Computing with tangled words

Nicolaie Popescu-Bodorin, IEEE Member,
Artificial Intelligence & Computational Logic Laboratory,
Department of Mathematics and Computer Science,
Spiru Haret University, Bucharest, Romania.
http://fmi.spiruharet.ro/bodorin/, bodorin ☺ ieee.org

In the end of February 2011, Professor L. Zadeh forwarded to the BISC Members the following text:

> *Can you raed this? Olny 55 plepoe out of 100 can. I cdnuolt blveiee that I cluod aulaclty uesdnatnrd what I was rdanieg. The phaonmneal pweor of the hmuan mnid, aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it dseno't mtaetr in what oerdr the ltteres in a word are, the olny iproamtnt tihng is that the frsit and last ltteer be in the rghit pclae. The rset can be a taotl mses and you can still raed it whotuit a pboerlm. This is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the word as a wlohe. Azanmig huh? Yaeh and I awlyas tghuhot slpeling was ipmorantt! If you can raed this […].*

This remaind me a problem that I didn't solved yet. In the context of Artificial Intelligence, the problem of reading bothers me for quite a while. Especially this topic exemplified above. Why is that? Because instead of defining an entity of artificial intelligence as being a formal structure (vocabulary) interconnected through a formal grammar which together endow the entity with the capacity to achieve a goal - such is reading a text, we could think to accept the possibility that the inverse problem is, in fact, the natural way of defining how intelligence starts and how it evolves: let's assume that both the formal structure (vocabulary) and the formal grammar evolve during reading exercises. And they evolve in such a manner, that the language become recognizable even when the letters within the words are tangled, i.e. the intelligent agent (human mind) surpasses its basic design requirements, i.e. for human mind, the meaning of some text is invariant with respect to a family of permutations of letters within the words. In evolutionary terms, the intelligent agent (human mind in this case) enlarges and learns its vocabulary, but also discovers and learns a theory of understanding over its vocabulary (a formal theory of reading).

This performance is possible due to the fact that for the human mind, the meaning is the unit of discourse and not the letter. And in the given example, the meaning is recognized as a familiar pattern, so familiar that its understanding proves very high fault-tolerance.

Hence, until now, I figured out that artificial reading is a problem of training an intelligent agent to recognize familiar patterns.

The only problem is that when the intelligent agent is human, the pattern is a 'meaning' - a volatile representation in his mind, whereas for an artificial intelligent agent the training examples and the test examples are just strings, i.e. non-meanings. If the meaning would be represented computationally as lite as it is represented in the human mind, that text would be easily understood by an artificial intelligent agent.

How to represent the meanings in a lite and versatile computational manner is an unsolved problem for me.

Who is the brick of this formal (computational) language that I'm searching for?

Another important difference between computing with meanings (human mind) and computing with strings (artificial intelligent agent) gives me shivers:

- for the human mind, the effortless experience of reading (understanding) that text is an exercise of practicing 'well-formed' english language.
- for an artificial agent, the understanding of that text could come only after recovering the 'well-formed' english language from it.

This difference is huge. Where did we go so wrong to allow this?

Nicolaie Popescu-Bodorin,
February 26, 2011.