a certain key.

Concluding Remarks

- We showed how to improve the technique, so that detection becomes more difficult.
- We showed an approach towards making the technique secure against arbitrators who do not have a certain key.
- We showed an approach towards making the technique secure against arbitrators who are not human.

Concluding Remarks

- We did not review the actual systems implemented so far.

Concluding Remarks

- We did not review the actual systems implemented so far.
- We did not review much related literature.

Concluding Remarks

- We did not review the actual systems implemented so far.
- We did not review much related literature.
- We did not show how to make the technique robust.

Concluding Remarks

- We did not review the actual systems implemented so far.
- We did not review much related literature.
- We did not show how to make the technique robust.
- We did not show how to use our linguistic properties for simple human interactive proofs.

Concluding Remarks

- We did not review the actual systems implemented so far.
- We did not review much related literature.
- We did not show how to make the technique robust.
- We did not show how to use our linguistic properties for simple human interactive proofs.

Please see the provided references for these things!

This slide-set is not a self-contained publication.
Please conduct the references instead.

In particular, note that sources were not properly cited in this slide-set. See the citations given in the project-report for reference on sources.

References

Natural Language Steganography
and an “AI-complete” Security Primitive

for reference, see:

- Richard Bergmair. Towards linguistic steganography: A systematic investigation of approaches, systems, and issues. April 2004.
- Richard Bergmair and Stefan Katzenbeisser. Towards human interactive proofs in the text-domain. September 2004.

Available online: `http://bergmair.cjb.net/`