Towards Human Interactive Proofs in the Text-Domain

Richard Bergmair
University of Derby in Austria

and

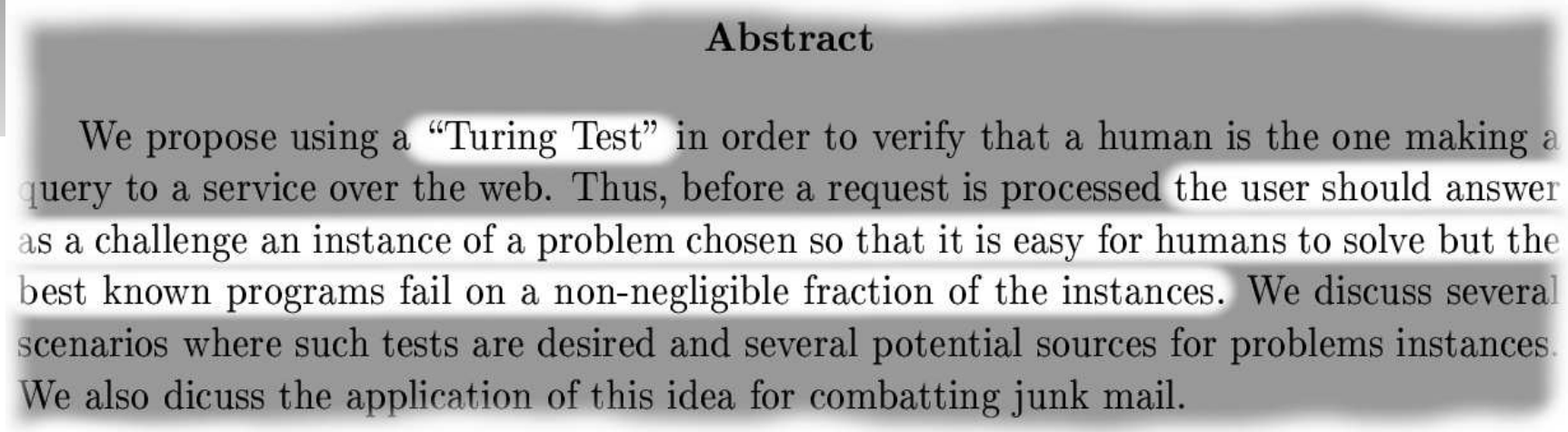
Stefan Katzenbeisser
Technische Universität München
Institut für Informatik

Introduction & Prior Work

Many serious threats to Information Security rely on attacks that can only be carried out by computers, not by humans:

- manipulation of online polls
- bulk subscription to web-services
- distribution of spam and worms
- privacy infringement by unwanted data mining
- denial-of-service attacks
- dictionary attacks

Introduction & Prior Work

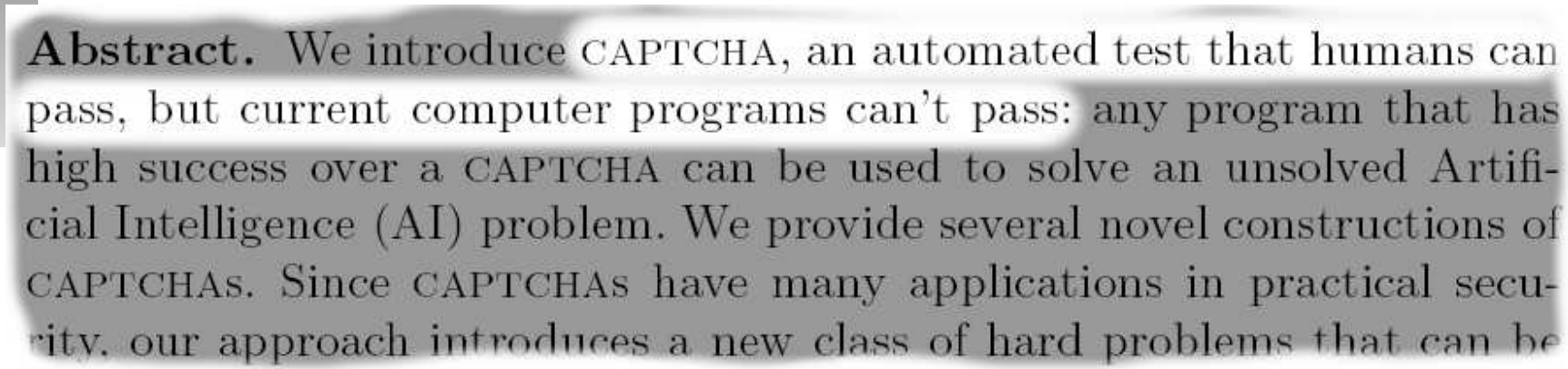


Abstract

We propose using a "Turing Test" in order to verify that a human is the one making a query to a service over the web. Thus, before a request is processed the user should answer as a challenge an instance of a problem chosen so that it is easy for humans to solve but the best known programs fail on a non-negligible fraction of the instances. We discuss several scenarios where such tests are desired and several potential sources for problems instances. We also dicuss the application of this idea for combatting junk mail.

Moni Naor. Verification of a human in the loop or identification via the turing test. Unpublished Manuscript. `http://www.wisdom.weizmann.ac.il/~naor/PAPERS/human.ps`, 1997.

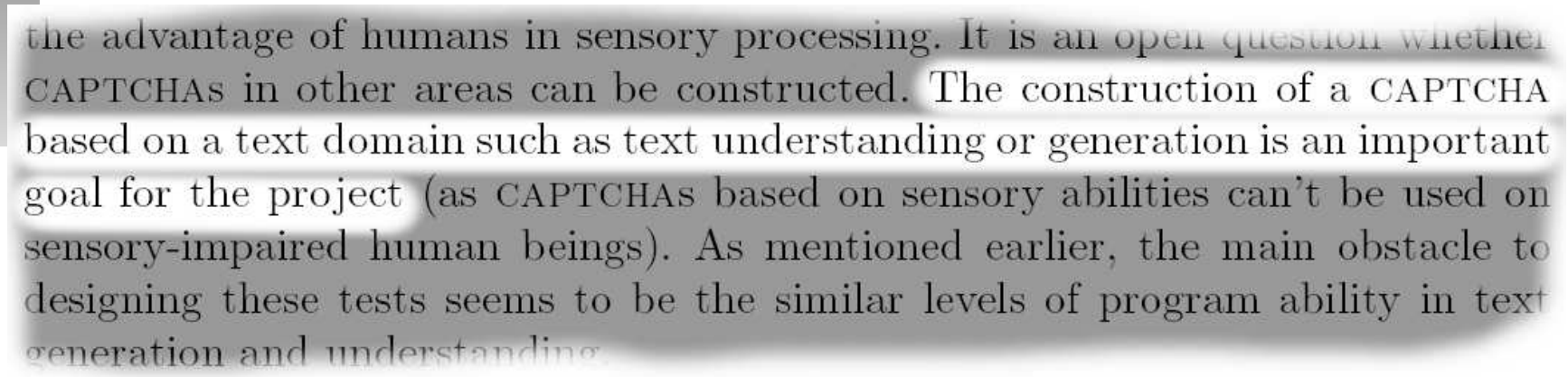
Introduction & Prior Work



Abstract. We introduce CAPTCHA, an automated test that humans can pass, but current computer programs can't pass: any program that has high success over a CAPTCHA can be used to solve an unsolved Artificial Intelligence (AI) problem. We provide several novel constructions of CAPTCHAs. Since CAPTCHAs have many applications in practical security, our approach introduces a new class of hard problems that can be

Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. CAPTCHA: using hard ai problems for security. In *Advances in Cryptology, Eurocrypt 2003*, May 2003.

Introduction & Prior Work


the advantage of humans in sensory processing. It is an open question whether CAPTCHAs in other areas can be constructed. The construction of a CAPTCHA based on a text domain such as text understanding or generation is an important goal for the project (as CAPTCHAs based on sensory abilities can't be used on sensory-impaired human beings). As mentioned earlier, the main obstacle to designing these tests seems to be the similar levels of program ability in text generation and understanding.

Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. CAPTCHA: using hard ai problems for security. In *Advances in Cryptology, Eurocrypt 2003*, May 2003.

Introduction & Prior Work

Unfortunately, images and sound alone are not sufficient: there are people who use the Web that are both visually and hearing impaired. The construction of a CAPTCHA based on a text domain such as text understanding or generation is an important open problem for the project.

Luis von Ahn, Manuel Blum, and John Langford. Telling humans and computers apart automatically. *Communications of the ACM*, 47(2):56–60, 2004.

Introduction & Prior Work

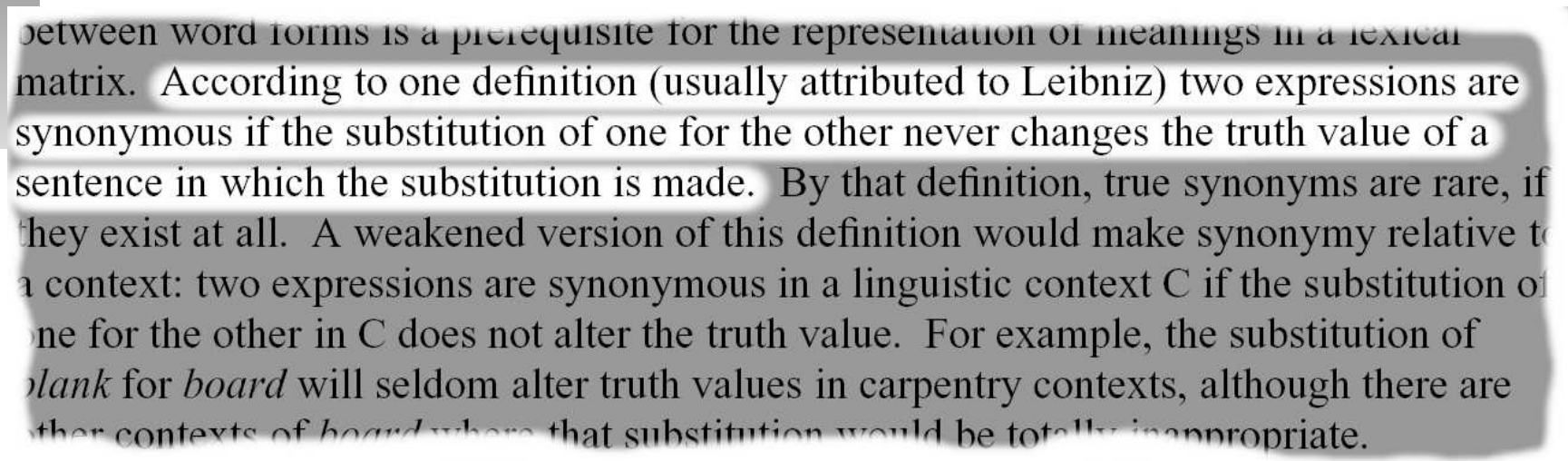
On the (im)possibility of a text-only CAPTCHA

Bartosz Przydatek CS Dept, CMU bartosz@cs.cmu.edu

I will present a formal model developed by Manuel Blum in collaboration with Luis von Ahn to capture properties of certain text-only CAPTCHAs. The model gives some insights why it is hard, if not impossible, to create a text-only CAPTCHA.

Unpublished Abstract from *First Workshop on Human Interactive Proofs*, January 2002.

Sense Ambiguity



between word forms is a prerequisite for the representation of meanings in a lexical matrix. According to one definition (usually attributed to Leibniz) two expressions are synonymous if the substitution of one for the other never changes the truth value of a sentence in which the substitution is made. By that definition, true synonyms are rare, if they exist at all. A weakened version of this definition would make synonymy relative to a context: two expressions are synonymous in a linguistic context C if the substitution of one for the other in C does not alter the truth value. For example, the substitution of *plank* for *board* will seldom alter truth values in carpentry contexts, although there are other contexts of *board* where that substitution would be totally inappropriate.

George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller.
Introduction to WordNet: An on-line lexical database.
`http://www.cogsci.princeton.edu/~wn/5papers.ps`,
August 1993.

Sense Ambiguity

- *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.*

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

Sense Ambiguity

- *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.*

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

Sense Ambiguity

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.

Sense Ambiguity

- *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.*

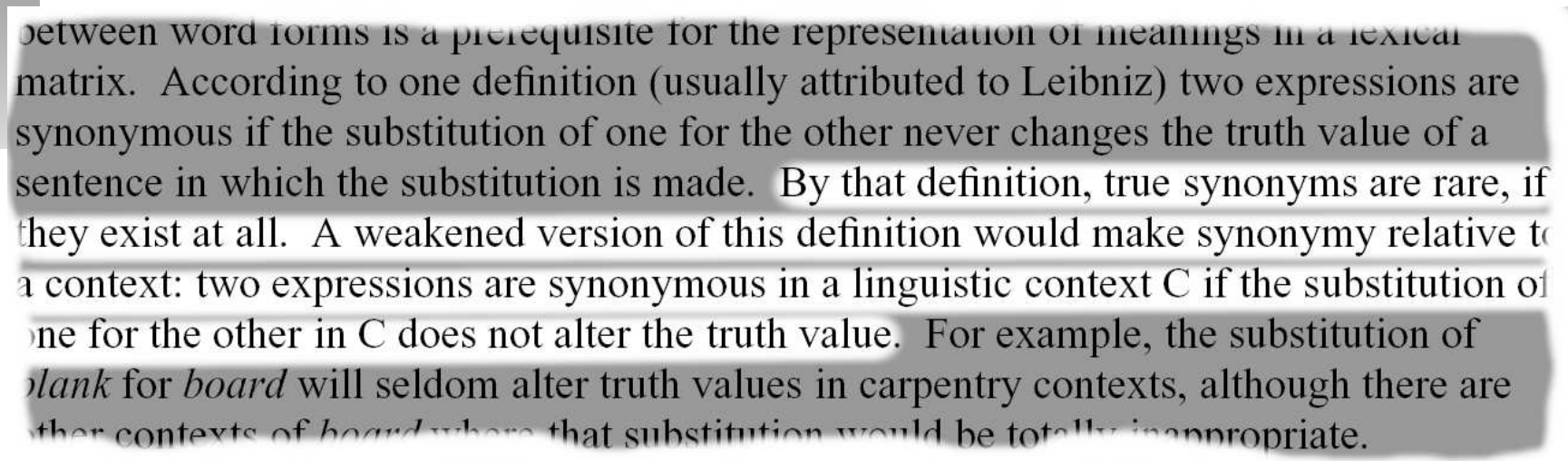
Sense Ambiguity

- *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.*

Sense Ambiguity



between word forms is a prerequisite for the representation of meanings in a lexical matrix. According to one definition (usually attributed to Leibniz) two expressions are synonymous if the substitution of one for the other never changes the truth value of a sentence in which the substitution is made. By that definition, true synonyms are rare, if they exist at all. A weakened version of this definition would make synonymy relative to a context: two expressions are synonymous in a linguistic context C if the substitution of one for the other in C does not alter the truth value. For example, the substitution of *plank* for *board* will seldom alter truth values in carpentry contexts, although there are other contexts of *board* where that substitution would be totally inappropriate.

George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller.
Introduction to WordNet: An on-line lexical database.
`http://www.cogsci.princeton.edu/~wn/5papers.ps`,
August 1993.

Sense Ambiguity

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 $c_1 \neq c_2$.

Sense Ambiguity

Pick the sentences that are meaningful replacements of each other:

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

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

Sense Ambiguity

The problem of automatic word-sense disambiguation has been under investigation in a computational context

- since the 1950s

and is of central importance for

- machine translation
- text mining
- spell checking
- text classification
- ...

Sense Ambiguity

	Fine	
	P	R
riori frequencies, by *requency*$^{\alpha}$ ($\alpha = 0.2$)	72.9	72.9
tic and syntagmatic info. , in an SVM classifier.	72.6	72.6
ech of neighbouring words, tions), in an SVM classifier.	72.4	72.4
of a-priori frequencies.	72.4	72.4
similarity, improved		

Rada Mihalcea, Timothy Chklovski, and Adam Kilgarriff. The senseval-3 english lexical sample task. In *Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, pages 25–28, Barcelona, Spain, July 2004.

Sense Ambiguity

We have introduced sense ambiguity making use of a function syn : $C \times W \mapsto 2^W$ that assigns to a word $w \in W$ used in context $c \in C$ the set $s \subset W$ of all words that are correct replacements of w.

We have presented evidence to suggest that no machine can reproduce syn with high accuracy. Humans can produce an annotation, by hand-crafting a table of associations sa $\subset$ syn, such that $|\text{sa}| \ll |\text{syn}|$.

Lexical HIP

What do we need?

- A public lexicon of words organized into sets of words that are synonymous in some linguistic context. (like WordNet)
- A corpus: A set of sentences that contain words also contained in multiple synsets of the dictionary.
- An initially hand-craftet secret annotation sa that is a subset of syn.

Lexical HIP: Generation Phase

t_1

- ◯ It should **move** through ... c
- ◯ It should **run** through ... $c_1 \in R(c)$
- ◯ It should **go** through ... $c_2 \in R(c)$
- ◯ It should **impress** through ... $c_3 \in Q(c)$
- ◯ It should **strike** through ... $c_4 \in Q(c)$

t_2

- ◯ We'll **send** your order ... d
- ◯ We'll **ship** your order ... $d_1 \in R(d)$
- ◯ We'll **broadcast** your order ... $d_2 \in Q(d)$

Lexical HIP: Testing Phase

t_1

⊙	It should **move** through ...	c
○	It should **run** through ...	$c_1 \in R(c)$
⊙	It should **go** through ...	$c_2 \in R(c)$
⊙	It should **impress** through ...	$c_3 \in Q(c)$
○	It should **strike** through ...	$c_4 \in Q(c)$

t_2

⊙	We'll **send** your order ...	d
⊙	We'll **ship** your order ...	$d_1 \in R(d)$
○	We'll **broadcast** your order ...	$d_2 \in Q(d)$

Lexical HIP: Verification Phase

t_1	⊙	It should **move** through ...	c	$\surd$
	○	It should **run** through ...	$c_1 \in R(c)$	$\times$
	⊙	It should **go** through ...	$c_2 \in R(c)$	$\surd$
	⊙	It should **impress** through ...	$c_3 \in Q(c)$	$\times$
	○	It should **strike** through ...	$c_4 \in Q(c)$	$\surd$

t_2	⊙	We'll **send** your order ...	d	$\surd$
	⊙	We'll **ship** your order ...	$d_1 \in R(d)$	$\surd$
	○	We'll **broadcast** your order ...	$d_2 \in Q(d)$	$\surd$

Lexical HIP: Learning

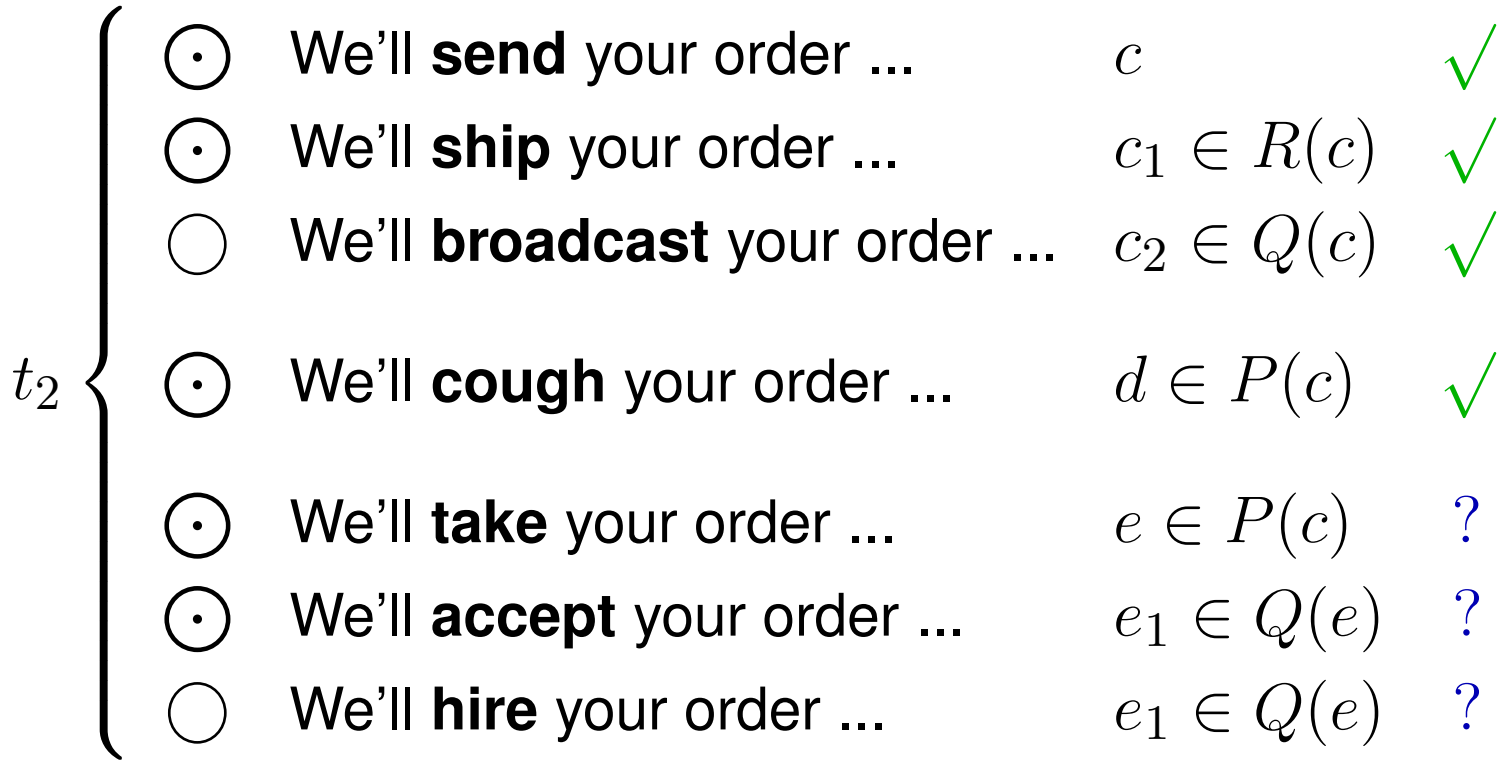
We have to trust in sa to be private at any time.

If we hand-craft it *once*, it will soon loose this property because whenever an association is used it is in fact published to the testee and to the adversary.

We have to think about sa as a dynamic resource, where we have to

- add new private associations
- remove associations if they are published

Lexical HIP: Learning Phase



Conclusions

In this contribution we have

- shown that the construction of text-based HIPs might in fact be possible.
- demonstrated word-sense ambiguity as a promising security primitive to build upon.
- presented the details of a construction automatically distinguishing computers and humans.

Conclusions

- The construction is **NOT** a CAPTCHA in the sense of a facility that does not rely on any private resources but a randomness source.
- **HOWEVER** we demonstrated that the security problems that arise from the use of a private database can be overcome by a learning approach.
- Details of such a learning construction were given.

Future Research

- We have pointed out the relevance of natural language semantics, and natural language learning to the construction of secure text-based HIPs.
- Since lexical methods provide only for the tip of the linguistic iceberg, we believe it will be fruitful to investigate the application of other methods as well, perhaps grammatical or ontological in nature.

Towards Human Interactive Proofs in the Text-Domain

Richard Bergmair
University of Derby in Austria

and

Stefan Katzenbeisser
Technische Universität München
Institut für Informatik