PERIODICA POLYTECHNICA SER. HUMAN and SOCIAL SCI. VOL. 2, NO. 1, PP. 57-68 (1994)

CAN ROBOTS LAUGH?

Tamás GERGELY and Miklós SZŐTS

Applied Logic Laboratory
1022 Budapest, Hankóczy J.u.7. Hungary
Phone and Fax: (36-1) 135-5386; (36-1)176-4916

Received: June 2, 1993

Abstract

For knowledge based systems knowledge, which is usually given as natural language descriptions, has to be transferred into formal representations. The authors argue that the expressive power of natural language lies partially in the possibility that it can be considered as a rich system of sublanguages. The category of theory morphisms is an adequate mathematical tool to handle the sublanguages of different subfields, different points of references and different levels of abstraction. To prove the claim jokes are analysed and it is shown that in this way a very abstract logical characterisation can be given. The paper tries to answer even the question in the title.

Keywords: humour, semantics, language hierarchy, theory morphism.

'Even a joke should
have some meaning...'

1. Introduction

Nowadays more and more so called knowledge based systems are manufactured. One may be surprised to learn that one of the most important bottlenecks is to describe the knowledge of a domain with precision enough. Even if the subconscious abilities of the experts are neglected, and good textbooks are considered to be available, it is a difficult task to translate the natural language descriptions into formal ones. Even if the domain has even mediocre complexity, then the adequate formal representation may need different sublanguages for different subdomains. Moreover, the formal description fixes the level of discussion while natural languages allow us to change the level of discussion without stating so. So at a formal, logical formalization we cannot do better but consider natural language as a conglomerate of sublanguages.

Let us see the most formalized field of human thinking: mathematics. Mathematical reasoning is governed by the rules of classical logic that

is predicate calculus. However, mathematical texts are never written in the strict, formal language of logic, but in a special fragment of natural language, namely in the so called mathematical language. More precisely mathematicians use a very complex hierarchy of languages. At the bottom there lays essentially the language of first order classical logic, since all sub-languages can be interpreted in it. Above it there are the languages of different fields of mathematics as well as the different dialects of the presentations of definitions, theorems and proofs. Language of formal mathematics can be strictly formalised on a higher level than predicate calculus, see GERGELY, VERSHININ (1981). If creative mathematical thinking is considered, then regions of non-formal, non-mathematical thinking also get an important role. Mathematical intuition is generally based on experiences coming from fields different the one of problem in question. The above said are true not only for mathematics, but in all fields of human intellect. Natural language can be considered as a rich system of closely connected languages. As for creative thinking KOESTLER (1977) shows that its underlying mechanism is the coupling of two self-consistent but habitually incompatible frames of reference. The sudden meeting of the different frames of reference may throw light on a really new solution (c.f. KOESTLER (1977) pp.101–120). This principle can be discussed from many different points of view, we want you to note only that if knowledge is stored as descriptions, the above mechanism calls for a highly structured collection of descriptions and can really effectively be discussed by using different languages.

So we think that using a complex system of languages (language hierarchy) instead of one language in AI systems would provide a way out of the present cul-de-sac. Note that this principle is not inconsistent with the present directions in AI. In theorem proving there is a great demand for metalanguages describing search strategies; there are more and more systems of hybrid reasoning, and the famous frame technique is based on hierarchically structured descriptions. However, we think of a highly complex usage of language hierarchies which would ensure smooth transfer between the languages, which would provide the system with the ability of using definitions, incomplete descriptions and recognizing analogies.

For the above reason the study of language hierarchies was started in ANDRÉKA, GERGELY, NÉMETI (1980 and 1981). Doing so we found that the usage of this notion may help to understand the mechanism of humour. This is supported also by KOESTLER (1977). He claims that the mechanism of humour and creativity in science (problem solving) is basically the same. So to test whether a theory of language hierarchies is powerful enough to be the theoretical basis of creative problem solving systems, in this paper we use it to explain the semantics of humour. We don't want to deal with

the social and psychological aspects of understanding humour, not denying that they have decisive importance. Our aim is a logical analysis that is we want to show that a text which considered funny can be characterized at a very abstract level.

2. What is Language?

To show the possibility of formal discussion, first we provide the definitions of the basic notions. Language can be studied at different levels (cf. GERGELY, SZŐTS (1982)). If it is not mentioned otherwise, definitions are at abstract level.

At the abstract level a language is considered as the pair of its syntax and semantics: $L =< F_L, S_L >$. The syntax is considered as the set of meaningful expressions, it may be given by some basic expressions and by those grammatical rules that generate the others. At abstract level the use of model theoretic semantics is adequate that is, semantics is the pair of class of models and meaning function: $S_L =< M_L, k_L >$. M_L is the class of possible worlds in which the expressions can be interpreted, and k_L is a function (defined on $M_L \times F_L$) giving the meaning of an expression in a model.So language can be considered also as a triple: $< F_L, M_L, k_L >$. In the case of the language of classical logic M_L is the class of relation structures, and k_L renders truth-value to the formulas. Function k_L connects syntax and models, though we also have syntactical tools to handle semantical questions, the so called calculi. These consist of syntactical text-rewriting rules preserving semantical features, e.g. a certain truth-value. The pair of a language and calculi is called logic.

GERGELY, SZABOLCSI (1979) show that model theoretic semantics is relevant also from the linguistic point of view. However, the usage of language, as the mechanism of understanding, cannot be properly handled at abstract level only. When speaking about representation of semantics we stipulate that a system using a language has some inner model of the environment. The representation of semantics consists of

- the set of the possible inner models,
- a connection between the syntax and these inner models, which simulate the interpretation function of semantics (cf. GERGELY, SZŐTS (1982)).

Natural language is highly complex, so it can be best handled as a complex system of sublanguages. Here we do not think of a predetermined division of the original language, but of the system of all possible sublanguages. Such a system is called a language hierarchy. MONTAGUE (1975) showed that an arbitrary wide fragment of a natural language can be in-

terpreted in a logical language similar to the classical one. So we may stipulate the existence of such interpretations, and instead of fragments of a natural language mathematically well defined languages can be considered, which it is interpreted in. So instead of the language hierarchy of a natural language a hierarchy of languages of a mathematical logic can be considered.

Language hierarchies are mathematically investigated in ANDRÉKA, GERGELY, NÉMETI (1980) and (1981). The connection between the languages is given by the interpretations. Let L_1 and L_2 be two languages. An interpretation from L_1 to L_2 is such a function from F_1 to F_2, which creates also a mapping between the meanings, and these two functions commute with meaning functions k_1, k_2. That is $f: \quad F_1 \to F_2$ is an interpretation, if there is f' such that

$$k_2(f(.)) = f'(k_1(.)).$$

Clearly the sublanguages with the interpretations form a category.

There may be or may not be interpretation between two languages. However, the investigations referred to above show that always can be found a connection between any pair of languages even if they are not connected. Namely there exist two languages with the following properties, respectively:

i) a more general language which can be interpreted in both languages;
ii) a less general language which both languages can be interpreted in, this one is said to be the common language of the original ones.

For those who speak the language of category theory it means that the category of the language hierarchy is complete and cocomplete.

Using several languages mixed together first were suggested by BURSTALL, GOGUEN (1977) for program specification. Language hierarchy is the mathematical version of their idea. Other formal tools can also be used for the same purpose, like situational structures in GERGELY, VERSHININ (1993). Language hierarchy seems to be the most general form, that is why we use it in the present paper.

3. The Basic Idea

KOESTLER (1977) states that the source of humour is 'the perceiving of a situation or idea in two self-consistent but habitually incompatible frames of reference'. RASKIN (1979) claims basically the same, and several of the classical theories can be fitted into this thesis. We formalize this idea according to our purposes:

(i) two different sublanguages meet in a text in an unusual and often irregular way, however
(ii) there is a link between them.

Note that this can be generalized to non-verbal humour, if we speak about 'language' of pictures, or 'language' of gestures or about any similar 'languages'.

From the above feature two directions of this study open up:
(i) the examination of what sublanguages are confronted with,
(ii) the examination of the links, connecting them.

As for (i) one can think about such sublanguages as
- the languages of different sciences,
- the languages of different arts,
- official language,
- nursery language,
- military language, and so on.

Any further study of this question is beyond of our line of investigation, it is rather subject to psychology or sociology. We can but only indicate the contrasted languages in square brackets, if they can be labelled easily. Our purpose is to study the links between the confronted languages.

4. Basic Mechanism: Double Meaning

(1) The Junior String Quartet **played** Brahms last night. Brahms lost. [music — sport]

The first sentence is one of the language of music, the second is constructed as one of the language of sport. The link is provided by the word **played** which can be interpreted both in the universe of music and the one of sport. When the whole text is interpreted, the second sentence has to be meant metaphorically: the performance was a loss for the case of music as well as for Brahms. Were this plainly stated the text would not be humorous at all. Clearly the humour in this message is in the way it is delivered that is in the sudden change of language.

The above shown mechanism can be discussed formally at abstract level as follows.

Let $L_1 =< F_i, M_i, k_i >$ and $L_j =< F_j, M_j, k_j >$ be two sublanguages, and the same syntactical unit (word or phrase) s be element of both F_i and F_j. However, let the value of $k_i(s)$ and $k_j(s)$ be different and/or the two classes of models be also different. Let α, β be texts of L_i and L_j, respectively, and let us consider the texts $\alpha s \beta$ (see *Fig. 1*).

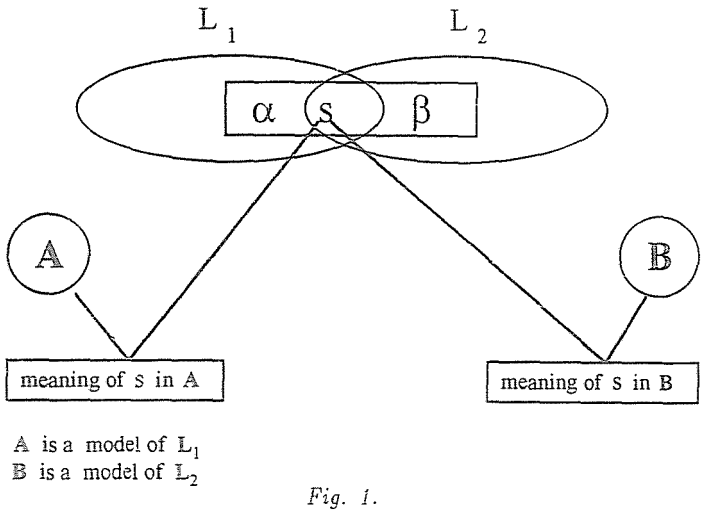


Fig. 1.

If text $\alpha s \blacksquare \beta$ is interpreted until the point marked by $\blacksquare$, the meaning may be definite and unambiguous, but perceiving β a second meaning enters, connected to the same text s. The whole text $\alpha s \beta$ may or may not have the meaning in both models but has to have some in one of them. This mechanism will be referred to as double meaning.

Naturally the scheme $\alpha s \beta$ is only a rough one. In (1) the word **played** plays the role of s. It is embedded in α (the first sentence), and interpreting β we remember it. Later on we see more sophisticated mixing of the elements of this scheme.

Note that the syntactical unit s is the link between the two languages. In certain situations such a syntactical link alone is sufficient to evoke laughter, but it can hardly be called humour. Usually there is also some connection between the two different meanings. In the case of (1) we can appreciate the joke, if we can refer the verb lost to the composer Brahms, whose opus was played.

In what follows we discuss some jokes with exactly the same scheme.

(2) An editor spent a whole afternoon cementing holes in the sidewalks in front of his house, only to have a kid come along on a bike and make ruts through the fresh **concrete**. The editor let go some sulphurous words. Hearing his language, the kid's mother protested.

'I thought you loved children' she said. 'I do' replied the editor. 'But in the abstract — not in the **concrete**'.

Here the phrase 'in the abstract — not in the concrete' belongs to a highly abstracted language, but because of the context of the first occurrence, the word 'concrete' may mean a really concrete material. The two meanings of the word concrete are independent from each other, the semantic link is created by the situation.

Generally the two meanings have a natural connection, like in the following joke.

(3) This girl reminds me Dreyfus, the army does not believe in her innocence.

[politics — sex]

Here the key words are:

$$\underbrace{\text{Dreyfus}}_{\text{politics}}\quad\underbrace{\text{innocence}}_{\text{sex}}\quad\text{girl}$$

Note the clever timing: the connection is stated in advance, and explained later, that is, the scheme takes the form $\alpha\beta s$.

'Non-fiction' stories also support our thesis about the role of different sublanguages in humour. The following one took place in 1944 while liberating Paris:

(4) The battle's principal victim was one of Gabriel's massive Corinthian columns, the fifth from the left along the facade of the Hotel Crillon. According to the legend, it was shot apart by the gunner of the tank destroyer 'Filibuster' after he had been warned by his commands to 'watch out for the 'fifth column'. The commander was referring to the collaborationist snipers.

[military — architecture]

5. The Role of Background Knowledge

(5) Ug, the caveman, observed his mate running to him in tears, her leopard-skin skirt in disorder. 'Ug', she cried, distraught, 'do something quickly. A sabre-toothed tiger has entered Mother's cave. Do something!' Ug grunted, picked up his well-gnawed buffalo bone and said, 'Why do anything? Who the hell cares what happens to a sabre-toothed tiger?'.

[jungle — family life [stone age — modern age]]

In (5) you do not find the syntactical link. However, let us remember that language is not merely a syntax, but the triple of syntax, class of mod-

els and meaning function. Contrasting of the two languages here happens on the level of models. The words of the wife belong to the common language, but Ug denies to interpret it there — he interprets it in the models of family life, where the mother-in-law is the most dangerous to him. It can be said that he constructs the common language badly.

However, this case can also be discussed by the help of syntax. In many cases the class of models can be described by sentences of the language, that is by axioms.

Let $L =< F, M, k >$ be a language, let Ax be a consistent set of formulas, and let $Mod(Ax)$ denote those models of M, which satisfy all sentences in Ax. Then $< F, Mod(Ax), k >$ is a sublanguage of L. If Ax_1 and Ax_2 define two languages, the common language can be defined by an Ax_3. However, in several cases AX_3 cannot be the union of Ax_1 and Ax_2. That is the case in (5), where an axiom of jungle, like

dangerous (tiger)

ought to overwrite a similar axiom of family life

dangerous (mother-in-law).

The set of axioms determining the class of models of different languages is a description of what we call background knowledge. Background knowledge is essential to understand jokes. Let us think for example of the special knowledge we have to appreciate (3).

In the following we give some typical constructions, where the syntactical link is not important or is missing altogether, and the double meaning is created primarily in a semantical way.

5.1 Double Meaning by Improper Language

(6) To bolster attendance at the church choir he directs, husband sent out this notice: 'Free admissions. New anthems with striking cover designs for liberals. Old anthems with familiar cover design for conservatives. Exiting choir responses. Fun benedictions. Stylish choir robes with two precut slits in the back for wing potential'.

[advertisement — religion [church-transcendental]]

Above the language of advertisement is connected to an ecclesiastic matter. To put the words of an improper language to the mouth of a person is often source of humour, if the words may remain meaningful.

Joke (6) merits our attention for some other factors, too. First, the elements of the two languages are mixed together unseparably. As to have got it one has to go back and forth between the two languages. Secondly,

note the sudden change of worlds in the last sentence. The reference to the angel's wings brings us from the atmosphere of an ecclesiastic choir to the transcendental world, which only widens the gap between the languages.

The mechanism of improper language is one of the most widely used: e.g. in the cases, when animals and things are anthropomorphised. It is also the case, when in scientific publications the author uses jokes or quotations from literature as mottos.

5.2 Improper Language by Situation

(7) After losing a bridge game, the wife glared at her husband and said

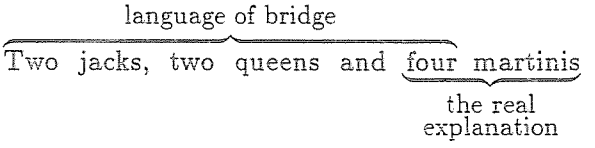
'I had four aces and three kings. What in the world did you bid no trump on?'

'Two jacks, two queens and four martinis'.

[bridge-party]

This is similar to the case discussed before, since a person says something he is not supposed to. However, the improper utterance is not because of a predetermined role: a husband having martinis is humorous only in the situation described in the joke.

Let us see the semantical link:



Note, that here the simple word **four** is the syntactical link. Proof by counter-example: a sensible answer, like 'I had only two jacks, two queens but the martinis I had had made me bid so' would not be humorous at all. A good example for the role of models is that the meaning of word **four** remains the same, but it has to be interpreted in different models.

6. The Role of Calculus

One can often meet such jokes, where humour is connected with some kind of implicitly presupposed knowledge. The latter can be considered as a set of axioms which is the base of the effect of jokes. A non-valid statement is not only stated as valid, but it is also used as a basis of further implication. Having the jokes, you have to trace back this implication to find the double meaning. Of course enjoying a joke this process is unconscious. Note that

double meaning exists also in this case, since the meaning of a proposition is its truth-value. Let us see an example:

(8) The prince, travelling through his domains, noticed a man who bore a striking resemblance to himself. He beckoned him over and asked:
'Was your mother ever employed in my place?
'No, Sire' the man replied 'But my father was'.

Here we first have to infer, what the question really is after; and similarly to infer the real meaning of the answer.
However the role of calculi may be more important.

(9) Any **big** men born around here?' a tourist asked in a condescending voice. 'No', responded the native. Best we can do is babies. Different in the city, I suppose.

Here the double meaning at word **big** is a simple business, but the last sentence suggests a strange analogy: 'the bigger the place, the bigger the new-born babies are'. So the native has a rule of inference, which is quite uncommon. In similar cases it is easier to contrast **logics** instead of contrasting languages.
Note the more effort has to be spent to trace the conflict in a text, the nearer we are to puzzles, moreover this way we get to the famous paradoxes of logic.

7. Conclusions in the Form of Questions and Answer

– Have we defined humour?

The semantic characterisation alone cannot define humour. As KOESTLER (1977) writes, the mechanism of the three characteristic human abilities — namely humour, art, and creative problem solving, — is the same. This means that it cannot be decided on semantic level, whether an utterance is humorous, has aesthetic value, or is a description of a creative step in solving a problem. According to Koestler the emotional attitude makes the difference between humour, art, and creative problem solving. Fitting it into this paper's terminology, we may say that the further characterisation of humour is not a semantical, but a pragmatical question.

– **Why do we laugh?**

While in the semantics of humour the main idea of this paper seems irrefutable, this question cannot be answered by studying languages, logics. It is a question of psychology. At any case laughter is a way to resolve stress. In the case of verbal humour while passing the text, the subject has to change languages (frames of references) unexpectedly. This surely may cause some stress. According to Koestler intellect can follow this change, but emotion, having greater inertia, cannot. He claims, that this incongruity is dissolved by laughter (KOESTLER (1977) pp.55–64).

– **Can Robots 'understand' humour?**

This paper has showed that at a semantic level humour can be characterised by perceiving a situation or idea in two self-consistent, but usually incompatible languages or frames of reference. Our tools for knowledge representation are theoretically able to detect such features. We used **logic** in our analysis, but RASKIN (1979) made similar investigations using **scripts** as representation of semantics, and from the formulation of the feature it is clear that frames are proper tools to detect it. However, in spite of the theoretical possibility it would be nearly impossible to describe the background knowledge needed for appercieving any of the jokes quoted here. The problem is the same, as in problem solving systems: the practical problems are too complex. However, if we have really intelligent robots, they will be armed with the ability to handle a complex hierarchy of languages. So they have to detect the semantic features of some jokes parsing its text. Naturally we do not think of jokes dealing with overcomplicated human relations (sex, politics, etc.) However, mathematics has its own humour, there are humorous chess puzzles, too. For example we think that a really intelligent chess program has to be able to detect a humorous (or beautiful) game.

And finally:

– **can robots laugh?**

According to the above said, understanding humour does not imply the ability of laughter. However, the sudden contrast of the incompatible languages will probably create some kind of 'stress' in the intelligent robots of the future, and this stress has to be neutralized somehow. Remember the sci-fi stories, where robots or intelligent supercomputers are driven mad by paradoxes! It does not seem impossible that robots will have an ability with the same function as laughter in human beings.

References

(Referring to language in general)

ANDRÉKA, H. – GERGELY, T. – NÉMETI, I. (1980): Model Theoretic Semantics for Many-purpose Languages and Language Hierarchies, COLING, Tokyo.

ANDRÉKA, H. – GERGELY, T. – NÉMETI, I. (1981): On Semantics of Language Hierarchies, SZÁMKI Working Paper, Budapest.

BURSTALL, R. M. – GOGUEN, J. A. (1977): Putting Theories together to Make Specifications. *Proc. 5th Int. Joint Conf. Artificial Intelligence*, Cambridge, Mass. p. 1045.

GERGELY, T. – SZABOLCSI, A. (1979): How to do Things with Model Theory, CL and CL, Vol. 13, Budapest.

GERGELY, T. – SZŐTS, M. (1982). About Representation of Semantics, Progress in Cybernetics and System Research, Vol. 11. p. 227. Hemisphere Publication Corporation.

GERGELY, T. – VERSHININ, K. P. (1981): Concept Sensitive Formal Language for Task Specification Mathematical Logic in Computer Science, ed. B. Dömölki, T. Gergely North Holland.

GERGELY, T. – VERSHININ, K. P. (1983): A Method of Formalization of Hardly Formalizable Disciplines (see ALL working paper, Budapest).

MONTAGUE, R. (1975): Formal Philosophy. Academic Press, New York.

(Referring to humour):

KOESTLER, A. (1977): The Act of Creation. Pan Books Ltd. London.

RASKIN, V. (1979): Semantic Mechanism of Humour, in *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistic Society C. Chiarello ed* University of California, Berkeley.

(The sources of the quoted jokes)

Our motto comes from L. Caroll's Through the Looking Glass. The Annotated Alice. Ed. by M. Gardner, Penguin Books, Harmondsworth, England, 1970.

Some of the jokes are also analyzed in the referred publications:

(8) by KOESTLER, A. (1977) p. 84, (1), (3), (9) by RASKIN, V. (1979) p. 332, p. 327, p. 333, respectively.

(The sources of the other jokes):

Reader's Digest, July, 1978: (2) – p. 75, (6) – p. 33, (7) – p. 74.

(4) – comes from: Is Paris burning? by L. Collins and D. Lapierre, Simon and Schuster of Canada Ltd. 1966, pp. 144–145.

(5) comes from the sci-fi story 'Jokester' by I. ASIMOV, see in: More Penguin Science Fiction (ed: Brion Aldiss), Penguin Books, Harmondrworth, England, 1976.