Developing the Incongruity-Resolution Theory

Graeme Ritchie
University of Edinburgh, 80 South Bridge, Edinburgh, Scotland EH1 1HN.
g.d.ritchie@ed.ac.uk

Abstract

The idea of incongruity-resolution has frequently been suggested as an account of many types of joke. However, there is no precise statement either of this "theory" nor of its main concepts (incongruity and resolution), and different authors may disagree on details. We concentrate on two particular variants and attempt to clarify what would be needed to make these into computational models.

1 Introduction

In recent years, there has been increasing interest within artificial intelligence in the modelling of humorous behaviour, a phenomenon which had for some time been regarded as particularly challenging. One of the difficulties in constructing computational models of humour-interpretation or humour-generation is that there are no detailed formal theories of humour. One possible approach to remedying this deficiency is to take some of the many ideas that have been proposed informally within the literature of philosophy, psychology, linguistics, and other non-computational disciplines, and try to make the proposals more precise. That is the aim of this paper. Starting from a widely advocated idea, the *incongruity-resolution* theory of humour, we develop some sketches of what processing steps might be needed in interpreting certain classes of joke, and we indicate some research subproblems. The assumed problem will be the *interpretation* of a joke, rather than the generation of a joke, or a declarative statement of its form. This is not just because we wish to consider a specific computational task, but also because the proposals we are starting from all take that perspective.

Within this short paper, it is not possible to move from the existing ideas to a completely detailed or formal model, but we hope that the analyses presented here are a useful first step towards a more precise account.

As a simplifying assumption, we will consider (as many authors do) just humour conveyed in language – *verbally expressed humour*.

2 Incongruity

If there is one generalisation that can be extracted from the literature about humour, it is that humour involves *incongruity*. This point, with varying terminology, has been made by numerous authors (Keith-Speigel (1972) lists 24, the earliest from 1759), and remains current, appearing in work as varied as Katz (1993) and Attardo (1997).

The main problem, if we are to develop a detailed theory of verbally expressed humour, is that the notion of 'incongruity' is not clearly defined, and it is not even obvious that all the writers on this subject have exactly the same concept in mind. This is a typical definition from an ordinary dictionary:

incongruous: 1. out of keeping or place; inappropriate; unbecoming. 2. not harmonious in character; inconsonant; lacking harmony of parts. 3. inconsistent. [Random House College Dictionary (Revised Edition).1988. Random House, NY.]

Some academic definitions do not go beyond this in their precision, and it is rare to see anything more detailed than the following widely-cited definition:

> Laughter arises from the view of two or more inconsistent, unsuitable, or incongruous parts or circumstances, considered as united in one complex object or assemblage, or as acquiring a sort of mutual relation from the peculiar manner in which the mind takes notice of them. (Beattie, 1776), quoted in Raskin (1985)

3 Resolving Incongruity

There is sometimes a debate as to whether incongruity alone is sufficient for humour, or whether it is necessary for the incongruity to be *resolved*; that is, to be shown to be logical, or at least less incongruous than was first thought. The usual statement of the incongruity-resolution (IR) model postulates that humour is created by a multi-stage process in which an initial incongruity is created, and then some further information causes that incongruity to be resolved. Shultz (1976) claims that this analysis is of 'immense heuristic value in accounting for vast samples of humour'.

The way of creating and resolving the incongruity can be quite varied, as are the analyses presented in the literature. Rothbart (1977) suggests that in the joke (1), the incongruity is presented by the question part – namely, that the question appears to have a surprisingly easy answer – and then resolved in the answer (which is also a surprise).

(1) What is grey, has four legs, and a trunk? A mouse on vacation.

With example (2), Shultz (1976) says that the answer is initially seen as incongruous, with *wafer* interpreted as "a type of cookie", but then resolution occurs with the realisation that there is an alternative interpretation, "away for".

(2) Why did the cookie cry? Because its mother had been a wafer so long.

Pepicello and Weisberg (1983) quote Shultz (1974) as saying that the word *wafer* actually *creates* the incongruity, and go on to claim that in an 'informal study' most people interpret *wafer* as "away for" initially, thus subverting the analysis of Shultz (1976). Rothbart and Pien (1977) say that (2) has three incongruities: a cookie crying, the 'surprisingness' of the answer, and the cookie having a mother.

Rothbart and Pien also say that in (3), the question presents an incongruous situation, and the answer both explains (resolves) it and adds a new incongruity.

(3) Why did the elephant sit on the marshmallow? Because he didn't want to fall into the hot chocolate.

In contrast, Pepicello and Weisberg (1983) claim that the comparably structured joke (4) is simply a parody of a riddle, in which a question is posed which is impossible to answer.

(4) Why do elephants paint their toenails red? So they can hide in cherry trees

Suls (1977) analyses (5) in IR terms, since the answer 'comes as a surprise' but can be made to follow from the question with a little thought.

(5) If your son flunks out of school and is illiterate and anti-social, what can he grow up to be?
An Italian policeman.

Despite this apparent disagreement, there are some particular variants which are more clearly defined. Two of these, which are somewhat similar, are the *two-stage* model of Suls (1972) and the notion shared by (for example) (Shultz, 1976, p.13), (Minsky, 1980, p.10) and Paulos (1980), which here will be referred to as the *surprise disambiguation* model. (Although Suls claims that his model is also applicable to cartoons, we will consider it here only as an account of verbally expressed humour). For both models, there is a certain basic arrangement, as follows. A joke is analysed as being in two main parts: the initial portion of text, the *set-up* (or *joke body* (Godkewitsch, 1976)), and the second part, the *punchline*. The set-up creates no particular incongruity that the audience is aware of, but the punchline, at least initially, does not make immediate sense. However, a way is found to allow the punchline to be congruous (the resolution). Where the two variants differ is as follows:

Surprise disambiguation: The set-up has two different interpretations, but one is much more obvious to the audience, who does not become aware of the other meaning. The meaning of the punchline conflicts with this obvious interpretation, but is compatible with, and even evokes, the other, hitherto hidden, meaning.

Two-stage (Suls): The punchline creates incongruity, and then a *cognitive rule* must be found which enables the content of the punchline to follow naturally from the information established in the set-up.

There is no clear definition of the IR theory within the literature, but these two versions capture the essence of the idea, and most authors do not distinguish between the two, nor between them and the more general notion of IR (e.g. (Attardo, 1997, p.397)). The surprise disambiguation version is perhaps more prominent as an embodiment of the IR approach, and is often taken to be what is meant by an IR theory (it is also essentially Raskin's SSTH; see Section 4.1).

As has been implied above and observed by various writers, the general idea of IR may not be universally applicable to verbal humour, and each one of these two models fails to cover all the examples that might, informally, be argued to be IR phenomena. However, it is clear that there are at least some subclasses of jokes which are very naturally accounted for, at least intuitively, by these models, so it is worthwhile examining them in more detail.

4 The surprise disambiguation model

4.1 The various interrelations

There are various entities centrally involved in the SD account:

M_1: the first (more obvious) interpretation of the set-up

M_2: the second (hidden) interpretation of the set-up

M_3: the meaning of the punchline.

There are also various relationships and properties that are of interest, based on various observations made informally in the literature (the labels used here are invented as *ad hoc* mnemonics – part of the problem is that there is no standard terminology for these):

OBVIOUSNESS: M_1 is more likely than M_2 to be noticed by the reader.

CONFLICT: M_3 does not make sense with M_1

COMPATIBILITY: M_3 does make sense with M_2

COMPARISON: there is some contrastive relationship, even a clash, between M_1 and M_2.

INAPPROPRIATENESS: M_2 is inherently odd, eccentric or preposterous, or is taboo, in that it deals with matters not conventionally talked of openly, such as sexual or lavatorial matters, or forbidden political sentiments. These differ in terms of which norms are being flouted: those of everyday logic – leading to ABSURDITY – or those of socially acceptable discourse – leading to TABOO effects.

Although some or all of these relations or properties are sometimes proposed as the *essential* ingredients in IR humour, it is possible to find texts which, although plausibly of the general form of an SD joke, display only some of these relationships, yet which are humorous.

(6) Why do birds fly south in winter?
It's too far to walk.

For example, (6) relies on a (hidden) ambiguity regarding the focus of the initial question, and this is disambiguated in a surprising way by the answer; that is, it falls within the SD model. Although it could be argued that the hidden meaning is ABSURD, there is no evidence that it additionally has some significant COMPARISON with the more obvious meaning. Similar remarks apply to (7).

(7) Postmaster: Here's your five-cent stamp.
Shopper (with arms full of bundles): Do I have to stick it on myself?
Postmaster: Nope. On the envelope.

Hence not all of these relations or properties are necessary to create humour, even within the SD class of jokes.

It could be argued that there is no reason why theorists should arrange their analyses to conform to the particular divisions chosen here. However, all discussions of SD-style jokes admit, as pre-theoretic (or informally theoretic) entities, the three items listed above (the meanings of set-up and punchline). It is therefore reasonable to ask how their theoretical commentaries can be stated as properties of, or as relationships between, these basic entities. The lack of standard terminology means that it is sometimes hard to be sure what types of entity a proposed relation is between, and whether one author's "incongruity" is comparable to another's. The six relations/properties used here (CONFLICT, etc.) are merely chosen as some commonly occurring notions, to organise this discussion. Whatever variants of these are adopted in a theory of humour, we need fuller definitions in order to proceed, whether we are interested in scientific falsification or in computer simulations.

The above distinctions are not always made clearly. Most confusingly, several of them (particularly CONFLICT, COMPARISON, and INAPPROPRIATENESS) are referred to as 'incongruity', thereby obscuring the relationships. Shultz says 'the incongruity consists in the relation between the last line, or punchline, and the part that precedes the last line'(Shultz, 1976, p.13) (i.e. CONFLICT). Of example (3), Rothbart and Pien (1977) observe that the resolution involves an elephant sitting in a cup of hot chocolate, and refer to this as 'an incongruous situation' (i.e. ABSURDITY), and they suggest that this residual "incongruity" is a potential problem for IR theory, in that the oddity has not been eliminated ("resolved").

The Semantic Script-based Theory of Humour (SSTH) (Raskin, 1985) concentrates on COMPARISON as the central factor, labelling it *script opposition*, but does not offer much detail of how this opposition functions. (The importance of script opposition seems to be the main reason that Raskin regards the SSTH as having some substance beyond the already familiar ideas which are here dubbed 'surprise disambiguation'.) The SSTH also seems to package INAPPROPRIATENESS into script opposition, by regarding this (or other properties of the second interpretation) as relative to the first interpretation, rather than as inherent properties of the less obvious meaning. Giora (1991) proposes that the main mechanism in SD-style jokes is that the punchline provides a *marked increase in informativeness*, and that the interpretations differ in their 'markedness'. The notion of markedness is not formally defined, but this seems to be tackling CONFLICT, OBVIOUSNESS, and possibly COMPARISON. De Palma and Weiner (1992) discuss a notion of *accessibility* for the meanings of words, thus addressing OBVIOUSNESS. Attardo (1997) suggests that SSTH's 'script opposition' could be defined in terms of Giora's and De Palma and Weiner's concepts, thus replacing a description of COMPARISON with an amalgam of accounts of CONFLICT and OBVIOUSNESS.

4.2 Prediction

A further factor which is often discussed in the context of SD-style jokes (and elsewhere) is "surprise", or *violated expectations*, but the term "expectation" also lacks a precise definition. In some cases, it may just mean the "expectation" that the more OBVIOUS interpretation is the correct one. For example, (Suls, 1972, p.90) gives the example (8), with the observation 'the ending of the joke is unexpected and incongruous... but can be so interpreted as to make sense.'

(8) O'Riley was on trial for armed robbery. The jury came out and announced, "Not guilty."
"Wonderful," said O'Riley, "does that mean I can keep the money?"

Rather than simply allowing for a CONFLICT relation between punchline and set-up, Suls feels the need to establish a

CONFLICT with some PREDICTION: 'It is then predicted that he will say "Does that mean I can go now?"'. However, the argument that there is CONFLICT with some specific PREDICTION, and that this causes the humorous effect, is unsupported. There is no evidence that this remark rather than any other (e.g. "Thank you") is predicted, nor that the humour of the actual punchline depends on a deviation from some precise expectation. Suls may feel the need to follow this line of reasoning owing to his definition: 'Incongruity of the joke's ending refers to how much the punch line violates the recipient's expectations.'(p.92).

Shultz uses a similar definition to Suls, implying that this captures the intentions of Schopenhauer (1819), Kant (1892), Maier (1932) and Koestler (1970): 'Incongruity is usually defined as a conflict between what is expected and what actually occurs in the joke'(Shultz, 1976, p.12). (Notice that this is not identical to the gloss of "incongruity" given by Shultz elsewhere in that article – see quotation in section 4.1 above.) In this definition, the CONFLICT is with something fairly specific which is a PREDICTION from the set-up, but which did not occur. That is, there is an implied text whose form and interpretation is different from that of the actual text; for example, (9), or the much-quoted (10).

(9) Isn't modern technology wonderful? I remember the excitement when we were the first family in our street to have cordless pyjamas. (Arnold Brown, early 1990s)

(10) One more drink and I'll be under the host. (Dorothy Parker)

To give (9) an SD analysis, we can segment it into a set-up of *Isn't modern technology .... to have cordless* and a punchline of *pyjamas*. The CONFLICT is between the punchline and the PREDICTED interpretation (in which the text ends with *(tele)phones*).

It may therefore be necessary to consider two possible subtypes of the CONFLICT relation: one sets the punchline against the (more OBVIOUS meaning of the) set-up, and the other sets the punchline against a predicted interpretation (derived from the set-up).

PREDICTION can happen at various linguistic levels – semantic, syntactic, lexical, etc. Such issues are familiar within computational linguistics and psycholinguistics, perhaps most relevantly in the study of "garden path" sentences. It could be argued that in (9), a prediction occurs at a semantic level, manifested lexically; alternatively, there might be an explanation in terms of statistics or semantic priming.

4.3 What is an interpretation?

Underlying all of the concepts listed here is the more general question of what counts as an INTERPRETATION – is it just the overall meaning of the text, or is the way of expressing it relevant? That is, could a text which unexpectedly switched between two different ways of expressing the same overall meaning act as a joke? Example (11) is like (9) and (10), in that the punchline is the final word, and that the CONFLICT is with a predicted ending rather than with an already completed meaning.[1]

(11) Kevin Keegan isn't fit to lace George Best's drinks.

A further point of interest, however, is that the overall import of both phrasings is the same (Keegan is a vastly inferior player to Best); what differs is the idiomatic phrasing. A similar observation could be made about (12).

(12) One prostitute said to another, "Can you lend me ten dollars until I get back on my back?" (Suls, 1972)

So if jokes such as (11) and (12) are to be analysed within the SD framework, the notion of INTERPRETATION may have to be generalised, to encompass variations in linguistic expression, or at least to allow for a level prior to decoding of idioms.

Example (13) could be analysed in SD terms if we allow an INTERPRETATION to include the pragmatic status of an utterance.

(13) What's the difference between an elephant and a watermelon?
(I don't know.)
You'd be a fine one to send to the store for a watermelon. (Dienhart, 1999)

4.4 A sketch of the processing

In attempting to make the above discussion more precise, it is instructive to consider the facilities that a processing model would have to provide in order to interpret jokes in the manner suggested by IR theories in general, and the SD model in particular.

4.4.1 Analysing the set-up.

The processor must be capable of extracting a symbolic representation of the meaning of a multi-sentence text, and making at least some inferences from it. That is, it must be a fairly general text-understanding module. Where the text has more than one interpretation, it must be able to select the most OBVIOUS one (either by having this inherent in its workings, or by having it separately compute all readings and then select one).

[1]This joke is heavily dependent on cultural knowledge, namely: George Best was a brilliantly gifted football player whose success led to severe alcohol problems; Kevin Keegan was a very successful and prominent, but less skilful, footballer; there is within football an idiomatic phrase of condemnation 'X is not fit to lace Y's boots'. The line is due to a UK sports journalist in the 1970s.

4.4.2 Predicting.

The system must be able to reason ahead so as to predict at least the meaning of likely continuations of the text, and even sometimes predict the actual lexical items which are most probable.

4.4.3 Detecting the punchline.

Whether the punchline is one word, one sentence, or a sequence of sentences, the processor must not only be able to interpret it (an ability already required for the set-up), it must be able to construct a separate structure to compare with the meaning of the set-up. The text will contain no explicit markers to separate the punchline from the preceding text (except in special cases such as riddles), so this is not trivial. The distinction raised above between CONFLICT with the OBVIOUS interpretation of the set-up, and CONFLICT with a PREDICTED ending becomes relevant here. For the first of these, the punchline can be recognised, almost by definition, by the fact that it does not make sense (is not COMPATIBLE) with the current (most OBVIOUS) interpretation of the text so far. That is, there must be some form of assimilation mechanism which attempts to form a coherent interpretation of all the material in the text; when this mechanism finds a difficulty, then this may indicate that a punchline has occurred. On the other hand, where the CONFLICT is with predicted material, some form of matching is needed. This will necessitate the system having a firm hypothesis about which levels (semantic, syntactic, lexical) within the prediction are crucial, as there may be cases, for example, where the same semantic prediction can be expressed equally well in different syntactic and lexical forms. In examples such as (9), (10), (11) and (12), the prediction is for a particular lexical item, and the matching should be easier than those cases where an unexpected concept is conveyed by a whole sentence.

4.4.4 Processing the punchline.

Even if the punchline can be detected, it is not trivial to make sense of it. In the detection stage, above, there were two cases to consider depending on the type of CONFLICT involved, but here they both reduce to much the same requirements.

Where there has been no particular prediction, but the punchline has failed to be assimilated into the ongoing interpretation the task is to find an interpretation which does cover both set-up and punchline. If all possible interpretations have been computed as the text has been scanned, the goal is to find one into which the meaning of the punchline can be assimilated. If only one interpretation has been maintained, some form of backtracking is necessary, in which the processor must ignore the original interpretation (which initially was the most OBVIOUS one), and find an alternative which is COMPATIBLE with the punchline.

Where the CONFLICT is with a predicted interpretation (or text), the situation is similar. The fact that the punchline CONFLICTS with the predicted material suggests that the interpretation must be reconsidered (although see the remarks above about (11) and (12)). Again, the goal is to find an alternative interpretation which covers both set-up and punchline.

4.4.5 The humorous effect.

In the outline so far, the CONFLICT between punchline and set-up has been used only to detect the punchline. Some authors (see quotations earlier from Suls and from Shultz) appear to propose that this is also the factor determining the incongruity, and hence the humour. Other accounts emphasise the role of what is labelled here COMPARISON (e.g. Raskin (1985)) or INAPPROPRIATENESS (e.g. Freud (1966)). Some of these apparent differences of opinion may be more an artefact of the lack of precise terminology or definitions, in that the CONFLICT between punchline and set-up may be conflated with, or confused with, COMPARISON between the revised interpretation and the initial interpretation. Whichever of these proposals is adopted, it means that the processor must make some assessment of the extent to which these factors (COMPARISON, INAPPROPRIATENESS, and perhaps CONFLICT) are present. This will not be trivial.

4.5 Building on the SD model

What does this dissection of the surprise disambiguation variant of incongruity-resolution theory tell us? Consideration of the above steps, particularly the final stage (subsection 4.4.5) shows that virtually all of the defining properties of the SD model relate to the *delivery* of the humorous content rather than the actual detection, comprehension or appreciation of humour. All the steps in the processing, except for the last, would equally well describe a simple misunderstanding, and not all misunderstandings are humorous. In other words, the SD account does not particularly illuminate one of the central concerns of IR theory (and of much theorising about humour): what kind of incongruity is funny? What the model does do is strip away some of the more peripheral aspects about how this particular genre of joke *conveys* the incongruity. In this way, it reduces the research problem to a set of component sub-problems. Most of the processing stages outlined above depend upon knowledge about language and (via linguistic meaning) the world, and research into language processing will illuminate these parts of the model. The subproblems are:

OBVIOUSNESS: What makes one potential interpretation more obvious than another?

CONFLICT/COMPATIBILITY: How difficult to assimilate must a piece of text be in order to stimulate a search for another interpretation? How can this search be guided by the portion of text that caused the re-assessment?

COMPARISON: What does it mean for two interpretations to differ in an amusing way (as opposed to merely not being the same)?

INAPPROPRIATENESS: What factors make an interpretation inherently more amusing?

Overall**:** What combinations of these factors combine to produce humour?

Of these, those concerning OBVIOUSNESS and CONFLICT (or COMPATIBILITY) are to some extent more general research issues in language processing, and the simplest assumption is to posit as few special mechanisms as possible beyond those needed for normal language comprehension. However, jokes may (in common with stories) require a certain semantic or pragmatic licence, so that the audience can "suspend disbelief": jokes can feature entities which would qualify as semantically ill-formed in some descriptions of the world, such as talking kangaroos or walking cauliflowers. This freedom could complicate the detection of punchline-induced CONFLICT.

Definitions of several of the necessary concepts might be forthcoming from other disciplines, such as psycholinguistic research into ambiguity and ambiguity resolution (OBVIOUSNESS); research into the interpretation and assimilation of sentences within a discourse (CONFLICT, COMPATIBILITY); semantic research, studies of belief systems and of social attitudes (COMPARISON, INAPPROPRIATENESS); sociology, psychology (INAPPROPRIATENESS); psycholinguistic and computational research into parsing and semantic interpretation during sentence processing (PREDICTION); semantics (INTERPRETATION). If these subsidiary concepts could be defined independently of humour theory (the *theory-external* approach) then the SD model of humour would reduce to the last question listed above ('Overall'). On the other hand, if it is necessary to have specific definitions of these concepts (for example, if only certain types of COMPARISON result in humour), then the humour-theory content of the SD would also include these *theory-internal* definitions.

The extent to which the hidden interpretation is INAPPROPRIATE (or has other suitable properties) may vary, as may the strength of COMPARISON between the interpretations. It is also conceivable that the sharpness of the CONFLICT between punchline and the first interpretation may vary, as may the OBVIOUSNESS of that interpretation. Hence whatever formalisation is developed may have to be capable of expressing degrees of some properties. A more detailed version of the SD model may involve a complex disjunction of conditions, where each condition is itself a specification of thresholds for certain properties. For example, maybe a text is humorous if either the COMPARISON yields a value above a certain threshold or the INAPPROPRIATENESS of the hidden interpretation reaches some other minimum level.

The last three questions in the list above are central to humour theory, and the solution of these would be a major step forward.

This is not to say that the processing arrangement of the SD model does not contribute to our understanding of humour. Many authors have agreed that the humorous effect usually depends upon the incongruity being brought to the attention of the audience abruptly, and the SD mechanism is one way to do this. Hence, it offers one solution to that particular sub-problem within a theory of humour.

5 Suls' two-stage model

5.1 Overview of processing

As noted briefly in Section 3 above, Suls (1972) outlines an IR-style processing model, which could be summarised follows (see also Figure 1):

- as text is read, make predictions
- while no conflict with predictions, keep going
- if input conflicts with predictions:
  - if not ending - PUZZLEMENT
  - if it is the ending, try to resolve:
    - no rule found - PUZZLEMENT
    - cognitive rule found - HUMOUR

This differs from the SD model in that Suls does not demand that any ambiguity be present in the set-up, which allows him to cover examples such as (5), (8) and (14), where it is hard to argue for any ambiguity in the set-up.

(14) Fat Ethel sat down at the lunch counter and ordered a whole fruit cake. "Shall I cut it into four or eight pieces?" asked the waitress. "Four," said Ethel, "I'm on a diet."

On the other hand, in examples where there is clear use of ambiguity, such as (7), (6), (15) or (16), Suls would have to postulate a cognitive rule which simulates the effect of re-interpreting the set-up.[2]

(15) Mr Fields, do you believe in clubs for young people?
Only when kindness fails.
(Shultz, 1976) and elsewhere

(16) Can you tell me how long cows should be milked?
They should be milked the same as short ones.
(Shultz, 1976)

[2] There is a puzzling passage on p.96, where Suls, discussing work by Kagan (1967) with infants, seems to imply that his model differs from Kagan's in that Suls' system does indeed reinterpret the set-up, contrary to both his diagram – Figure 1 here – and the implications of his examples.

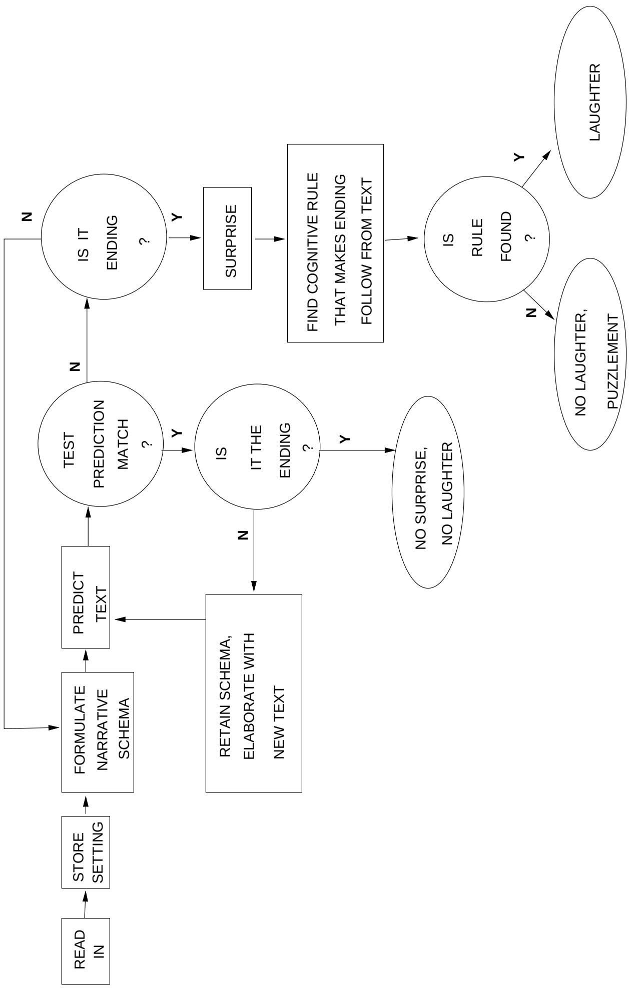

Figure 1: The Suls Two-Stage Model

Suls gives only a broad outline of his model, not much more detailed than the summary given above of the SD model, and his use of the term 'incongruity' is not very precise, although it seems to correspond, in the informal terminology of Section 4, to CONFLICT:

> 'the perceiver finds his expectations about the text disconfirmed by the ending of the joke....In other words, the recipient encounters an incongruity — the punch line.'(p.82)
> 'humor derives from experiencing a sudden incongruity which is then made congruous'.(p.82)
> 'Incongruity of the joke's ending refers to how much the punch line violates the recipient's expectations.' (p.92)

5.2 Building on the "Two-Stage" model

Suls quotes approvingly from authors who observe that perception (including language understanding) is an active process, so presumably he would accept that even in a smooth-flowing congruous text, the perceiver may be using cognitive rules to assimilate each sentence in turn. A cognitive rule is, after all, a very general construct: 'a logical proposition, a definition, or a fact of experience' (p.82). However, Suls gives no indication of what would push a potential punchline over the threshold between normal rule-based comprehension into incongruity-driven reasoning. Moreover, he gives no discussion of the distinction between a punchline simply not making sense (a misunderstanding) and being humorously incongruous. Since he is not postulating an ambiguous set-up, the notions called (in Section 4 above) COMPARISON and INAPPROPRIATENESS do not enter into the debate.

Suls also does not explore the possibility that the nature of the resolution (i.e. the content of the cognitive rule) may contribute to the humour, although he suggests that the complexity of the reasoning needed may be a factor. His claim is that it is both necessary and sufficient for humour if there is an incongruous punchline which can make sense using a suitable cognitive rule.

Like the SD model outlined earlier, this model leaves the hardest parts still to be done, and is more of a sketch of a "delivery" mechanism than an explanation of incongruity in humour.

The residual questions from Suls' framework are:

(a) How surprising must a portion of text be to count as a punchline? (Cf. discussion in Section 4.5 above)

(b) What is a 'cognitive rule'?

(c) What types of cognitive rule give rise to a humorous resolution rather than merely the patching up of a misunderstanding?

The last of these questions could be seen as being the central issue of much of humour research. It is quite common for writers on humour to observe that jokes often rely on a 'local logic' or 'para-logic' (Attardo, 1994, Section 4.0.2) – some distorted form of reasoning which is close to, or analogous with, sound reasoning, but which is sufficiently bizarre to produce humour. Inspecting Suls' examples (particularly (8) and (14) above), it is plausible that the humorous effect relies heavily on some faulty form of reasoning. This accords with the proposal by Attardo (1997) that the 'logical mechanism' of the 'General Theory of Verbal Humour' (Attardo and Raskin, 1991) can be regarded as the *resolution* within an IR-style model.

6 Comparing the two models

A superficial reading of articles on the two variants discussed above gives the impression that they are extremely similar, but closer inspection shows interesting differences.

(a) The two models cover, or attempt to cover, different subclasses of joke. The SD model requires an ambiguous set-up, whereas the two-stage model makes no mention of ambiguity.

(b) The SD model decomposes the humorous effect into slightly simpler concepts (particularly COMPARISON, ABSURDITY, and TABOO), and so at least starts to address the issue of "incongruity".

(c) The two-stage model relies on some (undefined) form of "humorous logic", and so leaves the difficult problem of "incongruity" relatively untouched.

As noted earlier, there are further jokes which have been postulated to be incongruity-resolution jokes (e.g. (2), (3)), but which are not covered by either of these approaches. (Although (2) involves ambiguity, the ambiguity is not in the set-up). That is, what we have here are two subclasses of IR humour, leaving others to be explored.

Acknowledgements

This paper was written while the author was a Visiting Fellow at Microsoft Research Institute, Division of Information and Communication Sciences, Macquarie University, Sydney, Australia.

References

Salvatore Attardo. *Linguistic Theories of Humour*. Mouton de Gruyter, Berlin, 1994.

Salvatore Attardo. The semantic foundations of cognitive theories of humor. *HUMOR*, 4(10):395–420, 1997.

Salvatore Attardo and Victor Raskin. Script theory revis(it)ed: joke similarity and joke representation model. *HUMOR*, 4(3):293–347, 1991.

James Beattie. An essay on laughter, and ludicrous composition. In *Essays*. William Creech, Edinburgh, 1776. Reprinted by Garland, New York, 1971.

Anthony J. Chapman and Hugh C. Foot, editors. *Humour and Laughter : Theory, Research and Applications*. Transaction Publishers, London, first edition, 1976.

Anthony J. Chapman and Hugh C. Foot, editors. *It's a Funny Thing, Humour*. Pergamon Press, Oxford, 1977.

Paul De Palma and E. Judith Weiner. Riddles: Accessibility and knowledge representation. In *Proceedings of COLING-92*, pages 1121–1125, 1992.

John M. Dienhart. A linguistic look at riddles. *Journal of Pragmatics*, 31(1):95–125, 1999.

Sigmund Freud. *Jokes and their relation to the unconscious*. Routledge & Kegan Paul, London, 1966. First published 1905.

Rachel Giora. On the cognitive aspects of the joke. *Journal of Pragmatics*, 16(5):465–485, 1991.

Michael Godkewitsch. Physiological and Verbal Indices of Arousal in Rated Humour. In Chapman and Foot (1976), chapter 6, pages 117–138.

Jeffrey H. Goldstein and Paul E. McGhee, editors. *The Psychology of Humor*. Academic Press, New York, 1972.

J. Kagan. On the need for relativism. *American Psychologist*, 22:131–147, 1967.

Immanuel Kant. Excerpt from "Critique of Judgement". In Morreall (1987), chapter 8, pages 45–50.

Bruce Katz. A neural resolution of the incongruity and incongruity-resolution theories of humour. *Connection Science*, 5:59–75, 1993.

Patricia Keith-Speigel. Early conceptions of humor: Varieties and issues. In Goldstein and McGhee (1972), pages 3–39.

Arthur Koestler. *The Act of Creation*. Pan Books, London, 1970. First published 1964, Hutchinson & Co.

N. R. F. Maier. A gestalt theory of humour. *British Journal of Psychology*, 23:69–74, 1932.

Paul E. McGhee and Jeffrey H. Goldstein, editors. *Handbook of Humor Research*. Springer-Verlag, New York, 1983.

Marvin Minsky. Jokes and the logic of the cognitive unconscious. AI Memo 603, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Cambridge, Mass., 1980.

John Morreall, editor. *The Philosophy of Laughter and Humor*. SUNY Press, Albany, NY, 1987.

John Allen Paulos. *Mathematics and Humor*. University of Chicago Press, Chicago and London, 1980.

William J. Pepicello and Robert W. Weisberg. Linguistics and humor. In McGhee and Goldstein (1983), pages 59–83.

Victor Raskin. *Semantic Mechanisms of Humour*. Reidel, Dordrecht, 1985.

Mary K. Rothbart. Psychological approaches to the study of humour. In Chapman and Foot (1977), pages 87–94.

Mary K. Rothbart and Diana Pien. Elephants and marshmallows: A theoretical synthesis of incongruity- resolution and arousal theories of humour. In Chapman and Foot (1977), pages 37–40.

Arthur Schopenhauer. Excerpt from "The World as Will and Idea". In Morreall (1987), chapter 9, pages 51–64.

Thomas R. Shultz. Development of the appreciation of riddles. *Child Development*, 45:100–105, 1974.

Thomas R. Shultz. A Cognitive-Developmental Analysis of Humour. In Chapman and Foot (1976), chapter 1, pages 11–36.

Jerry M. Suls. A two-stage model for the appreciation of jokes and cartoons: an information-processing analysis. In Goldstein and McGhee (1972), chapter 4, pages 81–100.

Jerry M. Suls. Cognitive and disparagement theories of humour: A theoretical and empirical synthesis. In Chapman and Foot (1977), pages 41–45.