Biophysics, Vol. 37, No. 2, pp. 242–248, 1992
Printed in Great Britain

0006–3509/92 $15.00+0.00

UDC 577.3

BIOPHYSICS OF COMPLEX SYSTEMS

COMPUTER MODEL OF A "SENSE OF HUMOUR" I. GENERAL ALGORITHM*

I. M. SUSLOV

Lebedev Physics Institute, Russian Academy of Sciences, Moscow

(*Received* 8 *August* 1991)

A computer model of a "sense of humour" is proposed. The humorous effect is interpreted as a characteristic malfunction in the course of information processing due to the need for the rapid deletion of the false version brought into consciousness. The biological function of the sense of humour consists in speeding up the bringing of information into consciousness and in the fuller use of the resources of the brain.

In ordinary life we use humour as entertainment, as a "means of obtaining pleasure directly from the psychic process" [1] and do not bother about what aim nature pursued in endowing us with it. The very existence of a complex biological mechanism ensuring the advent of specific muscular contractions — laughter — as a reaction to a particular combination of sound or visual images indicates that the "sense of humour" appeared at quite early stages of evolution,† in the harsh conditions of the struggle for existence when the possibility of obtaining pleasure was hardly of fundamental importance. The present work seeks to answer the question of the biological function of humour.

In the proposed scheme, the humorous effect is interpreted as a characteristic malfunction in the course of information processing due to the biological need to speed up the bringing of the processed information into consciousness and the fuller use of the resources of the brain. In the present work we formulate a general algorithm of the origin of the humorous effect; the following paper deals with its realization in neuronal networks and discusses the mechanism of laughter [3].

HUMOUR FROM THE PSYCHOLOGICAL STANDPOINT

In psychology several views on humour exist [4, 5], the best grounded of which is the concept of incongruity already advanced in 1776 by the Scottish poet Beattie [6]. Different authors differ in its concrete interpretations; we adopt a view close to that outlined in [4]: the humorous effect

---

* *Biofizika*, 37, No. 2, 318–324, 1992.

† According to Darwin [2] anthropoid monkeys possess a clearly marked sense of humour.

is associated with the "switching" in human consciousness of two mutually exclusive images (versions, evaluations). Thus, in the example:*

"On a mound of earth outside his house
sits grandpa Anton twisting the goat's little leg.
The goat cries." (1)

Translator's note: in Russian, "a goat's little leg" is an idiom, signifying "a self-twist", i.e. a home-made hand-rolled cigarette.

The expression "the goat's little leg" at first is perceived as a "self-twist" and then assumes the meaning of the "leg of the goat". In this case we have a typical example of a play on words when switching occurs at the level of the meanings of the individual word; however, it may also occur at the level of more complex images:

"The horse dealer: " . . . if you sit on this horse at four in
the morning you will be in Pittsburg by seven in the morning." (2)
The buyer: ". . . and what will I do in Pittsburg at seven in the morning?"

Here the words of the dealer perceived as the characteristic of the swiftness of the horse are interpreted as an "explanation" of the way of dropping into Pittsburg at seven in the morning. The example

"Is this the place where the Duke of Wellington uttered
his famous words?" (3)
"Yes, this is the place but he never said them."

shows that switching may occur along the lines of a general evaluation of the phrase: the second retort at first gives the impression of being "logical" or "natural" but is then perceived as "absurd".

The presence of two mutually exclusive versions can be established in all witty expressions. Distinct switching of versions occurs in about half the cases. The principle of construction of other jokes may be called the "double entendre scheme". In the example:

"Criticism from below never reaches true heights." (4)

the word "heights" may be interpreted as "heights in the artistic sense" or as "the higher echelons of power". Characteristic here is the presence of the practically total equivalence of the two versions: accordingly, no definite sequence of their appearance in consciousness exists. Apparently, it may be considered that in such cases there is switching of the versions although the order of their appearance is determined by random factors: repeated passage from one version to another is also possible.

The humorous effect is associated not only with the "wit" discussed above but also with the "comical" (exaggerated movements of the clown, grimaces, caricatures, parodies, etc.). Apparently, the main sign of the comical may be considered to be "deviation from the norm"; accordingly, the humorous effect on its perception is associated with repeated "norm"–"not the norm" switchings.†

Laughter on tickling may be linked with the attempt of the brain to localize the site of irrita-

---

* A classification of the technical aspects of wit is given in Freud's book [1] (see also [7]); we give a simplified classification in relation to the goals of the present work: in principle, it embraces all the cases considered by Freud. Examples (2) and (3) are taken from Freud [1], example (4) from the *Literaturnay gazeta* (No. 20, 1990), and example (1) from the folklore of the students of the Moscow Physicotechnical Institute.

† Of course, not every deviation from the norm appears funny; however, it should be borne in mind that customary oft-repeated deviations from the norm are similar to "stale" jokes (see below) and slight deviations from the norm are readily vulnerable to ousting by other emotions [3].

tion of the skin; the result of such localization is continuously rejected since the site of irritation changes unpredictably (therefore, someone else must do the tickling).

INFORMATION PROCESSING

Let us begin the formulation of the computer model of the "sense of humour" by analysing information processing. Let a "text" come into the brain from the outside world — a sequence of symbols $A_1, A_2, A_3 \ldots$ — for example, a sequence of words read or heard. In the brain with each symbol A_n is associated a certain set of images $\{B_n\}$ — for example, to each word-symbol is counterposed the totality of its meanings, i.e. a dictionary paragraph. The task of information processing is to choose from each set of images $\{B_n\}$ one image B_n^{in} which is understood in the given context. We shall consider that the text is "understood" if to the sequence of symbols $A_1, A_2, \ldots$ is counterposed the sequence of images $B_1^{i1}, B_2^{i2}, \ldots$, which may be represented as the trajectory in the space of images (Fig. 1).

In principle, the algorithm of information processing consists of the following: (1) all possible trajectories in the space of the images are composed; (2) on the basis of the information stored in the memory on the compatibility of images, a certain probability is ascribed to each trajectory; and (3) the trajectory with the maximum probability is chosen. Only step 2 is non-trivial, i.e. the algorithm of determination of the probability of the given trajectory. Such an algorithm may be based, for example, on the pair correlations between adjacent images: then a set of probabilities p_{ij} must be stored in the memory that in the comprehended text the image i is followed by the image j and the probability of the trajectory $ijkl\ldots$ is written as $p_{ij}p_{jk}p_{kl}\ldots$. The probabilities p_{ij} may be obtained from statistical analysis in the course of "learning" consisting in the introduction into the brain of a sufficiently long piece of "decoded" text (written in images and not symbols). One may use a more complex algorithm allowing for the correlations not of two but n adjacent images: then the memory must contain probabilities $P\{i_1, i_2, \ldots i_{n-1}, i_n\}$ that the sequence of $(n-1)$ images $i_1 i_2 \ldots i_{n-1}$ followed by the image i_n. One may allow for the pair correlations of the images but with account being taken of their syntactic links,* etc. Algorithms of such a kind are worked out in machine translation studies [8], the concrete form of the algorithm being of no importance for the further discussion.

The realization of any algorithm of such a kind requires a number of operations exponentially rising with the length of the text. Therefore, fragments of the text not exceeding a certain number N of symbols may be processed directly in such a manner. How to process longer texts? A natural possibility is to keep not one but several (M) most probable trajectories on processing the first N symbols. Then it is necessary to shift by one step — look at the segment from the second to the $(N+1)$th symbol and for each of the M retained trajectories construct all possible continuations. From them again the M most probable is kept, and so forth. It is rational to make the number M variable, so that at each step just as many trajectories are memorized as can be placed in the volume of the operative memory set aside for this. As a whole the process appears as follows

---

* The syntactic structure of a sentence may be represented in the form of a tree so that each dependent word is united with its "host" ([8], p. 58). The probability of the trajectory in the space of the images is represented by the product of the probabilities of the paired links between images corresponding to the structure of the tree. The practice of machine translation [8] shows that the syntactic structure in most cases is clearly established by purely grammatical analysis (word order, adherence to a part of speech, harmonization of the endings, etc.) and for the purpose of the present work may be taken as known.

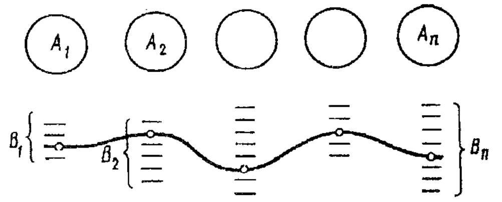

Fig. 1. Scheme of information processing: to each of the symbols A_n is counterposed the set of images $\{B_n\}$ from which it is necessary to choose one image B_n^{in}; the sequence $B_1^{i1}, B_2^{i2}, \ldots$ may be represented as a trajectory in the space of the images.

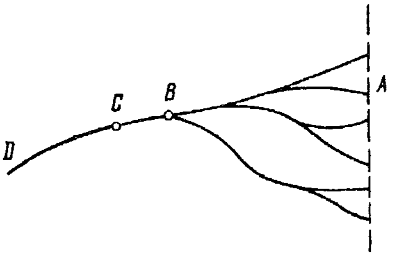

Fig. 2. Graphic representation of the information processing: fine lines are the trajectories contained in the subconscious, A is the leading edge, B is the point of termination of branching, DC is the portion of the trajectory brought into the consciousness.

(Fig. 2): close to the leading edge A the trajectory heavily branches; at a certain point B branching terminates (the distance between A and B is limited by the volume of the operative memory assigned for memorization of the trajectories) and, finally, the decoded trajectory DC with a certain delay AC is brought into the consciousness of the person and is perceived by him as a thought (the act of processing itself occurs in the subconscious and is not directly sensed by the person).

THE ROLE OF THE EMOTIONS IN INFORMATION PROCESSING

If the numbers M and N are sufficiently large and the algorithm of determination of the probability of the trajectory with a length of N symbols is sufficiently refined then the scheme of information processing described will work successfully. Nevertheless, because of the probabilistic character of the algorithm, errors are inevitable in its work: therefore, a mechanism reducing their after-effects to a minimum is necessary. Such a mechanism exists and consists in the fact that into consciousness comes a certain piece of information on how processing is proceeding in the subconsciousness; the human subject perceives this information in the form of emotions.

For example, of importance are such parameters of the process as the probability of the derived trajectory $p_{\max}$ and the probability of the most probable and competing trajectories p_{comp}. For high values of $p_{\max}$ and the ratio $p_{\max}/p_{\text{comp}}$ processing proceeds well and is accompanied by positive emotions (joy, certainty): the resulting information is considered reliable. Low values of $p_{\max}$ and $p_{\max}/p_{\text{comp}}$ indicate an unsatisfactory course of processing and are accompanied by negative emotions (dismay, doubt); for very low $p_{\max}$ no version at all may come into consciousness (total incomprehension). The concrete possibility of a link of the emotions with the parameters of the process may be illustrated on the basis of the semi-empirical "formula of emotion" established by Simonov [9]

$$E=N(I-I_0),$$

where E is the strength of the emotion (expressed objectively in the frequency of the pulse, blood pressure, etc.); N is the strength of a certain need; I_0 is the information necessary to satisfy this need; I is the information available to the subject (both items of information in subjective representation); for $I>I_0$ the emotion is positive ($E>0$), for $I<I_0$ it is negative. It may be assumed that in information processing N is the need for information and the different parameters of the process control different emotions. For example, $p_{\max}$ may be used as I when as E figures as the emotion "the joy of understanding — the disappointment of incomprehension" (accordingly I_0

is a characteristic value of $p_{\max}$ ensuring the reliable work of the algorithm). Similarly, $p_{\max}/p_{\text{comp}}$ may be used as I if E is the emotion "certainty–doubt" etc.

These ideas lead to the conclusion that an emotion consisting in the sensation of a humorous effect is also associated with a certain characteristic situation arising during information processing.

HUMOROUS EFFECT

Let us discuss the nature of the lag of point C behind the leading edge A (Fig. 2). At first sight in a rationally organized system point C must always be behind point B or coincide with it: we would act just so in drawing up a program for a computer. However, for man as in general for any living creature, such a variant is quite unacceptable. The point is that with the lag of point C relative to the leading edge A is linked a certain time interval τ_{AC} during which the information arriving at the brain is not delivered into consciousness (for example, the subject already sees a bear but is not aware of this). It is clear that the protraction of the interval AC may be dangerous; at the same time the interval AB may be drawn out for objective reasons (the subject cannot decide whether he sees a bear or a bush in the shape of a bear). Therefore, the interval AC must be upwardly limited in time by a certain value $\tau_{\max}$: if the time delay τ_{AB} corresponding to the interval AB is less than $\tau_{\max}$ then the point C coincides with the point B (Fig. 3*a*), if $\tau_{AB} > \tau_{\max}$, then $\tau_{AC} = \tau_{\max}$ and the point C begins to overtake the point B (Fig. 3*b*). In the last case the most probable version DE beings to come into consciousness (Fig. 3*b*) while the competing versions (DE') persist in the subconscious — their dismissal is undesirable since the possibilities of the brain allow one to continue the analysis.

If on further movement of the leading edge A the probability of the trajectory DE remains maximal, then the competing trajectory DE' will be dismissed and time will be gained. If on movement of the leading edge the probability of the trajectory DE drops below the probability of DE' then in the brain the possibility of correcting the error will exist. However, a characteristic malfunction appears: it is necessary to delete rapidly the portion BC brought into consciousness and to begin the derivation of the trajectory BE'. Psychologically this is perceived as interference of two mutually exclusive versions — the version BC preserved in long-time memory and the newly appearing version BE'. The characteristic malfunction described may be equated with the "humorous effect".

In fact, the situation described is exactly reproduced on perception of jokes: as in the example (1) in analysing the first sentence there appear in the subconscious two competing versions in one of which (DE) "the goat's little leg" is interpreted as a "self-twist" and in the other (DE') as the "leg of the goat". In the context of this sentence the version DE is more probable and

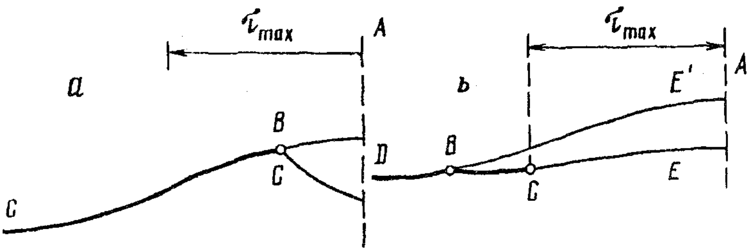

Fig. 3. The parameter $\tau_{\max}$ determined the maximum lag time of point C relative to the leading edge A (DC is the portion of the trajectory brought into consciousness); *a*, $\tau_{AB} < \tau_{\max}$; *b*, $\tau_{AB} > \tau_{\max}$.

thanks to a natural pause at the end of the sentence begins to come into consciousness. The appearance of the phrase "the goat cries" makes the version DE less probable but sharply increases the probability of the version DE' which also produces the humour.

We would stress that the existence of the humorous effect is in no way fatalistic: nature might avoid it in one of two ways: (1) through a delay in deriving the version BE until the version BE' is discarded in a natural way; and (2) through forcing of the derivation of BE with simultaneous rejection of BE'. However, in the first case the entry of the information into consciousness is delayed and in the second the resources of the brain are not fully utilized: therefore, nature solves this problem at the expense of psychological "confusion".

In the course of development an optimal value of the time delay, $\tau_{\max}$ is elaborated ensuring a compromise between the reliability of the information and the speed with which it is obtained. The optimal value of $\tau_{\max}$, as a rule, ensures fulfilment of the inequality $\tau_{AB}<\tau_{\max}$; therefore, in natural conditions the humorous effect appears relatively seldom; however, it may be readily induced with the aid of specially created witticisms and comical remarks.

Some consequences

The model described offers a natural explanation for some generally known facts.

1. The absence of a humorous effect of a joke heard many times is connected with the fact that the person knowing beforehand of the presence of two mutually exclusive images avoids bringing into consciousness the version known to be false (thus, knowing as in the example (1) "the goat's little leg" in the end is actually the leg of the goat he is not tempted to interpret it as the "hand-made cigarette").

2. The role of intonation in telling a joke is essentially linked with the temporal characteristics — tempo, position and duration of the pauses which may be taken into account by introducing into the sequence A_n the corresponding number of "gaps". For too rapid tempo the false version does not have time to be fixed in consciousness and the portion BC (Fig. 3) proves small. For too slow a tempo the length of the trajectories is increased through the "gaps" so that the competing trajectory BE' (Fig. 3) is discarded from the operative memory and switching of versions does not occur. Thus, in the example, (1) an optimal pause at the end of the first sentence is required.*

3. The different susceptibility of persons with an identical intellectual level to humour† is associated with the difference in the length of the time of the delay $\tau_{\max}$. In persons with a large value of $\tau_{\max}$ there rarely appears a situation when the point C overtakes the point B and accordingly the humorous effect seldom appears; in contrast, in persons with a low $\tau_{\max}$ the humorous effect appears even when in the view of the majority there is nothing funny. Apparently, in most persons the $\tau_{\max}$ value decreases under the influence of alcohol which leads to merriment without cause. For a fixed $\tau_{\max}$ susceptibility to humour correlates with the volume of the operative memory which determines the mean length of the interval AB (Fig. 2).

4. Nervous laughter. If a mass of unpleasant impressions impinges on a person and the threat of overstress of the nervous system is created then the body reacts to this by a forced rejection

---

* The dependence of the extent of the humorous effect on the duration of the pause is well described in Mark Twain's essay "Public Speaking".

† We consider the susceptibility in principle to humour, leaving aside the cases when the individual features lead to an inadequate reaction to a specific joke. Examples may be the incomprehension of a joke because of the absence in memory of the necessary image, the ousting of laughter through secondary emotions, non-standard notion of the "norm" on perception of the comical, etc.

of the unpleasant information and its replacement by another, neutral information — this induces reflex laughter.

CONCLUSION

A sense of humour is governed by the biological need to speed up the bringing of information into consciousness and the full use of the brain's resources. Therefore, the act of obtaining pleasure from laughter which Freud [1] was inclined to consider the main cause of the existence of a sense of humour (the person discovers the possibility of extracting pleasure from the psychic process and begins unconsciously and then consciously to exploit it) has no fundamental significance — just as sneezing and coughing due to the biological need to clear the airways exist independently of the pleasure derived from the first and displeasure from the second. Of course, if laughter caused displeasure then the social function of humour would change — society would try to get rid of it by censorial restrictions, persecution of witty persons, etc.

Is the creation of an effective computer program which would "laugh" in the same cases as man realistic? In our view it is quite realistic if one confines oneself to very simple jokes based on switching of the meanings of the individual words (example (1)); the corresponding program would be little more complex than the average program for machine translation [8]. Computer modelling of more complicated jokes involves the need to identify that vast set of images which the brain of the average person contains and to establish between them the correct associative connections — this requires many years of painstaking work of psychologists and programmers.

The author is grateful to L. A. Prozorova for discussion of the linguistic aspects of the work and to D. S. Chernavskii for fruitful discussions.

REFERENCES

1. S. Freud, in *Wit and its Relation to the Unconscious*, Moscow (1925).
2. C. Darwin, in *Collected Works*, pp. 773–775, USSR Academy of Sciences, Moscow (1953).
3. I. M. Suslov, *Biofizika*, **37**, 325 (1992).
4. J. A. Paulos, in *Mathematics and Humor*, University of Chicago Press, Chicago (1980).
5. In *Handbook of Humor Research*, (Eds P. E. McGhee and J. H. Goldstein), Vol. 1, Springer, New York (1983).
6. J. Beattie, in *Essays*, Edinburgh (1776).
7. A. N. Luk, in *A Sense of Humour and Wit*, Iskusstvo, Moscow (1968).
8. *Machine Translation* (Eds. O. S. Kulagina and I. A. Mel'chuk), Progress, Moscow (1971).
9. P. V. Simonov, in *The Emotional Brain*, Nauka, Moscow (1981).