Computational Humour

Carlo Strapparava, Oliviero Stock, and Rada Mihalcea

Abstract Computational humour is a challenge with connections and implications in many artificial intelligence areas, including natural language processing, intelligent human–computer interaction, and reasoning, as well as in other fields such as cognitive science, linguistics, and psychology. Of particular interest is its connection to emotions. In this chapter we overview the basic theories of humour and present the main contributions made in the field of computational verbal humour, including applications for automatic humour generation and humour recognition.

1 Introduction

The interaction between humans and computers needs to evolve beyond usability and productivity. There is an agreement in the field of human–computer interaction that the future stands in themes such as entertainment, fun, emotions, aesthetic pleasure, motivation, attention, engagement. Humour is an essential element in communication: it is strictly related to the themes mentioned above, and arguably humans cannot survive without it. While it is generally considered as merely a way to induce amusement, humour provides an important way to influence the mental state of people to improve their activity. Even though humour is a complex capability to reproduce, it is realistic to model some types of humour production and to aim at implementing this capability in computational systems.

Humour is a powerful generator of emotions. As such, it has an impact on people's psychological state, directs their attention, influences the processes of memorization and of decision-making, and creates desires and emotions. Actually, emotions are an extraordinary instrument for motivation and persuasion because those who are capable of transmitting and evoking them have the power to influence other people's opinions and behaviour. Humour, therefore, allows for conscious and

C. Strapparava (✉)
Fondazione Bruno Kessler-Irst, Povo, Trento, Italy
e-mail: strappa@fbk.eu

P. Petta et al. (eds.), *Emotion-Oriented Systems*, Cognitive Technologies,
DOI 10.1007/978-3-642-15184-2_31,

constructive use of the affective states generated by it. Affective induction through verbal language is particularly interesting; and humour is one of the most effective ways of achieving it. Purposeful use of humourous techniques enables us to induce positive emotions and mood and to exploit their cognitive and behavioural effects. For example, the persuasive effect of humour and emotions is well known and widely employed in advertising. Advertisements have to be both short and meaningful, to be able to convey information and emotions at the same time.

Humour acts not only upon emotions but also on human beliefs. A joke plays on the beliefs and expectations of the hearer. By infringing on them, it causes surprise and then hilarity. Jesting with beliefs and opinions, humour induces irony and accustoms people not to take themselves too seriously. Sometimes simple wit can sweep away a negative outlook that places limits on people's desires and abilities. Wit can help people overcome self-concern and pessimism that often prevents them from pursuing more ambitious goals and objectives.

Humour encourages creativity as well. The change of perspective caused by humourous situations induces new ways of interpreting the same event. By stripping away clichés and commonplaces, and stressing their inconsistency, people become more open to new ideas and points of view. Creativity redraws the space of possibilities and delivers unexpected solutions to problems. Actually, creative stimuli constitute one of the most effective impulses for human activity. Machines equipped with humourous capabilities will be able to play an active role in inducing users' emotions and beliefs and in providing motivational support.

There are many practical settings where computational humour adds value. Among them there are business world applications (such as advertisement, e-commerce), general computer-mediated communication and human–computer interaction, increase in the friendliness of natural language interfaces, educational and edutainment systems.

There are also important prospects for humour in automatic information presentation. In the Web age, presentations will become more and more flexible and personalized and will require humour contributions for electronic commerce developments (e.g. product promotion, getting selective attention, help in memorizing names) more or less as it happened in the world of advertisement within the old broadcast communication.

In this chapter we focus mainly on "verbal humour", which is the most tangible and perhaps the most widely researched form of humour. Although other forms of humour (e.g. visual or situational) have also received attention from the research community, we concentrate our work and consequently this survey on the linguistic expressions of humour.

The chapter is structured as follows. Section 2 briefly surveys the main theories of humour in philosophy, psychology, and linguistics. Sect. 3 summarizes the main research attempts in computational humour. In Sect. 4 we illustrate an example of humour generation system (HAHAcronym), while in Sect. 5 we describe an approach to deal with humour recognition. Finally, Sect. 6 concludes the chapter with some prospects on this field.

2 Background in Humour Research

In this section, we summarize the main theories of humour that emerged from philosophical and modern psychological research and survey the past and present developments in the fields of theoretical and computational linguistics. We also briefly overview related research work in the fields of psychology, sociology, and neuroscience.

2.1 Theories of Humour

There are three main theories of humour, which emerged primarily from philosophical studies and research in psychology.

2.1.1 Incongruity Theory

The incongruity theory suggests that humour is due to the mixing of two disparate interpretation frames in one statement. One of the earliest references to an incongruity theory of humour is perhaps due to Aristotle (350 BC) who found that the contrast between expectation and actual outcome is often a source of humour. He is also making a distinction between surprise and incongruity, where the later is presumed to have a resolution that was initially hidden from the audience. The incongruity theory has also found a supporter in Schopenhauer (1819), who emphasizes the element of surprise by suggesting that "the greater and more unexpected [...] the incongruity is, the more violent will be [the] laughter". The incongruity theory has been formalized as a necessary condition for humour and used as a basis for the Semantic Script-based Theory of Humour (SSTH) (Raskin, 1985) and later on the General Theory of Verbal Humour (GTVH) (Attardo and Raskin, 1991).

2.1.2 Superiority Theory

The superiority theory argues that humour is a form of expressing the superiority of one over another. As suggested by Hobbes (1840), laughter is "nothing else but sudden glory" triggered by a feeling of superiority with respect to others or with respect to ourselves in a previous moment. A closely related theory is the one supported by Solomon (2002), who suggests that humour is due to feelings of inferiority, which led to the so-called inferiority theory. Although the superiority and inferiority theories of humour have been typically perceived as diametrically opposed, they are in fact intimately related, as the "superior"/"inferior" distinctions are often due to a different point of view. In fact, it can be argued that laughter is triggered by our feelings of superiority with respect to others or ourselves in a previous moment, which are equivalent to feelings of inferiority felt by others or by ourselves in a past moment.

2.1.3 Relief Theory

The third major theory is the *relief theory*, which suggests that humour is a form of bypassing certain censors that prevent us from having "prohibited thoughts". Eluding these censors results in a release of the energy inhibited by these censors and consequently the feeling of relief. One of the strongest supporters of the relief theory is Freud (1905), who draws a connection between jokes and the unconscious, and Spencer (1860), who suggests that laughter is a form of "nervous energy". Some of these ideas have been later embraced by Minsky in his theory of humour (Minsky, 1980), to which he adds a cognitive element that attempts to explain the "faulty logic" that is typically encountered in jokes, which is normally suppressed in order to avoid "cognitive harm".

2.2 Linguistics Research on Humour

A significant fraction of the research on humour that has been carried out to date has concentrated on the linguistic characteristics of humour. Among the linguistic theories, the most influential is perhaps the General Theory of Verbal Humour (GTVH) (Attardo and Raskin, 1991), which is an extension of the earlier Semantic Script-based Theory of Humour (SSTH) (Raskin, 1985).

2.2.1 Semantic Script-Based Theory of Humour

SSTH is based on the representation of jokes as *script opposition*, which is an idea closely related to the incongruity resolution theory. Briefly, SSTH defines the structure of a joke as consisting of a *set-up* and a *punchline*. The set-up has at least two possible interpretations out of which only one is obvious, and consequently the humourous effect is created by the punchline which triggers the second less obvious interpretation in a surprising way.

The central hypothesis in SSTH is that a text is humourous if the following two conditions are satisfied. First, the humourous text has to be compatible with at least two different interpretations (scripts). And second, the two interpretations have to be opposed to each other. For instance, the following example taken from Raskin (1985) illustrates this theory: "The first thing that strikes a stranger in New York is a big car." The set-up has two possible interpretations: strike as in "impress" or "hit", which are opposed to each other ("impress" being a positive action and "hit" triggering negative feelings). The first interpretation is more obvious and thus initially preferred. However, the punchline "by a big car" will change the preference to the second interpretation, which generates the humourous effect.

According to SSTH, the opposition between scripts is binary and can fall into one of the following three generic types: actual/non-actual, normal/abnormal, possible/impossible, which in turn can be broken down into more specific oppositions, such as positive/negative or good/bad.

2.2.2 General Theory of Verbal Humour

Following SSTH, the GTVH (Attardo and Raskin, 1991) extends the script opposition theory and adds other possible knowledge resources for a humourous text. While SSTH was primarily focused on semantics, the GTVH is more general and includes other areas in linguistics such as pragmatics and style. GTVH defines six main knowledge resources that can be organized on six levels from concrete (low level) to abstract (high level).

- *Script opposition*, which is a knowledge source based on the main idea of SSTH of opposing interpretations that are both compatible with the text.
- *Logical mechanism*, which provides a possible resolution mechanism for the incongruity between scripts.
- *Situation*, which defines the context of the joke in terms of location, participants, and others.
- *Target*, which is the person or group of persons that are targeted by the joke.
- *Narrative strategy*, which defines the style of the joke, i.e. whether it is a dialogue, a riddle, or a simple narrative.
- *Language*, which defines the "surface" of the joke in terms of linguistic aspects such as lexicon, morphology, syntax, semantics.

For example, Attardo and Raskin (1991) exemplify the knowledge resources using the following joke: "How many Poles does it take to screw in a light bulb? Five. One to hold the light bulb and four to turn the table he is standing on." The script opposition is formed between the expected normal behaviour of a person when screwing in a light bulb and the "dumb" resolution proposed by the punchline; the logical mechanism is that of "reversal" of a normal behaviour; the situation is "bulb changing"; the target of the joke are the "Poles"; and finally, the narrative structure is a "riddle" (Ritchie, 2003).

An interesting experiment centred around the GTVH theory is reported by Ruch et al. (1993), where three jokes are transformed into variants that differed from the original joke in one of the GTVH parameters. A group of 500 subjects were asked to rate the similarity between each of the variants and the original joke on a scale of 1–4. The findings indicate that higher similarity is observed for those variants that differ in a low-level parameter in the GTVH hierarchy, thus suggesting that the higher level parameters such as script opposition and logic mechanism are more humour related (Ritchie, 2003).

While GTVH is perhaps the most extensive linguistic theory of humour that has been proposed to date, it has been criticized by Ritchie (2003) as lacking theoretical grounds. Ritchie raises doubts about the falsifiability of the GTVH and about the lack of systematic examples where some of the GTVH knowledge resources are missing, thus resulting in a lack of humourous effect, along with humourous examples that include the missing knowledge resources.

2.2.3 Related Work in Linguistics

Besides the SSTH and the GTVH theories, other research work in linguistics has focused mainly on the analysis of the lexical devices used in humourous text. The syntactic ambiguity often encountered in humour is analysed by Hetzron (1991), who describes the structure of jokes and punchlines and analyses the logical devices found in verbal humour. Oaks (1994) proposes an interesting account on syntactic ambiguity in humour and identifies several ambiguity "enablers". He focuses mainly on part-of-speech ambiguity and identifies verbs, articles, and other parts-of-speech that can introduce ambiguity in language (e.g. *bite* that can be either a verb or a noun).

The lexical and syntactic ambiguity as a source of humour is also studied by Bucaria (2004), who analyses the linguistic ambiguity in newspaper headlines. She identifies three main types of ambiguity: lexical (e.g. "Actor sent to jail for not finishing *sentence*."), syntactic (e.g. "Eye *drops* off shelf"), and phonological (e.g. "Is there a ring of debris around *Uranus*"). She also identifies two additional schemata for humourous ambiguity, including the disjunctor/connector model (e.g. "New study on *obesity* looks for *larger* group.") and the double ambiguity model (e.g. "Farmer *Bill dies* in house."). In an analysis of 135 headlines, the lexical and syntactic forms of ambiguity were found to be dominant (71 lexical and 63 syntactic), covering a significant fraction of the corpus, and thus providing support for the incongruity theory of humour.

2.3 Multidisciplinary Research on Humour

In addition to the research work in linguistics humour has been also studied in other areas, e.g. sociology, neuroscience, and last but not least recent efforts in computational linguistics.

2.3.1 Sociology

In sociology, humour has been frequently associated with studies concerned with patterns of communication in different groups. For instance, Duncan (1984) shows that cohesive and non-cohesive work groups have different humour patterns, suggesting a correlation between the type of humour practiced in a group and the structure of the group.

Studies have also investigated the association between gender and humour, by analysing the type and role of humour for female, male, and mixed groups. Hay (1995) used a taxonomy of humour in a gender-oriented analysis, which revealed the preference of women for observational humour and the tendency of male groups for insults and role play. Interestingly, a correlation was also observed between the gender of these groups and the function of humour; women groups used humour primarily as a social element, whereas men groups often used it as a means for increasing status. Finally, Hay's study also reported on the association between gender and humour topics, suggesting that women use more frequently humour on topics involving people, while men joke more about politics, computers, and work;

this observation correlates with recent conclusions drawn in corpus-based gender studies (Liu and Mihalcea, 2007).

Another aspect of interest in sociology is the relation between culture and humour. Work in this area has highlighted the relation between cultural background and humour appreciation, showing that the set of values and norms of a culture largely determine the content and style of humour (Hertzler, 1970). Focused studies have highlighted differences between various cultures, as for instance the study reported by Nevo (1984), which shows how Arab and Jewish communities developed a different sense of humour explained by their diverse background and different social status.

2.3.2 Psychology

Humour research in psychology has been mainly concerned with the correlation between humour and individual development. There are several studies that considered the cognitive aspects of humour and the role that humour can play in infants and children development. For instance, it has been found that humour has an important role in improving text comprehension (Yuill, 1997).

Other studies have been concerned with the relation between personality profiles and sense of humour. Along these lines, it has been suggested that extroversion and neuroticism can be predicted from humour perception (Mobbs et al., 2005). Similarly, humour was found to be related to other personality characteristics such as simplicity–complexity, intelligence, or mood (Ruch, 1998).

2.3.3 Neuroscience

In recent years, given the advances made in brain imaging techniques (fMRI or MEG), researchers have started to investigate the brain activity observed during humour detection and comprehension. Recent research findings suggest that the left and the right hemispheres are both involved in humour appreciation, which is an effect that has been observed in verbal humour as well as visual humour (Bartolo et al., 2006). Moreover, studies have also observed the activation of the amygdala and midbrain regions (also known as the "pleasure centre"), which is probably due to the pleasurable effect created by humour (Watson et al., 2007).

It is also worth noting the study reported in Mobbs et al., (2005), which shows connections between gender, personality (i.e. extroversion and neuroticism), and humour appreciation, observed using brain imaging techniques. Such associations have been typically identified through surveys conducted in psychological studies, and the study reported in Mobbs et al., (2005) confirms these previous findings by identifying patterns of brain activity occurring during humour comprehension.

3 Computational Humour: State of the Art

While humour is relatively well studied in fields such as linguistics (Attardo, 1994) and psychology (Freud, 1905; Ruch, 2002), to date only a limited number of research contributions have been made towards the construction of computational humour prototypes. Most of the computational approaches to date on style classification have focused on the categorization of more traditional literature genres, such as fiction, sci-tech, legal, and others (Kessler et al., 1997), and much less on creative writings such as humour.

The most systematic effort in this area is perhaps Ritchie's book on the linguistic analysis of jokes, which brings together research on linguistic theories and artificial intelligence. In addition to a comprehensive overview of the main research contributions in humour, Ritchie is also proposing a classification of jokes into propositional and linguistic and suggests a structural description of the jokes (Ritchie, 2003).

There are two main research directions in computational humour: (1) *humour generation*, which attempts to build computational models to generate humourous text, and (2) *humour recognition*, which deals with the problem of identifying humour in natural language.

3.1 Humour Generation

One of the first attempts in humour generation is the work described by Binsted and Ritchie (1997), where a formal model of semantic and syntactic regularities was devised, underlying some of the simplest types of puns (*punning riddles*). The model was then exploited in a system called JAPE that was able to automatically generate amusing puns. A punning riddle is a question–answer riddle that uses phonological ambiguity. The three main strategies used to create phonological ambiguity are syllable substitution, word substitution, and metathesis. Their system generates punning riddles from a fixed linguistic model of pun schemata. An example: "What do you call a murderer with fiber?" *A cereal killer*.

Tinholt and Nijholt (2007) describe a first attempt at automatically generating jokes based on cross-reference ambiguity. The idea is that when a given cross-reference ambiguity results in script opposition it is possible to generate a punchline based on this ambiguity. An example of dialogue is "User: Did you know that the cops arrested the demonstrators because they were violent?" "System: The cops were violent? Or the demonstrators? :)"

Another humour generation project is the HAHAcronym project (Stock and Strapparava, 2003), whose goal was to develop a system able to automatically generate humourous versions of existing acronyms or to produce a new amusing acronym constrained to be a valid vocabulary word, starting with concepts provided by the user. The comic effect was achieved mainly by exploiting incongruity (e.g. finding a religious variation for a technical acronym). We describe in detail this system in Sect. 4.

3.2 Humour Recognition

There are only a few studies addressing the problem of humour recognition. The study reported in Taylor and Mazlack (2004) is devoted to the problem of humour comprehension, focusing on a restricted type of wordplays, namely the

"Knock-Knock" jokes. The goal of the study was to evaluate to what extent wordplay can be automatically identified in "Knock-Knock" jokes and if such jokes can be reliably recognized from other non-humourous text. The algorithm was based on automatically extracted structural patterns and on heuristics heavily based on the peculiar structure of this particular type of jokes. While the generic wordplay recognition gave satisfactory results (67% accuracy), the identification of wordplays that had a humourous effect turned out to be significantly more difficult (12% accuracy).

In our own previous work (Mihalcea and Strapparava, 2005b; Mihalcea and Pulman, 2007), humour recognition was formulated as a text classification task, and machine learning algorithms were run on large collections of humourous texts (oneliners or humourous news articles). Both content and stylistic features were evaluated, including n-gram models, alliteration, antonymy, and adult slang, with performance figures significantly higher than a priori known baselines. We will describe in detail the used methodology in Sect. 5.

Another humour-recognition study was reported by Purandare and Litman (2006), where the recognition experiments were performed using both content features and spoken dialogue prosody features (tempo, energy, and pitch). The experiments were run on dialogues from the TV series "Friends", with significant improvements observed over the baseline. They also reported a gender study, with the improvement obtained for humour recognition in male dialogues being higher than the one obtained for female dialogues, suggesting perhaps that the humourous features are more prominent for males than for females.

4 Humour Generation: HAHAcronym

HAHAcronym was the first European project devoted to computational humour.[1] The main goal of HAHAcronym was the realization of an acronym ironic re-analyser and generator as a proof of concept in a focalized but non-restricted context. In the first case the system makes fun of existing acronyms; in the second case, starting from concepts provided by the user, it produces new acronyms, constrained to be words of the given language. And, of course, they have to be funny.

The realization of this system was proposed to the European Commission as a project that we would be able to develop in a short period of time (less than a year), that would be meaningful and well demonstrable, that could be evaluated along some pre-decided criteria, and that was conducive to a subsequent development in a direction of potential applicative interest. So for us it was essential that

1. the work could have many components of a larger system, simplified for the current setting;
2. we could reuse and adapt existing relevant linguistic resources;
3. some simple strategies for humour effects could be experimented.

[1]EU project IST-2000-30039 (partners: ITC-irst and University of Twente), part of the Future Emerging Technologies section of the Fifth European Framework Program.

One of the purposes of the project was to show that using "standard" resources (with some extensions and modifications) and suitable linguistic theories of humour (i.e. developing specific algorithms that implement or elaborate theories), it is possible to implement a working prototype.

4.1 Resources

In order to realize the HAHAcronym prototype (see Fig. 1), we refined existing resources and developed general tools useful for humourous systems. A fundamental tool is an incongruity detector/generator that makes the system able to detect semantic mismatches between word meaning and sentence meaning (i.e. in our case the acronym and its context). For all tools, particular attention was put on reusability.

The starting point consisted in making use of some standard resources, such as WORDNET DOMAINS (Magnini et al., 2002) (an extension of the well-known English WORDNET) and standard parsing techniques.

Wordnet. WORDNET is a thesaurus for the English language inspired by psycholinguistics principles and developed at the Princeton University by George Miller (Fellbaum, 1998). Lemmata (about 130,000 for version 1.6) are organized in synonym classes (about 100,000 *synsets*). A synset contains all the words by means of which it is possible to express a particular meaning: for example, the synset `knight, horse` describes the sense of "horse" as a chessman. The main relations present in WORDNET are *synonymy*, *antonymy*, *hyperonymy–hyponymy*, *meronymy–holonymy*, *entailment*, *troponymy*.

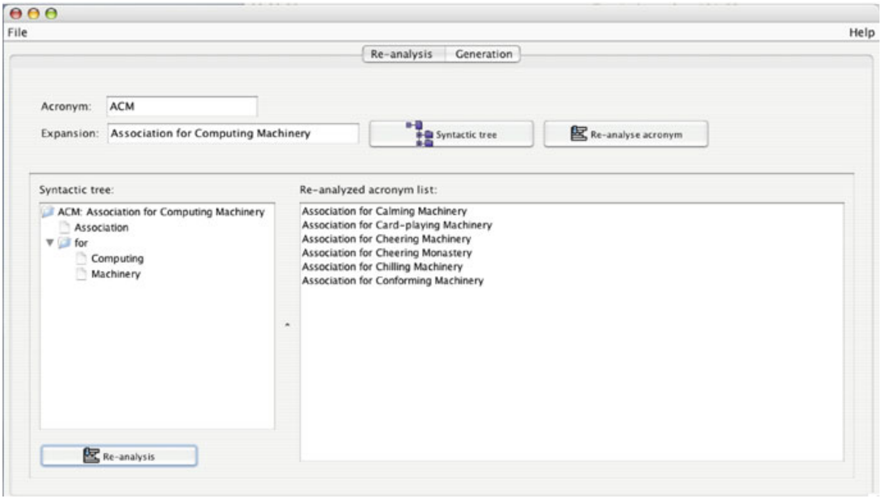

Fig. 1 A screenshot of a reanalysis in HAHAcronym

Wordnet Domains. Domains have been used both in linguistics (i.e. Semantic Fields) and in lexicography (i.e. Subject Field Codes) to mark technical usages of words. Although this is useful information for sense discrimination, in dictionaries it is typically used for a small portion of the lexicon. WORDNET DOMAINS[2] is an attempt to extend the coverage of domain labels within an already existing lexical database, WORDNET. The synsets have been annotated with at least one domain label, selected from a set of about 200 labels hierarchically organized.

The 250 domain labels are organized in a hierarchy (exploiting Dewey Decimal Classification), where each level is made up of codes of the same degree of specificity: for example, the second level includes domain labels such as BOTANY, LINGUISTICS, HISTORY, SPORT, and RELIGION, while at the third level we can find specialization such as AMERICAN_HISTORY, GRAMMAR, PHONETICS, and TENNIS.

Opposition of Semantic Fields. On the basis of well-recognized properties of humour accounted for in many theories (e.g. incongruity, semantic field opposition, apparent contradiction, absurdity) an independent structure of domain opposition was modelled, such as RELIGION VS. TECHNOLOGY, SEX VS. RELIGION. Opposition is exploited as a basic resource for the incongruity generator.

Adjectives and Antonymy Relations. Adjectives play an important role in modifying and generating funny acronyms. WORDNET divides adjectives into two categories. *Descriptive adjectives* (e.g. `big, beautiful, interesting, possible, married`) constitute by far the largest category. The second category is called simply *relational adjectives* because they are related by derivation to nouns (i.e. `electrical` in `electrical engineering` is related to the noun `electricity`). To relational adjectives, strictly dependent on noun meanings, it is often possible to apply similar strategies as those exploited for nouns. Their semantic organization, though, is entirely different from that of the other major categories. In fact it is not clear what it would mean to say that one adjective "is a kind of" (ISA) some other adjective. The basic semantic relation among descriptive adjectives is antonymy. WORDNET proposes also that this kind of adjectives is organized in clusters of synsets associated by semantic similarity with a focal adjective. Figure 2 shows clusters of adjectives around the direct antonyms *fast/slow*.

Exploiting the Hierarchy. It is possible to exploit the network of lexical and semantic relations built in WORDNET to make simple ontological reasoning. For example, if a noun or an adjective has a geographic location meaning, the pertaining country and continent can be inferred.

Rhymes. The HAHAcronym prototype takes into account word rhymes and the rhythm of the acronym expansion. To cope with this aspect the CMU Pronouncing Dictionary[3] was organized with a suitable indexing. The CMU Pronouncing Dictionary is a machine-readable pronunciation dictionary for North American English that contains over 125,000 words and their transcriptions.

[2] It is freely available for research purposes at http://wndomains.itc.it (visited 30 May 2010).

[3] Available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict (visited 30 May 2010).

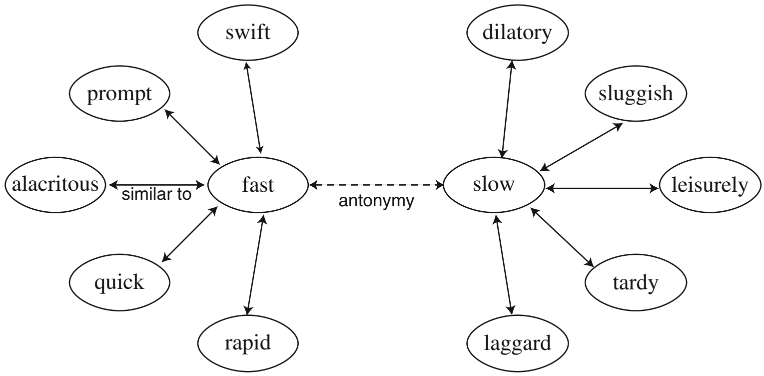

Fig. 2 An example of adjective clusters linked by antonymy relation

Parser, Grammar and Morphological Analyser. Word sequences that are at the basis of acronyms are subject to a well-defined grammar, simpler than a complete noun phrase grammar, but complex enough to require a nontrivial analyser. A well-established nondeterministic parsing technique was adopted. As far as the dictionary is concerned, the full WordNet lexicon was used, integrated with an ad hoc morphological analyser. Also for the generation part the grammar is exploited as the source for syntactic constraints. All the components are implemented in Common Lisp augmented with nondeterministic constructs.

Other Resources. An "a-semantic" or "slanting" dictionary is a collection of hyperbolic/attractive adjective/adverbs. This is a last resource, that sometimes can be useful in the generation of new acronyms. In fact a slanting writing refers to that type of writing that springs from our conscious or subconscious choice of words and images. We may load our description of a specific situation with vivid, connotative words and figures of speech. Some examples are *abnormally*, *abstrusely*, *adorably*, *exceptionally*, *exorbitantly*, *exponentially*, *extraordinarily*, *voraciously*, *weirdly*, *wonderfully*. This resource is handmade, using various dictionaries as information sources.

Other lexical resources are a euphemism dictionary, a proper noun dictionary, lists of typical foreign words commonly used in the language with some strong connotation.

4.2 Reanalysis and Generation

To get an ironic or "profaning" reanalysis of a given acronym, the system follows various steps and strategies. The main elements of the algorithm can be schematized as follows:

- acronym parsing and construction of a logical form
- choice of what to keep unchanged (typically the head of the highest ranking NP) and what to modify (e.g. the adjectives)

- look up for possible substitutions
- exploitation of semantic field oppositions
- granting phonological analogy: while keeping the constraint on the initial letters of the words, the overall rhyme and rhythm should be preserved (the modified acronym should sound similar to the original as much as possible)
- exploitation of WORDNET antonymy clustering for adjectives
- use of the slanting dictionary as a last resource

Figures 3 and 4 show a sketch of the HAHAcronym system architecture.

HAHAcronym, making fun of existing acronyms, amounts to an ironical rewriting, desecrating them with some unexpectedly contrasting, but otherwise consistently sounding expansion.

As far as acronym generation is concerned, the problem is more complex. To make the task more attractive – and difficult – we constrain resulting acronyms to be words of the dictionary (APPLE is good, IBM is not). The system takes in input concepts (actually synsets, possibly resulting from some other process, for instance sentence interpretation) and some minimal structural indication, such as the semantic head. The primary strategy of the system is to consider words that are in ironic relation with the input concepts as potential acronyms. By definition acronyms have

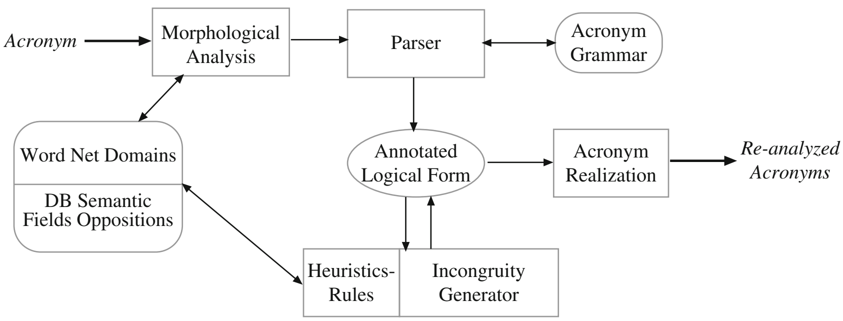

Fig. 3 The HAHAcronym system architecture

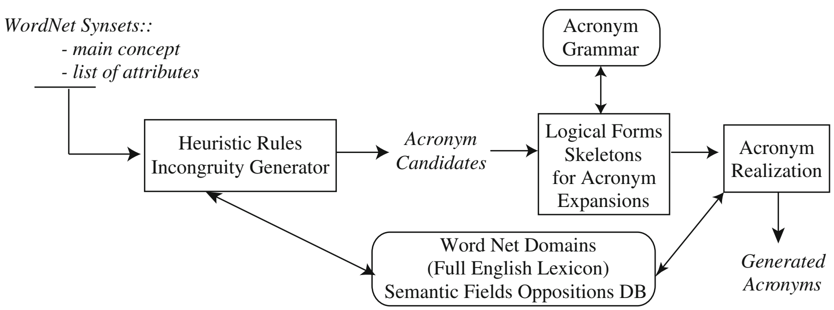

Fig. 4 Acronyms generation in the HAHAcronym system

to satisfy constraints – to include the initial letters of lexical realization, granting that the sequence of initials to satisfy the overall acronym syntax. Ironic reasoning comes mainly at the level of acronym choice and in the selection of the fillers of the *open slots* in the acronym.

For example, giving as input “fast” and “CPU”, we get static, torpid, dormant. The complete synset for “CPU” is {processor#3, CPU#1, central_proces-sing_unit#1, mainframe#2}; so we can use a synonym of “CPU” in the acronym expansion. The same happens for “fast”. Once we have an acronym proposal, a syntactic skeleton has to be filled to get a correct noun phrase. For example, given in input “fast” and “CPU”, the system selects TORPID and proposes as syntactic skeletons

$$\langle adv\rangle_T\langle adj\rangle_O \; \texttt{Rapid Processor}\langle prep\rangle\langle adj\rangle_I\langle noun\rangle_D$$

or

$$\langle adj\rangle_T\langle adj\rangle_O \; \texttt{Rapid Processor}\langle prep\rangle\langle noun\rangle_I\langle noun\rangle_D$$

where “rapid” and “processor” are synonyms, respectively, of “fast” and “CPU” and the notation <Part_of_Speech>$_{Letter}$ means a word of that particular part_of_speech with *Letter* as initial. Then the system fills this syntactic skeleton with strategies similar to those described for reanalysis.

4.3 Examples

Here below some examples of acronym reanalysis are reported. As far as semantic field opposition is concerned we have slightly tuned the system towards the domains Food, Religion, and Sex. We report the original acronym, the reanalysis, and some comments about the strategies followed by the system.

ACM – Association for Computing Machinery
→ Association for Confusing Machinery
FBI – Federal Bureau of Investigation
→ Fantastic Bureau of Intimidation

The system keeps all the main heads and works on the adjectives and the PP head, preserving the rhyme and/or using the a-semantic dictionary.

CRT – Cathodic Ray Tube
→ Catholic Ray Tube
ESA – European Space Agency
→ Epicurean Space Agency
PDA – Personal Digital Assistant
→ Penitential Demoniacal Assistant
→ Prenuptial Devotional Assistant
MIT – Massachusetts Institute of Technology
→ Mythical Institute of Theology

Some re-analyses are RELIGION oriented. Note the rhymes.

As far as generation from scratch is concerned, a main concept and some attributes (in terms of synsets) are given as input to the system. Here below we report some examples of acronym generation.

Main concept: *processor* (in the sense of CPU);
Attribute: *fast*

`OPEN` – On-line Processor for Effervescent Net
`PIQUE` – Processor for Immobile Quick Uncertain Experimentation
`TORPID` – Traitorously Outstandingly Rusty Processor for Inadvertent Data_processing
`UTMOST` – Unsettled Transcendental Mainframe for Off-line Secured TCP/IP

We note that the system tries to keep all the expansions of the acronym coherent in the same semantic field of the main concept (COMPUTER_SCIENCE). At the same time, whenever possible, it exploits some incongruity in the lexical choices.

4.4 Evaluation

Testing the humourous quality of texts is not an easy task. There have been relevant studies though, such as those in Ruch (1996). For HAHAcronym, a simpler case, an evaluation was conducted under the supervision of Salvatore Attardo at Youngstown University, Ohio. Both reanalysis and generation have been tested according to criteria of success stated in advance and in agreement with the European Commission, at the beginning of the project.

The participants in the evaluation were 40 students. They were all native speakers of English. The students were not told that the acronyms had been computer generated.

No record was kept of which student had given which set of answers (the answers were strictly anonymous). No demographic data were collected. However, generally speaking, the group was homogeneous for age (traditional students, between the ages of 19 and 24) and mixed for gender and race.

The students were divided into two groups. The first group of 20 was presented the reanalysis and generation data. We tested about 80 reanalysed and 80 generated acronyms (over twice as many as required by the agreement with the European Commission). Both the reanalysis module and the generation module were found to be successful according to the criteria spelled out in the assessment protocol (see Table 1).

The acronyms reanalysis module showed roughly 70% of acronyms having a score of 55 or higher (out of a possible 100 points), while the acronym generation module showed roughly 53% of acronyms having a score of 55 or higher. The thresholds for success established in the protocol were 60 and 45%, respectively.

Table 1 Evaluation results

Acronyms	Scored > 55%	Success thresholds (%)
Generation	52.87	45
Reanalysis	69.81	60
Random	7.69	

One could think that a random selection of fillers could be often funny as well. A special run of the system was performed with lexical reasoning and heuristics disabled, while only the syntactical constraints were operational. If the syntactical rules had been disabled as well, the output would have been gibberish and it would be not fairly comparable with normal HAHAcronym production. This set of acronyms was presented to a different group of 20 students. The result was that less than 8% of the acronyms passed the 55 points score test; we conclude that the output of HAHAcronym is significantly better than random production of reanalysis.

A curiosity that may be worth mentioning: HAHAcronym participated in a contest about (human) production of best acronyms, organized by RAI, the Italian National Broadcasting Service. The system won a jury's special prize.

5 Humour Recognition: One-Liners Recognition

Previous work in computational humour has focused mainly on the task of humour generation (Stock and Strapparava, 2003; Binsted and Ritchie, 1997), and very few attempts have been made to develop systems for automatic humour recognition (Taylor and Mazlack, 2004; Mihalcea and Strapparava, 2005b). This is not surprising, since, from a computational perspective, humour recognition appears to be significantly more subtle and difficult than humour generation.

In this section, we describe experiments concerned with the application of computational approaches to the recognition of verbally expressed humour. In particular, we investigate whether automatic classification techniques represent a viable approach to distinguish between humourous and non-humourous text, and we bring empirical evidence in support of this hypothesis through experiments performed on very large data sets.

Since a deep comprehension of humour in all of its aspects is probably too ambitious and beyond the existing computational capabilities, we chose to restrict our investigation to the type of humour found in *one-liners*. A one-liner is a short sentence with comic effects and an interesting linguistic structure: simple syntax, deliberate use of rhetoric devices (e.g. alliteration, rhyme), and frequent use of creative language constructions meant to attract the readers' attention. While longer jokes can have a relatively complex narrative structure, a one-liner must produce the humourous effect "in one shot", with very few words. These characteristics make this type of humour particularly suitable for use in an automatic learning setting, as the humour-producing features are guaranteed to be present in the first (and only) sentence.

We attempt to formulate the humour-recognition problem as a traditional classification task and feed positive (humourous) and negative (non-humourous) examples to an automatic classifier. The humourous data set consists of one-liners collected from the Web using an automatic bootstrapping process. The non-humourous data are selected such that it is structurally and stylistically similar to the one-liners. Specifically, we use four different negative data sets: (1) Reuters news titles; (2) proverbs; (3) sentences from the British National Corpus (BNC); (4) commonsense statements from the Open Mind Common Sense (OMCS) corpus. The classification results are encouraging, with accuracy figures ranging from 79.15% (One-liners/BNC) to 96.95% (one-liners/Reuters). Regardless of the non-humourous data set playing the role of negative examples, the performance of the automatically learned humour recognizer is always significantly better than a priori known baselines.

5.1 Humourous and Non-humourous Data Sets

To test our hypothesis that automatic classification techniques represent a viable approach to humour recognition, we needed in the first place a data set consisting of humourous (positive) and non-humourous (negative) examples. Such data sets can be used to automatically *learn* computational models for humour recognition and at the same time *evaluate* the performance of such models.

Humourous data. While there are plenty of non-humourous data that can play the role of negative examples, it is significantly harder to build a very large and at the same time sufficiently "clean" data set of humourous examples. We use a dually constrained Web-based bootstrapping process to collect a very large set of one-liners. Starting with a short *seed* set consisting of a few one-liners manually identified, the algorithm automatically identifies a list of webpages that include at least one of the seed one-liners, via a simple search performed with a Web search engine. Next, the webpages found in this way are HTML parsed, and additional one-liners are automatically identified and added to the seed set. The process is repeated several times, until enough one-liners are collected. As with any other bootstrapping algorithm, an important aspect is represented by the set of constraints used to steer the process and prevent as much as possible the addition of noisy entries. Our algorithm uses (1) a *thematic* constraint applied to the theme of each webpage, via a list of keywords that have to appear in the URL of the webpage, and (2) a *structural* constraint, exploiting HTML annotations indicating text of similar genre (e.g. lists, adjacent paragraphs)

Two iterations of the bootstrapping process, started with a small seed set of 10 one-liners, resulted in a large set of about 24,000 one-liners. After removing the duplicates using a measure of string similarity based on the longest common subsequence, we are left with a final set of 16,000 one-liners, which are used in the

Table 2 Sample examples of one-liners, Reuters titles, proverbs, OMC and BNC sentences

One-liners	
Take my advice; I don't use it anyway. I get enough exercise just pushing my luck. Beauty is in the eye of the beer holder.	
Reuters titles	*Proverbs*
Trocadero expects tripling of revenues. Silver fixes at two-month high, but gold lags. Oil prices slip as refiners shop for bargains.	Creativity is more important than knowledge. Beauty is in the eye of the beholder. I believe no tales from an enemy's tongue.
OMCS sentences	*BNC sentences*
Humans generally want to eat at least once a day. A file is used for keeping documents. A present is a gift, something you give to someone.	They were like spirits, and I loved them. I wonder if there is some contradiction here. The train arrives three minutes early.

humour-recognition experiments. A more detailed description of the Web-based bootstrapping process is available in Mihalcea and Strapparava (2005a). The one-liners humour style is illustrated in Table 2, which shows three examples of such one-sentence jokes.

Non-humourous data. To construct the set of negative examples required by the humour-recognition models, we tried to identify collections of sentences that were non-humourous, but similar in structure and composition to the one-liners. We do not want the automatic classifiers to learn to distinguish between humourous and non-humourous examples based simply on text length or obvious vocabulary differences. Instead, we seek to enforce the classifiers to identify humour-specific features, by supplying them with negative examples similar in most of their aspects to the positive examples, but different in their comic effect.

We tested four different sets of negative examples, with three examples from each data set illustrated in Table 2. All non-humourous examples are enforced to follow the same length restriction as the one-liners, i.e. one sentence with an average length of 10–15 words.

1. *Reuters* titles, extracted from news articles published in the Reuters newswire over a period of 1 year (20 August 1996–19 August 1997) (Lewis et al., 2004). The titles consist of short sentences with simple syntax and are often phrased to catch the readers' attention (an effect similar to the one rendered by the one-liners).
2. *Proverbs* extracted from an online proverb collection. Proverbs are sayings that transmit, usually in one short sentence, important facts or experiences that are considered true by many people. Their property's of being condensed, but memorable sayings make them very similar to the one-liners. In fact, some one-liners attempt to reproduce proverbs, with a comic effect, as in e.g. *"Beauty is in the eye of the beer holder"*, derived from *"Beauty is in the eye of the beholder"*.

3. *British National Corpus (BNC)* sentences, extracted from BNC – a balanced corpus covering different styles, genres, and domains. The sentences were selected such that they were similar in content with the one-liners: we used an information retrieval system implementing a vectorial model to identify the BNC sentence most similar to each of the 16,000 one-liners.[4] Unlike the Reuters titles or the proverbs, the BNC sentences have typically no added creativity. However, we decided to add this set of negative examples to our experimental setting, in order to observe the level of difficulty of a humour-recognition task when performed with respect to simple text.
4. *Open Mind Common Sense (OMCS)* sentences. OMCS is a collection of about 800,000 commonsense assertions in English as contributed by volunteers over the Web. It consists mostly of simple single sentences, which tend to be explanations and assertions similar to glosses of a dictionary, but phrased in a more common language. For example, the collection includes such assertions as "keys are used to unlock doors" and "pressing a typewriter key makes a letter". Since the comic effect of jokes is often based on statements that break our commonsensical understanding of the world, we believe that such commonsense sentences can make an interesting collection of "negative" examples for humour recognition. For details on the OMCS data and how it has been collected, see Singh (2002). From this repository we use the first 16,000 sentences.[5]

To summarize, the humour recognition experiments rely on data sets consisting of humourous (positive) and non-humourous (negative) examples. The positive examples consist of 16,000 one-liners automatically collected using a Web-based bootstrapping process. The negative examples are drawn from (1) Reuters titles; (2) proverbs; (3) BNC sentences; and (4) OMCS sentences.

5.2 Features for Automatic Humour Recognition

We experiment with automatic classification techniques using (a) heuristics based on humour-specific stylistic features (alliteration, antonymy, slang); (b) content-based features, within a learning framework formulated as a typical text classification task; and (c) combined stylistic and content-based features, integrated in a stacked machine learning framework.

[4]The sentence most similar to a one-liner is identified by running the one-liner against an index built for all BNC sentences with a length of 10–15 words. We use a *tf.idf* weighting scheme and a cosine similarity measure, as implemented in the Smart system (`ftp.cs.cornell.edu/pub/smart`, visited 30 May 2010).

[5]The first sentences in this corpus are considered to be "cleaner", as they were contributed by trusted users (Push Singh, p.c.).

5.2.1 Humour-Specific Stylistic Features

Linguistic theories of humour (e.g. Attardo, 1994) have suggested many *stylistic features* that characterize humourous texts. We tried to identify a set of features that were both significant and feasible to implement using existing machine-readable resources. Specifically, we focus on alliteration, antonymy, and adult slang, previously suggested as potentially good indicators of humour (Ruch, 2002; Bucaria, 2004).

Alliteration. Some studies on humour appreciation (Ruch, 2002) show that structural and phonetic properties of jokes are at least as important as their content. In fact one-liners often rely on the reader awareness of attention-catching sounds, through linguistic phenomena such as alliteration, word repetition, and rhyme, which produce a comic effect even if the jokes are not necessarily meant to be read aloud. Note that similar rhetorical devices play an important role in wordplay jokes and are often used in newspaper headlines and in advertisement. The following one-liners are examples of jokes that include alliteration chains:

Veni, Vidi, Visa: I came, I saw, I did a little shopping.
Infants don't enjoy infancy like adults do adultery.

To extract this feature, we identify and count the number of alliteration/rhyme chains in each example in our data set. The chains are automatically extracted using an index created on top of the CMU Pronuncing Dictionary.

Antonymy. Humour often relies on some type of incongruity, opposition, or other forms of apparent contradiction. While an accurate identification of all these properties is probably difficult to accomplish, it is relatively easy to identify the presence of *antonyms* in a sentence. For instance, the comic effect produced by the following one-liners is partly due to the presence of antonyms:

A clean desk is a sign of a cluttered desk drawer.
Always try to be modest and be proud of it!

The lexical resource we use to identify antonyms is WordNet (Miller, 1995), and in particular the *antonymy* relation among nouns, verbs, adjectives, and adverbs. For adjectives we also consider an indirect antonymy via the *similar-to* relation among adjective synsets. Despite the relatively large number of *antonymy* relations defined in WordNet, its coverage is far from complete, and thus the *antonymy* feature cannot always be identified. A deeper semantic analysis of the text, such as word sense or domain disambiguation, could probably help in detecting other types of semantic opposition, and we plan to exploit these techniques in future work.

Adult Slang. Humour based on adult slang is very popular. Therefore, a possible feature for humour recognition is the detection of sexual-oriented lexicon in the sentence. The following represent examples of one-liners that include such slang:

The sex was so good that even the neighbors had a cigarette.
Artificial Insemination: procreation without recreation.

To form a lexicon required for the identification of this feature, we extract from WORDNET DOMAINS[6] all the synsets labelled with the domain SEXUALITY. The list is further processed by removing all words with high polysemy (≥ 4). Next, we check for the presence of the words in this lexicon in each sentence in the corpus and annotate them accordingly. Note that, as in the case of antonymy, WORDNET coverage is not complete, and the *adult slang* feature cannot always be identified. Finally, in some cases, all three features (alliteration, antonymy, adult slang) are present in the same sentence, as for instance the following one-liner:

> *Behind every great$_{al}$ man$_{ant}$ is a great$_{al}$ woman$_{ant}$, and behind every great$_{al}$ woman$_{ant}$ is some guy staring at her behind$_{sl}$!*

5.2.2 Content-Based Learning

In addition to stylistic features, we also experimented with *content-based features*, through experiments where the humour-recognition task is formulated as a traditional text classification problem. Specifically, we compare results obtained with two frequently used text classifiers, Naïve Bayes and Support Vector Machines, selected based on their performance in previously reported work and for their diversity of learning methodologies.

5.3 Experimental Results

Several experiments were conducted to gain insights into various aspects related to an automatic humour-recognition task: classification accuracy using stylistic and content-based features, learning rates, impact of the type of negative data, impact of the classification methodology. All evaluations are performed using stratified 10-fold cross validations, for accurate estimates. The baseline for all the experiments is 50%, which represents the classification accuracy obtained if a label of "humourous" (or "non-humourous") would be assigned by default to all the examples in the data set.

5.3.1 Heuristics Using Humour-Specific Features

In a first set of experiments, we evaluated the classification accuracy using stylistic humour-specific features: alliteration, antonymy, and adult slang. These are numerical features that act as heuristics, and the only parameter required for their application is a threshold indicating the minimum value admitted for a statement to be classified as humourous (or non-humourous). These thresholds are learned automatically using a decision tree applied on a small subset of humourous/non-humourous

[6] WORDNET DOMAINS assigns each synset in WORDNET with one or more "domain" labels, such as SPORT, MEDICINE, ECONOMY. See http://wndomains.itc.it.

Table 3 Humour-recognition accuracy using alliteration, antonymy, and adult slang

	One-liners			
Heuristic	Reuters (%)	BNC (%)	Proverbs (%)	OMCS (%)
Alliteration	74.31	59.34	53.30	55.57
Antonymy	55.65	51.40	50.51	51.84
Adult slang	52.74	52.39	50.74	51.34
All	76.73	60.63	53.71	56.16

examples (1000 examples). The evaluation is performed on the remaining 15,000 examples, with results shown in Table 3.[7]

Considering the fact that these features represent *stylistic* indicators, the style of Reuters titles turns out to be the most different with respect to one-liners, while the style of proverbs is the most similar. Note that for all data sets the alliteration feature appears to be the most useful indicator of humour, which is in agreement with previous linguistic findings (Ruch, 2002).

5.3.2 Text Classification with Content Features

The second set of experiments was concerned with the evaluation of content-based features for humour recognition. Table 4 shows results obtained using the four different sets of negative examples, with the Naïve Bayes and SVM classifiers.

Once again, the content of Reuters titles appears to be the most different with respect to one-liners, while the BNC sentences represent the most similar data set. This suggests that joke content tends to be very similar to regular text, although a reasonably accurate distinction can still be made using text classification techniques. Interestingly, proverbs can be distinguished from one-liners using content-based features, which indicates that despite their stylistic similarity (see Table 3), proverbs and one-liners deal with different topics.

Table 4 Humour-recognition accuracy using Naïve Bayes and SVM text classifiers

	One-liners			
Classifier	Reuters (%)	BNC (%)	Proverbs (%)	OMCS (%)
Naïve Bayes	96.67	73.22	84.81	82.39
SVM	96.09	77.51	84.48	81.86

5.3.3 Combining Stylistic and Content Features

Encouraged by the results obtained in the first two experiments, we designed a third experiment that attempts to jointly exploit stylistic and content features for

[7]We also experimented with decision trees learned from a larger number of examples, but the results were similar, which confirms our hypothesis that these features are heuristics, rather than learnable properties that improve their accuracy with additional training data.

Table 5 Humour-recognition accuracy for combined learning based on stylistic and content features

One-liners			
Reuters	BNC	Proverbs	OMCS
96.95%	79.15%	84.82%	82.37%

humour recognition. The feature combination is performed using a stacked learner, which takes the output of the text classifier, joins it with the three humour-specific features (alliteration, antonymy, adult slang), and feeds the newly created feature vectors to a machine learning tool. Given the relatively large gap between the performance achieved with content-based features (text classification) and stylistic features (humour-specific heuristics), we decided to do the meta-learning using a rule-based learner, so that low-performance features are not eliminated in favour of the more accurate ones. We use the Timbl memory-based learner (Daelemans et al., 2001) and evaluate the classification using a stratified 10-fold cross-validation. Table 5 shows the results obtained for the four data sets.

Combining classifiers results in a statistically significant improvement ($p < 0.0005$, paired t-test) with respect to the best individual classifier for the One-liners/Reuters and one-liners/BNC data sets, with relative error rate reductions of 8.9 and 7.3%, respectively. No improvement is observed for the one-liners/proverbs and one-liners/OMCS data sets, which is not surprising since, as shown in Table 3, proverbs and commonsense statements cannot be clearly differentiated using stylistic features from the one-liners, and thus the addition of these features to content-based features is not likely to result in an improvement.

The experimental results prove that computational approaches can be successfully used for the task of humour recognition. An analysis of the results shows that the humourous effect can be identified in a large fraction of the jokes in our data set using surface features such as alliteration, word-based antonymy, or specific vocabulary. Moreover, we also identify cases where our current automatic methods fail, which require more sophisticated techniques such as recognition of irony, detection of incongruity that goes beyond word antonymy, or commonsense knowledge. Finally, an analysis of the most discriminative content-based features identified during the process of automatic classification helps us point out some of the most predominant semantic classes specific to humourous text, which could be turned into useful features for future studies of humour generation.

6 Prospects for Computational Humour

Humour is an important mechanism for communicating new ideas and for changing perspectives. On the cognitive side humour has two very important properties: it helps getting and keeping people's attention and it helps remembering. Type and rhythm of humour may vary and the time involved in building the humourous effect may be varied in different cases: sometimes there is a context, as in joke telling, which allows you to expect since the beginning the humourous climax, even if it occurs after a long while. Other times the effect is obtained in almost no time, with one perceptive act. This is the case of static visual humour, of ironic pictures, when some well-established convention is reversed through the combination with an evocative surprising utterance. Many advertisement-oriented expressions have this property. The role of variation of a known expression seems to be of high importance and studies have also shown the positive impact on the audience of forms of incongruity in the resulting expressions.

As for memorization, it is a common experience to connect in our memory new knowledge with humourous remarks or events. In foreign language acquisition, it sometimes happens that involuntary humourous situations are created because of so-called false friends words that sound similar in the two languages, but have a very different meaning. The "false friends" acknowledgment is conducive to remembering the correct use of the word. Similarly, as shown experimentally, a good humourous expression has exceptionally good recall quality not only per se but also for product type and brand. For a large number of verbal expressions what it takes is the ability to perform *optimal innovation* (Giora, 2002) of existing material, with a humourous connotation. When the novelty is in a complementary relation to salience (familiarity), it is "optimal" in the sense that it has an aesthetics value and "induces the most pleasing effect". Therefore the simultaneous presence of novelty and familiarity makes the message potentially surprising, because this combination allows the recipient's mind to oscillate between what is known and what is different from usual.

A good strategy is to start from well-known expressions that are firm points for the audience and to creatively connect them to the concept or element we intend to promote. We should then be able to perform variations either in the external context, in case the material is ambiguous and the audience can be lured to a different interpretation, or, more often, within the expression itself, changing some material of the expression appropriately, while still preserving full recognizability of the original expression. For instance a good advertising expression for a soft drink is *"Thirst come, thirst served"*, an obvious alteration of a known expression.

In most fields of AI the difficulties of reasoning on deep world knowledge have been recognized for a long while. There is a clear problem in scaling up between toy experiments and meaningful, large-scale applications. The more so in an area such as humour, by many called "AI-complete", where good quality expressions require subtle understanding of situations and are normally the privilege of talented individuals. A goal of computational humour should be to produce general mechanisms limited to the humourous revisitation of verbal expressions, but meant to work in unrestricted domains. To this end, we consider the use of affective terms a critical aspect in communication and in particular for humour. Valence (positive or negative polarity) of a term and its intensity (the level of arousal it provides) are fundamental factors for persuasion and also for humourous communication. Making fun of biased expression or alluding to related "coloured" concepts plays an important role for humourous revisitation of existing expressions.

From an application point of view we think the world of advertisement has a great potential for the adoption of computational humour. Perception of humour in promotional messages produces higher attention and in general a better recall than non-humourous advertisement of the product category, of the specific brand, and of the advertisement itself (Perry et al., 1997).

The future of advertisement will include three important themes: (1) reduction in time to market and extension of possible occasions for advertisement; (2) more attention to the wearing out of the message and for the need for planning variants and connected messages across time and space; (3) contextual personalization, on the basis of audience profile and perhaps information about the situation. All three cases call for a strong role for computer-based intelligent technology for producing novel appropriate advertisements. We believe that computational humour will help produce those kinds of messages that have been so successful in the "slow" or non personalized situation we have lived in.

References

Aristotle. Rhetoric. 350 BC

Attardo S (1994) Linguistic theories of humor. Mouton de Gruyter, Berlin

Attardo S, Raskin V (1991) Script theory revis(it)ed: joke similarity and joke representation model. Humour 4(3):293–347

Bartolo A, Benuzzi F, Nocetti L, Baraldi P, Nichelli P (2006) Humor comprehension and appreciation: an fmri study. J Cogn Neurosci, 18(11):1789–1798

Binsted K, Ritchie G (1997) Computational rules for punning riddles. Humor 10(1):25–76

Bucaria C (2004) Lexical and syntactic ambiguity as a source of humor. Humor 17(3):279–309

Daelemans W, Zavrel J, van der Sloot K, van den Bosch A (2001) Timbl: Tilburg memory based learner, version 4.0, reference guide. Technical report, University of Antwerp

Duncan W (1984) Perceived humor and social network patterns in a sample of task-oriented groups: a reexamination of prior research. Hum Relat 37(11):895–907

Fellbaum C (1998) WordNet. An electronic lexical database. The MIT Press, Cambridge, MA

Freud S (1905) Der Witz und Seine Beziehung zum Unbewussten. Deutike, Leipzig

Giora R (2002) Optimal innovation and pleasure. In: Proceedings of the April fools day workshop on computational humour (TWLT20), Trento, Italy

Hay J (1995) Gender and humour: beyond a joke. Master's thesis, Victoria University of Wellington

Hertzler J (1970) Laughter: a social scientific analysis. Exposition Press, New York, NY

Hetzron R (1991) On the structure of punchlines. Int J Humor Res 4(1):61–108

Hobbes T (1840) Human nature in english works. Molesworth

Kessler B, Nunberg G, Schuetze H (1997) Automatic detection of text genre. In: Proceedings of the 35th annual meeting of the association for computational linguistics (ACL97), Madrid

Lewis D, Yang Y, Rose T, Li F (2004, December) RCV1: A new benchmark collection for text categorization research. J Mach Learn Res 5:361–397

Liu H, Mihalcea R (2007) Of men, women, and computers: data-driven gender modeling for improved user interfaces. In: Proceedings of international conference on weblogs and social media, Boulder, CO

Magnini B, Strapparava C, Pezzulo G, Gliozzo A (2002) The role of domain information in word sense disambiguation. J Nat Lang Eng 8(4):359–373

Mihalcea R, Pulman S (2007) Characterizing humour: an exploration of features in humorous texts. In: Proceedings of the conference on intelligent text processing and computational linguistics, Mexico City

Mihalcea R, Strapparava C (2005a) Bootstrapping for fun: web-based construction of large data sets for humor recognition. In: Proceedings of the workshop on negotiation, behaviour and language (FINEXIN 2005), Ottawa, Canada

Mihalcea R, Strapparava C (2005b, October) Making computers laugh: investigations in automatic humor recognition. In: Proceedings of the joint conference on human language technology/empirical methods in natural language processing (HLT/EMNLP), Vancouver

Miller G (1995) Wordnet: a lexical database. Commun ACM 38(11):39–41

Minsky M (1980) Jokes and the logic of the cognitive unconscious. Technical report, MIT Artificial Intelligence Laboratory, 1980. AI memo 603

Mobbs D, Hagan C, Azim E, Menon C, Reiss A (2005) Personality predicts activity in reward and emotional regions associated with humor. Proc Nat Acad Sci, USA, 102, 16502–16506

Nevo O (1984) Appreciation and production of humor as an expression of aggression. J Cross-Cult Psychol 15(2):181–198

Oaks D (1994) Creating structural ambiguities in humor: getting english grammar to cooperate. Int J Humor Res 7(4):377–402

Perry S, Jenzowsky S, King C, Yi H, Hester J, Gartenschlaeger J (1997) Humorous programs as a vehicle of humorous commercials. J Commun 47(1):20–39

Purandare A, Litman D (2006) Humor: Prosody analysis and automatic recognition for F*R*I*E*N*D*S*. In: Proceedings of the 2006 conference on empirical methods in natural language processing, Sydney, Australia

Raskin V (1985) Semantic mechanisms of humor. Reidel, Dordrecht

Ritchie G (2003). The linguistic analysis of jokes. Routledge, London

Ruch W, Attardo S, Raskin V (1993) Toward an empirical verification of the general theory of verbal humor. Int J Humor Res 6(2):123–136

Ruch W (1996) Special issue: measurement approaches to the sense of humor. Humor 9(3/4)

Ruch W (1998) The sense of humor: explorations of a personality characteristic. Mouton de Gruyte, Berlin

Ruch W (2002) Computers with a personality? lessons to be learned from studies of the psychology of humor. In: Proceedings of the April fools day workshop on computational humour (TWLT20), Trento, Italy

Schopenhauer A (1819) The World as Will and Idea. Kessinger Publishing, Whitefish, MT

Singh P (2002) The public acquisition of commonsense knowledge. In: Proceedings of AAAI spring symposium: acquiring (and using) linguistic (and world) knowledge for information access., Palo Alto, CA

Solomon R (2002) Are the three stooges funny? Soitainly! (or When is it OK to laugh?). In: Rudinow J, Graybosch A, (eds) Ethics and values in the information age. Wadsworth, New York, NY

Spencer H (1860) The physiology of laughter. Macmillan's Magazine 1, 395–402

Stock O, Strapparava C (2003, August) Getting serious about the development of computational humour. In: Proceedings of the 8th international joint conference on artificial intelligence (IJCAI-03), Acapulco, Mexico

Stock O, Strapparava C, Nijholt A (eds) (2002) Proceedings of the April fools day workshop on computational humour (TWLT20), Trento, Italy

Taylor J, Mazlack L (2004) Computationally recognizing wordplay in jokes. In: Proceedings of 26th annual meeting of the cognitive science society, Chicago, August 2004

Tinholt HW, Nijholt A (2007) Computational humour: utilizing cross-reference ambiguity for conversational jokes. In: Masulli F, Mitra S, Pasi G, (eds) 7th international workshop on fuzzy logic and applications (WILF 2007), Lecture Notes in Artificial Intelligence, Berlin, 2007. Springer, Berlin, pp 477–483

Watson K, Matthews B, Allman J (2007) Brain activation during sight gags and language-dependent humor. Cereb cortex 17(2):314–324

Yuill N (1997) A funny thing happened on the way to the classroom: Jokes, riddles and metalinguistic awareness in understanding and improving poor comprehension in children. In Disabilities: Processes and Intervention. Erlbaum, Hillsdale, NJ, pp 193–220