HAL
open science

Jokes and Belief Revision

Florence Dupin de Saint-Cyr, Henri Prade

To cite this version:

Florence Dupin de Saint-Cyr, Henri Prade. Jokes and Belief Revision. 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020), IJCAI: International Joint Conferences on Artificial Intelligence Organization, Sep 2020, Rhodes, Greece. pp.336-340, 10.24963/kr.2020/34 . hal-03036921

HAL Id: hal-03036921

https://hal.science/hal-03036921v1

Submitted on 2 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Jokes and Belief Revision*

Florence Dupin de Saint-Cyr
florence.bannay@irit.fr

Henri Prade
henri.prade@irit.fr

IRIT-CNRS, Toulouse University,
118, route de Narbonne, 31062 Toulouse Cedex 9

Abstract

The paper deals with a topic little studied in artificial intelligence: the understanding of humor. In this preliminary study, we try to identify the basic mechanism at work in quips and narrative jokes. It seems that in many cases a belief revision process is operating, leading to an unexpected conclusion through the punchline of the jest. We propose a formal modeling of jokes based on belief revision. Namely the punchline, which triggers a revision, is both surprising and explains perfectly what was reported in the beginning of the joke. This also suggests a way of ranking jokes in terms of surprise and strength of explanation, using possibilistic logic.
Keywords: Humor, belief revision, surprise.

1 Introduction

Aristotle already[1] pointed out that laughing is proper to human. Humor is one of those cognitive topics, as also, e.g., aesthetic judgments, that has been little studied in AI. Presumably such topics do not look important and may seem tricky. Their understanding both involve perception, analysis and synthesis issues. In both cases, a "sender", the joker or the artist, presents an expressive content to a "receiver", the audience, supposedly capable of appreciating what is intended for them. Humor and art also have in common that they often play with, or even shake up the prohibitions and prejudices. In both cases, there seems to be a surprise effect (Raccah 2016; Williams 1996) for the public when discovering what is presented. Very schematically, while art addresses first of all our sensitivity and arouses our emotions, it seems that jokes solicit above all a logical apprehension of the world (one speaks of "understanding a joke").

The question whether an intelligent system is able to understand humor remains largely open. The aim of this paper is to propose a representation of the cognitive comprehension of jokes and more precisely to present a preliminary modeling of the mechanism that triggers laughter in a joke. We first review the AI literature related to humor then briefly discuss the modeling of surprises before presenting the formalization of the mechanism based on belief revision, which is at work in most jokes[2].

2 Humor in AI

A large proportion of the AI literature on humor concerns the detection of humor or irony in texts, audio signals or images, or the evaluation of a level of humor. Some other works focus on the generation of puns, or the derivation of hijacked acronyms.

*This is a draft version, the article is published in the Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning pages 336-340

[1]Greek jokes were early compiled, nearly 2,000 years ago; see, e.g., (Anonymous 2008). Here is an example: *An intellectual, who is short on cash, has to sell all his books. He then wrote to his father: "I have good news: I'm beginning to make a living from my books"*. The reader may check that the approach presented in this paper would apply to this old joke.

[2]This is a fully rewritten version of a previous work in French (Dupin De Saint-Cyr and Prade 2019) which contains a more extensive introduction and more references.

Marvin Minsky (1984) discusses the Freudian theory of jokes (Freud 1961) and highlights that beyond questions of prohibitions, humor is also a matter of knowledge about knowledge, and of inconsistency resolution. According to (Attardo and Raskin 1991) there are three main theories about humor: the theory of relief (humor would help relieve nervous tension), the theory of superiority (humor would make one feel superior), and the theory of incongruity (humor consists of disobeying mental patterns and expectations). This last theory is the dominant theory which is also the perspective in which we place ourselves in this article (even if we can argue that the punchline of the joke restores somehow the consistency and calms a cognitive dissonance, which may seem to go also in the direction of the first theory). The same authors proposed the General Theory of Verbal Humor (Attardo and Raskin 1991; Raskin 2008). According to GTVH, the six main resources involved in jokes are (in decreasing order of importance): script opposition (SO), logical mechanism (LM), situation (SI), target (TA), narrative strategy (NS) and language (LA). Then the paper by (Hempelmann, Raskin, and Triezenberg 2006) uses ontologies to analyze and generate jokes automatically following predefined scripts. In this paper, we mainly consider the first two resources which can be related to the idea of surprise and to the logical machinery triggered by the punchline of a joke.

We can also mention (Stock and Strapparava 2003) which presents HAHACRONYM a system for producing humorous acronyms from existing acronyms (e.g., MIT becomes "Mythical Institute of Theology") which explains the series of letters in an unexpected way. Another research work by (Shahaf, Horvitz, and Mankoff 2015) concerns the automatic analysis of the degree of humor associated with an image. For this aim, the authors design a learning mechanism based on a database of thousands of funny captions associated with images. The captions were annotated in pairs indicating the funnier one of the two. The characteristics used to formulate the learning task are: the unusual nature of the language used, the unpredictability of the legend, the grammatical complexity.

Lastly, Raccah (2016) draws a parallel between metaphor and *bon mot*, both of which rely on a "manipulative" effect in communication. According to him, humor requires three conditions: i) a story; ii) an intention to make people laugh; iii) that the story causes laughter. Raccah offers as an explanation of a *bon mot* that makes us laugh "the feeling of having fallen into an unexpected and inevitable trap: it is the punchline that makes us fall into this trap".

3 Surprise in AI

Surprise is triggered by the occurrence of something considered as having low possibility. The first version of the theory of possibility proposed by the English economist George Shackle (1961) was based on the notion of the degree of surprise associated with an event, which was in fact a degree of impossibility of that event, calculated from a possibility distribution reflecting our uncertain knowledge about the current state of the world. A fact is then all the more surprising as it is less consistent with what we thought was possible. Formally,

$$\texttt{degree of surprise}(A) = 1 - \texttt{possibility}(A)$$

In AI, the issue of surprises was discussed in (Lorini and Castelfranchi 2007). These authors have formulated two different notions of surprise: i) *mismatch-based surprise* which reflects the incompatibility between what we perceive and what we expected to perceive (*scrutinized expectation*); ii) *astonishment* or *surprise in recognition* which reflects the difficulty of accepting what we perceive because this percept is very implausible in the absolute. In both cases the degree of surprise or astonishment is associated with a probability. In the first case, the greater the probability associated with the initial expectation, the greater the surprise; in the second case, the greater the surprise, the more unlikely the perceived thing is. Lorini and Castelfranchi propose a logic of probabilistic quantified beliefs using a semantics given in (Fagin and Halpern 1994) to which a standard action operator is added as well as constructors to talk about what is being examined and the perceived data. They proposed to model the cognitive integration that follows a surprise by a belief change operator that they call *update process*. Thus, for these authors, the notion of surprise is first defined by a form of inconsistency with respect to explicit expectations or with respect to prior knowledge.

In our study, we propose a different approach since we use belief change theory in order to define surprise (and not to model its integration by the listener): the new information (the punchline of the joke)

generates a revision of beliefs whose result is surprising in the sense that it is inconsistent with what was initially believed. Note that in this article the notion of surprise will not correspond to a surprising evolution of the world (Dupin de Saint-Cyr and Lang 2011), but rather to a surprising evolution of knowledge about the world. Thus, our work relates to belief revision rather than to update (Winslett 1988; Katsuno and Mendelzon 2003).

4 Modeling Jokes

In humor, according to Raccah (2016), there is a manipulation of what the joker wants the listener to believe, in order to make him fall into a trap. In this article on humor, we propose to translate Raccah's analysis in terms of belief revision. Thus, the manipulation of the listener will be seen as a two-phase process:

- The first phase corresponds to the description of the con-text of the joke. It leads to a revision of listener's beliefs by this intentionally incomplete description made in order to suggest a conclusion which will turn out to be false.

- The second phase corresponds to the arrival of the punchline of the joke, which on the one hand is surprising because it is incompatible or incongruent with the previous provisional conclusion. On the other hand, the punchline seems inevitable what we express by the fact that it logically explains the initial information.

The second lesson learnt from Raccah's article concerns humor intensity: according to him, "the punchline is all the more funny as the trap was unexpected and inevitable". This prompts us to propose a way to compare jokes in terms of the discrepancy between the levels of plausibility of the worlds obtained in the two phases on the one hand, and in terms of the inevitability of the explanation of the initial situation provided by the punchline on the other hand.

4.1 Notations and Background

We consider a propositional language $\mathcal{L}$, where propositions are noted by lowercase Greek symbols. The symbols $\perp$, $\vee$, $\wedge$, $\neg$, $\rightarrow$, $\equiv$, $\vdash$, $\models$ denote respectively contradiction, logical connectors "or", "and", "not", material implication, logical equivalence, classical inference, satisfaction.

In the following, we use a *belief revision* operator (Alchourrón, Gärdenfors, and Makinson 1985; Katsuno and Mendelzon 1991). The revision of a set of propositional formulas K by a piece of information φ is noted $K \circ \varphi$.

Note that belief revision semantics is based on sphere systems (Grove 1988) and epistemic entrenchments (Gärdenfors 1988), two notions that have their exact counterpart in possibility theory (under the form of qualitative possibility distributions and necessity relations (Dubois and Prade 1991)). A possibility distribution is qualitative in the sense that it only plays the role of an ordinal ranking function on the set of possible worlds (contrarily to what a probability distribution would do). Possibility distributions are used here for implementing the revision operator $\circ$ because what is required is simply a plausibility ordering as in sphere systems and no numerical value is neither needed nor available. However the approach presented in this paper holds for any kind of belief revision operator satifisfying Katsuno-Mendelzon postulates (1991).

Since we need to encode common sense knowledge, we use $\rightsquigarrow$ to represent a default rule: $\alpha \rightsquigarrow \beta$, with $\alpha, \beta \in \mathcal{L}$; meaning that when α is true it is more plausible that β is true than false. This kind of rule is easily translated in the possibility theory setting (Dubois and Prade 1988) as the constraint $N(\beta|\alpha) > 0$. Due to the duality between a necessity measure N and its associated possibility measure Π, i.e., $N(\alpha) = 1 - \Pi(\neg\alpha)$, the constraint $N(\beta|\alpha) > 0$ is equivalently written as $\Pi(\alpha \wedge \beta) > \Pi(\alpha \wedge \neg\beta)$ which expresses that, in context α, β being true is the normal situation (since strictly more possible than β false). This writing supposes the existence of a plausibility relation on worlds, which is represented by a possibility distribution π, i.e., a mapping from the set of interpretations of $\mathcal{L}$ to a linearly ordered scale bounded by 0 and 1, where 0 stands for impossible and 1 for fully possible.

4.2 Formalization

In this section, we propose a *static* formalization of jokes in propositional logic, i.e., we consider that a joke describes a situation α and that its punchline completes this description with a piece of information β (α, β are propositions in $\mathcal{L}$).

We consider that the joke is addressed to a listener whose knowledge base is a set of propositional formulas denoted K. Moreover, the listener is also characterized by her way $\circ$ to revise her belief.

In the following, a (simple) joke is viewed as a story, which is a pair (<context>, <punchline>) having some further characteristics. We first define a story.

Definition 1 (simple story, cascading story)**.** *A* simple story *is a pair* (α, β) *of formulas of* $\mathcal{L}$*. A* cascading story *is a tuple* $(\varphi_1, \ldots, \varphi_n)$ *with* $n > 2$ *formulas in* $\mathcal{L}$*.*

The cognitive situation induced by a story is described by the knowledge base K iteratively revised by the elements of the story. The punchline of a joke is considered *surprising* if the result of the revision of the knowledge base K by the initial description of the situation, is contradictory to what is obtained after revision by the entire story including the punchline.

Definition 2 (surprising story)**.** *A story* (α, β) *is* surprising *with respect to a listener associated to* $(K, \circ)$ *if*

$(K \circ \alpha)$ *consistent*[3] *and* $(K \circ \alpha) \cup (K \circ (\alpha \wedge \beta)) \vdash \bot$

A cascading story $(\varphi_1, \ldots, \varphi_n)$ *is surprising with respect to* $(K, \circ)$ *at the stage* i *if*

$(K \circ \psi)$ *consistent and* $(K \circ \psi) \cup (K \circ (\psi \wedge \varphi_i)) \vdash \bot$

with $\psi = \varphi_1 \wedge \varphi_2 \ldots \wedge \varphi_{i-1}$.

Note that the above definition excludes the possibility that the listener already knows the joke. Indeed in that latter case α and β would be already in K, forbidding any surprise. Besides, we do not require that $K \circ (\alpha \wedge \beta)$ is consistent, including absurd punchlines in Definition 2[4].

In order to understand the joke, its logic has to appear implacable once the punchline is revealed, so the punchline is both completely admissible and explains the situation. In other words, if we had known β from the beginning, it would have explained α, which we can translate as follows:

Definition 3 (revealing punchline)**.** *Given a story* (α, β)*, its punchline* β *is* revealing *with respect to* $(K, \circ)$ *if*

$(K \circ \beta)$ *consistent*[3] *and* $K \circ \beta \models \alpha$

Laughter can be thought to relieve the tension of cognitive dissonance caused by the inconsistency between what one expected to hear and the punchline of the joke. This relief can only come once the story is understood, with the punchline playing the role of a revelation.

Definition 4 (potentially funny story)**.** *A story is* potentially funny *if it is surprising and its punchline is revealing.*

This definition of a potentially funny story is directly inspired by Raccah (2016), except that we have tried to define the revealing character of the punchline rather than its inevitability (which could be understood as $K \circ \alpha \vdash \beta$, denying the surprising character of β).

Observe also that the definition relies both on K and $\circ$, which may differ from a listener to another, acknowledging the fact that a joke may not be found funny by anybody.

Note that the expression "potentially funny" is used in Definition 4 because one can imagine examples of surprising stories with revealing punchline that are not necessarily funny for their listeners. For example, a scientific discovery can be surprising and revealing, just like a well-crafted detective story, or certain political revelations, for the audience concerned in each case. This remark shows that ingredients other than the above belief revision mechanism are also necessary for laughter to occur. It includes elements such as the intention to make people laugh, lack of empathy or playing with taboos.

[3]This condition is automatically ensured by the usual revision postulate (KM3) (Katsuno and Mendelzon 1991): $K \circ \alpha$ is consistent as soon as α is consistent.

[4]However, absurd punchlines will not be revealing in the sense of the forthcoming Definition 3 due to postulates KM5 and KM6: (KM5): $(K \circ \varphi) \cup \{\psi\} \models K \circ (\varphi \wedge \psi)$; (KM6): if $(K \circ \varphi) \cup \{\psi\}$ is consistent then $K \circ (\varphi \wedge \psi) \models (K \circ \varphi) \cup \{\psi\}$.

We are now able to formalize a first example (from a French collection of "funny" stories (Nègre 1970)) expressed in natural language that we have transposed into logic. In this paper, we disregard the stage of the transition from natural language to logical representation. However, we acknowledge the fact that the listener uses default knowledge for understanding a story. Hence, in the examples, the listener knowledge is expressed by a knowledge base $\Sigma = P \cup \Delta$ consisting in a set P of propositional formulas and a set Δ of defeasible rules of the form $\alpha \rightsquigarrow \beta$ where $\alpha, \beta \in \mathcal{L}$. From the set Σ, one can compute (Benferhat, Dubois, and Prade 1997) a possibility distribution π_Σ on worlds that satisfies the constraints induced by Δ and assigns 0 to every world not satisfying P. For the sake of shortness, in the examples, all the revision mechanism is hidden behind the notation "$\Sigma \circ \alpha$", which is computed by the conditionalization of π_Σ by α (namely $\pi_\Sigma(\omega|\alpha) = \pi_\Sigma(\omega)$ if $\omega \models \alpha$, and 0 otherwise), yielding a revision operator satisfying Katsuno and Mendelzon postulates (Benferhat et al. 2002).

Example 1. *A man just got hit by a car. The driver gets out of the car and says, "You're lucky we're just in front of a doctor's office. Yes! Except the doctor is me!" Modeling:*
$\alpha = injured \wedge doctorNearby,$
$\beta = injured \wedge doctorHimself$
We assume that the listener has the following knowledge:

$$\Sigma = \left| \begin{array}{l} injured \wedge doctorNearby \rightsquigarrow treatedRapidly \\ doctorHimself \rightsquigarrow doctorNearby \\ injured \wedge \neg treatingDoctor \rightsquigarrow \neg treatedRapidly \\ injured \wedge doctorHimself \rightsquigarrow \neg treatingDoctor \end{array} \right.$$

Due to postulate (KM1) (Katsuno and Mendelzon 1991) $\Sigma \circ \alpha \models \alpha$. *Moreover, the most plausible models of* $\Sigma \circ \alpha$ *are satisfying* $\varphi = \alpha \wedge treatedRapidly$.
The most plausible models of $\Sigma \circ (\alpha \wedge \beta)$ *satisfy* ψ = $doctorNearby \wedge doctorHimself \wedge \neg treatingDoctor \wedge \neg treatedRapidly$. *The punchline is surprising since* $\varphi \wedge \psi \vdash \bot$.

Besides, the punchline explains α: $\Sigma \circ \beta \vdash injured \wedge doctorNearby$. *Indeed* $doctorNearby$ *is the part of the story that manipulated us by leading us to think that the wounded man was lucky.*

Note that Example 1 can also be seen as a cascading joke $(\varphi_1, \varphi_2, \varphi_3, \varphi_4)$ with $\varphi_1 = injured$, $\varphi_2 = lucky$, $\varphi_3 = doctorNearby$, $\varphi_4 = doctorHimself$. Indeed, saying "You're very lucky" after seeing that the person is injured can already provoke laughter. Then the revelation of being "in front of a doctor's office" helps to understand. Finally, the second part of the joke corresponds to what has already been analyzed. Let us note that in the first part of the cascading story, φ_2 is surprising but not revealing, it is φ_3 which explains φ_2. So for a cascading joke, the surprising and the revealing effects are not necessarily simultaneous.

The following property shows the importance of the narrative order in the joke.

Proposition 1 (Never give the punchline before the end)**.** *If* (α, β) *is a potentially funny story for a listener knowing* K*, then* (β, α) *is not potentially funny for that listener.*

Proof. Since (α, β) is potentially funny, then the punchline is revealing, that is, $K \circ \beta \nvdash \bot$ and $K \circ \beta \models \alpha$. But that prevents the story (β, α) from being surprising, and therefore potentially funny. Indeed, using (KM5) and (KM6)[4] (Katsuno and Mendelzon 1991), $K \circ (\beta \wedge \alpha) \equiv (K \circ \beta) \cup \{\alpha\}$ so $K \circ \beta$ is consistent with $K \circ (\beta \wedge \alpha)$. □

The following property indicates the virtue of brevity.

Proposition 2 (The shortest jokes ...)**.** *If* (α, β) *is a potentially funny story then adding* $\alpha' \in L$*, such that* $\alpha \nvdash \alpha'$ *can make the story* $(\alpha \wedge \alpha', \beta)$ *not potentially funny.*

Proof. Indeed here is a counterexample: let $\Sigma = \{b \rightsquigarrow f,\ p \rightarrow b,\ p \rightsquigarrow \neg f,\ b \wedge a \rightsquigarrow p\}$ a knowledge base expressing that: generally birds (b) fly (f), penguins (p) are birds, generally penguins (p) don't fly, most Antarctic birds (a) are penguins. Let us imagine a story (b, p), it is both surprising since $\Sigma \circ b \models b \wedge \neg p \wedge \neg a \wedge f$ and $\Sigma \circ (b \wedge p) \models b \wedge p \wedge \neg f$ and the "punchline" p is revealing since $\Sigma \circ p \models b$. However, by adding information $\alpha' = a$ that the bird is from Antarctica we can create an unsurprising situation with respect to $\beta = p$, so $(b \wedge a, p)$ is no longer a surprising story. Another counterexample consists in adding

information α' that is not related to β then we can obtain a punchline that is not completely revealing: let us consider $\alpha' = t$, where t means "is called Tweety" then $\Sigma \circ p \vdash b$ does not allow us to deduce that $\Sigma \circ p \vdash t$. □

A particular case of Proposition 2, when $\alpha' = \beta$, is another way to reveal the punchline before the end. This would correspond in Example 1 to specify that the driver has just knocked down a doctor in front of that same doctor's office. In this case, the punchline "Yes! Except that I am the doctor!" is no longer funny (because we already know that).

We can notice that Proposition 2 does not impose that the punchline be as short as possible, so we can refine the definition of "potentially funny story" into an "efficient story" by imposing minimality.

Definition 5 (conciseness of the punchline)**.** *Let $\alpha, \beta, \beta' \in \mathcal{L}$ and $\beta \not\equiv \beta'$, if $\beta \models \beta'$ then the story (α, β) is said to be* more efficient *than the story (α, β').*

In other words, the punchline is said more efficient if it is less precise and thus less specific. Here is another example from (Nègre 1970) for illustration.

Example 2. *A guy and a woman walk into a doctor's office. The doctor turns to the lady and says:*
- If you're sick, please take off your clothes ...
But the girl has manners. She is reluctant. She is looking down on the guy. So the doctor repeats:
- But, madam, please get undressed! I'm a doctor. There's no indecency, here!
Then she starts fidgeting and suddenly she bursts into tears. Disconcerted, the doctor asks the guy:
- What's wrong with your wife? Is she always this nervous?
- I don't know. I just met her in your waiting room...
$\varphi_1 = together$, $\varphi_2 = reluctant$, $\varphi_3 = \neg\, knowEachOther$

$$\Sigma = \left| \begin{array}{l} together \rightsquigarrow couple \\ couple \rightsquigarrow \neg reluctant \\ \neg knowEachOther \rightsquigarrow reluctant \\ \neg knowEachOther \rightarrow \neg couple \\ \neg couple \rightarrow \neg married \end{array} \right.$$

Let us compute $\Sigma \circ \varphi_1$, the most plausible models satisfy $\psi = together \wedge couple \wedge \neg reluctant$. $\Sigma \circ (\varphi_1 \wedge \varphi_2)$ satisfies $reluctant$. φ_2 is surprising since $\psi \wedge (\Sigma \circ (\varphi_1 \wedge \varphi_2)) \vdash \bot$. Moreover the punchline explains φ_2: $(\Sigma \circ \varphi_3) \models reluctant$.

Suppose now that the guy replied φ_3' ="Sorry, we are not married". Clearly $\varphi_3 \models \varphi_3'$. Hence, the new punchline φ_3' is less efficient. Note that the new story is no more funny since $(\Sigma \circ \varphi_3') \not\models reluctant$, which could have been obtained if Σ would contain $\neg married \rightsquigarrow \neg couple$.

Finally, in order to give substance to Raccah's intuition (2016) that the greater the surprise and inevitability, the funnier the joke, we may propose the use of two quantities, in order to build a partial order between jokes for the same listener:

Definition 6 (ingredients for humor)**.** *Given a base $\Sigma = \Delta \cup P$, a joke (α, β) is associated to the possibilistic levels:*
- a surprise level: $Incons((\Sigma \circ \alpha) \cup (\Sigma \circ (\alpha \wedge \beta)))$
- a revealing level: $N(\alpha | \Sigma \circ \beta)$

This definition assumes that the default rules Δ of Σ are encoded in possibilistic logic (Benferhat, Dubois, and Prade 1998), where $Incons(A)$ represents the level of inconsistency of the possibilistic knowledge base A. The second term is the conditional necessity of α knowing $\Sigma \circ \beta$, reflecting the ineluctability of the revelation.

In addition, other elements may contribute to the degree of humor: the register, the comic effects of the narrator, the mood of the listener, her ability to inhibit her emotions, the cognitive effort required to understand the joke ...

Besides, we might have considered iterated revision (Konieczny and Pino Pérez 2000) and abductive expansion (Pagnucco 1996). However, as we wanted to capture cognitive aspects of the understanding of joke, we have privileged the simplicity of basic revision theory where epistemic states may be drastically

changed in the revision process. More precisely the main idea is to select the best epistemic state rather than carrying the whole preference ordering as in iterated revision. Indeed the surprise would not be the same if the listener had in mind all the potential explanations. In that respect, a similar argument applies to abductive expansion whose aim is to provide reasons for the new epistemic state.

5 Conclusion

The proposed approach highlights the key role of belief revision and surprises in the mechanism underlying jokes. It has been illustrated with a few examples. It is clear that a validation of the relevance of these ideas would require their verification on a large corpus of jokes, although one may feel that many jokes actually work this way. As far as we know, this opens a new direction in the study of jokes from an artificial intelligence point of view.

References

Alchourrón, C. E.; Gärdenfors, P.; and Makinson, D. 1985. On the logic of theory change: Partial meet contraction and revision functions. *The J. of Symbolic Logic* 50(2):510–530.

Anonymous. 2008. *Va te marrer chez les Grecs (Philogelos)*. French version, afterword by A. Zucker. Mille et une Nuits.

Attardo, S., and Raskin, V. 1991. Script theory revis(it)ed: Joke similarity and joke representation model. *Humor-International Journal of Humor Research* 4(3-4):293–348.

Benferhat, S.; Dubois, D.; Prade, H.; and Williams, M.-A. 2002. A practical approach to revising prioritized knowledge bases. *Studia Logica* 70(1):105–130.

Benferhat, S.; Dubois, D.; and Prade, H. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. *Artif. Intell.* 92(1-2):259–276.

Benferhat, S.; Dubois, D.; and Prade, H. 1998. Practical handling of exception-tainted rules and independence information in possibilistic logic. *Appl. Intell.* 9(2):101–127.

Dubois, D., and Prade, H. 1988. *Possibility Theory. An Approach to Computerized Processing of Uncertainty*. Plenum Press (with H. Farreny, R. Martin-Clouaire, C. Testemale).

Dubois, D., and Prade, H. 1991. Epistemic entrenchment and possibilistic logic. *Artif. Intell.* 50(2):223–239.

Dupin de Saint-Cyr, F., and Lang, J. 2011. Belief extrapolation (or how to reason about observations and unpredicted change). *Artificial Intelligence* 175:760–790.

Dupin De Saint-Cyr, F., and Prade, H. 2019. Peut-on rire en IA? Révision de croyances et modélisation de plaisanteries. In *Conf. Nationale d'Intelligence Artificielle*, 113–121.

Fagin, R., and Halpern, J. Y. 1994. Reasoning about knowledge and probability. *J. of the ACM* 41(2):340–367.

Freud, S. 1961. Humour. In *The Standard Edition of the Complete Psychological Works*, volume 21. 161–166.

Gärdenfors, P. 1988. *Knowledge in Flux*. The MIT press.

Grove, A. 1988. Two modellings for theory change. *Journal of philosophical logic* 157–170.

Hempelmann, C.; Raskin, V.; and Triezenberg, K. 2006. Computer, tell me a joke... but please make it funny: Computational humor with ontological semantics. In Sut-cliffe, G., and Goebel, R., eds., *Proc. 19th Int. Florida Artif. Intell. Res. Soc. Conf., FLAIRS'06*, 746–751. AAAI Press.

Katsuno, H., and Mendelzon, A. O. 1991. Propositional knowledge base revision and minimal change. *Artificial Intelligence* 52(3):263–294.

Katsuno, H., and Mendelzon, A. O. 2003. On the difference between updating a knowledge base and revising it. In Gär-denfors, P., ed., *Belief revision*. Cambridge U.P. 183–203.

Konieczny, S., and Pino Pérez, R. 2000. A framework for iterated revision. *J. Applied Non-Class. Logics* 10:339–367.

Lorini, E., and Castelfranchi, C. 2007. The cognitive structure of surprise: looking for basic principles. *Topoi* 26(1):133–149.

Minsky, M. 1984. Jokes and the logic of the cognitive unconscious. In Vaina, L., and Hintikka, J., eds., *Cognitive Constraints on Communication - Representations and Processes*. D. Reidel Publ. Comp. 175–200.

Nègre, H. 1970. *Dictionnaire des Histoires Drôles*. Fayard.

Pagnucco, M. 1996. *The role of abductive reasoning within the process of belief revision*. Univ. Sydney Dissertation.

Raccah, P.-Y. 2016. Humour et métaphore: quelques éléments d'une analogie pour la construction d'un sens inattendu. illustration sur un corpus de citations de G. B. Shaw. *Revue de Sémantique et Pragmatique* 39:75–94.

Raskin, V. 2008. *The Primer of Humor Research*, volume 8. Walter de Gruyter.

Shackle, G. L. S. 1961. *Decision, Order and Time in Human Affairs*. UK: (2nd edition), Cambridge University Press.

Shahaf, D.; Horvitz, E.; and Mankoff, R. 2015. Inside jokes: Identifying humorous cartoon captions. In *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 1065–1074.

Stock, O., and Strapparava, C. 2003. Getting serious about the development of computational humor. In *Proc. IJCAI'03*, volume 3, 59–64.

Williams, M.-A. 1996. Aesthetics and the explication of surprise. *Languages of Design. Formalisms for word, image and sound* 3:145–157.

Winslett, M. 1988. Reasoning about action using a possible models approach. In *Proc. AAAI'88*, 89–93.