FUZZY PUNCTUATION OR THE CONTINUUM OF GRAMMATICALITY

by

Thomas T. Ballmer

Memorandum No. ERL-M590

February 1976

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Research sponsored by the U. S. Army Research Office Grant DAHCO4-75-G0056.

I. The aim of this short paper is to eludicate once more the rule of punctuation signs in a theory of grammar. Moreover the notions of fuzzy category and fuzzy punctuation are introduced to render the power of grammars based on punctuation signs more explicit.[1]

----

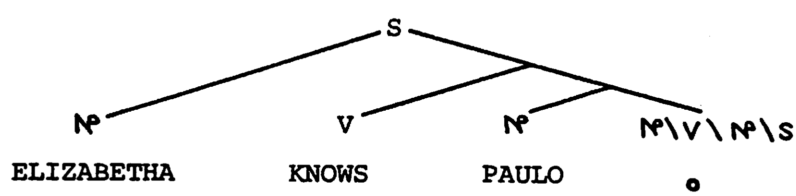
1) The reader not acquainted with such grammars should consult "Sprachrekonstruktionssysteme" ("Language Reconstruction Systems") Ballmer (1974), the Dissertation of the author.

II. A punctuation based grammar (PBG) is a grammar which characterizes the grammatical strings of morphemes (or words) with help of certain of these morphemes, called punctuation signs (or less colorfully: secondary morphemes). Punctuation signs determine sentence patterns, they are--syntactically--sentence patterns. As a clear case consider the following sentence:

(1) ELIZABETHA KNOWS PAULO.

The sentence pattern of this sentence is $NP\ V\ NP$, if only the primary words are taken into account. If all words are considered the sentence pattern is $NP\ V\ NP\ \circ$. The whole object is a sentence, i.e. of category S. Hence, in categorial grammar, the category of the punctuation sign is $NP \backslash V \backslash NP \backslash S$. Hence we have:

(2)

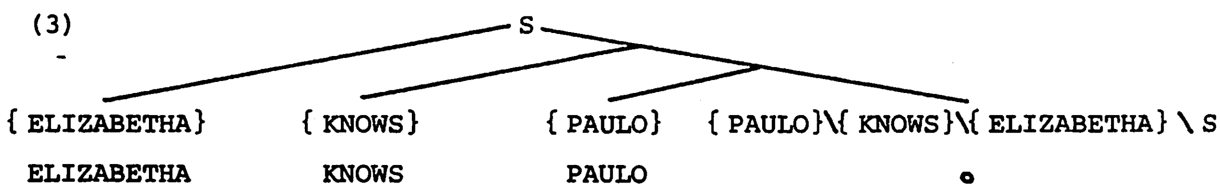


The sentence pattern determined by punctuations are seen to be rich enough to account for every possible succession of categories. Every such succession can be singled out. Take the general case of n categories $A_1, \ldots, A_n$. The categories $A_{j1} \backslash A_{j2} \backslash \ldots \backslash A_{jn} \backslash S$ account for every possible case, if $(j_1, \ldots, j_n)$ are taken to be permutations of $(1, \ldots, n)$. The lexicon should contain exactly those punctuations with sentence patterns appropriate for the language, of course.

IIa. The essential question for a grammar is to determine the categories of the primary morphemes (words).

The most trivial grammar is that which has the finest grid of categories: Every word (or even better, every disambiguated word) constitutes its own category. The lexicon contains beside these primary words also the punctuation signs, which then determine every possible sentence of the language by simply providing a slot for every single sentence of that language. There is no guarantee that this lexicon is finite, nor does such a lexicon allow for any "generalizations." It is, moreover, virtual, in the sense that it is not explicitly statable in simple terms. Its advantage is its correctness and completeness, if only virtual. As an example of an analysis in such a grammar take the following:

(3)



Call this grammar G_{∞}. It has possibly infinitely many rules (punctuation signs), one for each sentence of the language. It is very simple, it assigns to each sentence a similar (namely rightbranching) structure, it is, in principle, very easily described. Though simple, it is only of marginal theoretical interest.

IIb. Another (very) uninteresting grammar is G_0 which has no category besides S. No restrictions on the succession of the morphemes (words) are statable. More interesting grammars lie between G_{∞} and G_0. A grammar type, which we could call classical, makes use of finitely many sharp categories. The term sharp is meant to express that for each morpheme (word) of the language it is either fully true or completely false that it belongs to a definite category.

III. A moment's reflection--a reflection which could have been performed and widely accepted years ago--shows that classical grammars are bound to fail. The tiny differences of language force one to subdivide categories further and further in order to account especially for degrees of grammaticality. Thus the program to stick to few finitely many categories cannot be successful in that framework.

In deviation from the concept of sharp categories,[2] one could introduce

---

2) Cf. Ross (1972), Ross (1973).

---

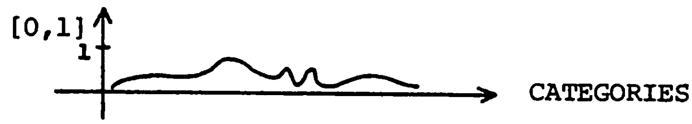
fuzzy categories. The hope for few, finitely many basic categories is then again justified. For this reason we shall try to work this proposal out in somewhat more detail. The approach based on punctuation signs coding sentence patterns lends itself most easily to this objective.

As far as I can see Ross's analysis does not yet provide for a conclusive argument that grammatical categories are fuzzy, or squishy in his own words. In his example he subdivides the alleged continuum between, say, the sentences and nounphrases into approximately eight segments. But nothing prevents the view that there are just about eight sharp categories.

What could it mean that there is an underlying continuum. The observation that the sample sentences considered vary their degree of grammaticality accounts not even for the fact that the category of sentences is fuzzy: There is only a small finite number of categories in play.[3]

---

3) As it appears we have to distinguish strictly, at least conceptually, between the degree to which a sequence of morphemes (or words) is grammatical and the degree to which that sequence is a sentence or an other category. A possible way to look at these things is to say that an expression is assigned a function from categories to real numbers between 0 and 1 characterizing the degree to which the expression belongs to that category.



The maximum value of the function for all categories is taken to be the (overall) grammaticality of the expression. The value for the argument C is taken to be the C-grammaticality. Linguists are predominantly interested in S-grammaticality. Observe that we distinguish between the context independent notion of grammaticality and the context dependent notion of acceptability.

III a. A proper way to prove that there is a (quasi) continuum of categories to show that iterative alterations of categories can be made, and that every arbitrary degree of category can be approximated. It suffices to show that this is the case for at least one category, however.

Before we discuss this purely syntactic issue, let us look at a semantic example of fuzziness where approximation of degrees of truth, say, seems to be possible with linguistic means. Consider the following sequence:

(4) true
nearly true but not quite true
more true than nearly true but less than not quite true
more than more true than nearly true but less than less than not quite true
.
.
.

This procedure of defining a degree of truth reminds somewhat of the Dedekind cut procedure, in fact it is a nonmetrical generalization thereof. Dedekind cuts are one representation of the reals, i.e. the continuum. We thus arrived at linguistically determining a continuum of truthvalues, when passing to the limit.

III b. Now let us consider the case of the continuum of categories. Three examples are of interest to us in that respect:

1. Ross (1973) gives the following sequence of sentences:

$$\text{(5a) THEY ALL} \left\{\begin{array}{l} \text{THREE} \\ \text{?NINE} \\ \text{*729} \end{array}\right\} \text{HATE SPAM.}$$

In extension to these we could speculate:

$$\text{(5b) THEY ALL} \left\{\begin{array}{l} \text{*ZERO} \\ \text{*ONE} \\ \ \text{BOTH} \\ \ \text{TWO} \\ \ \text{THREE} \\ \ \text{SEVEN} \\ \ \text{EIGHT} \\ \text{?NINE} \\ \text{??TEN} \\ \text{???ELEVEN} \\ \quad \vdots \\ (N-8)\cdot ?_{\ldots N \ldots} \end{array}\right\}$$

This allows us to simply inherit the continuum of numbers into the continuum of the category of sentences, if we can assume that the degrees of grammaticality depend homeomorphically on the numbers.

2. Again, Ross (1973) gives the following sequence of sentences:

(6) a. THAT NOT EVERYONE WILL REFUSE OUR OFFER IS EXPECTED.
b. ? FOR NOT EVERYONE TO REFUSE OUR OFFER IS EXPECTED.
c. ? UNDER WHAT CIRCUMSTANCES NOT EVERYONE WILL REFUSE OUR OFFER IS THE SUBJECT OF A HEATED DEBATE.
d. ?? NOT EVERYONE REFUSING OUR OFFER WAS EXPECTED.
e. ?* NOT EVERYONE'S REFUSING OUR OFFER WAS EXPECTED.
f. ** NOT EVERYONE'S REFUSING OF OUR OFFER WAS A SURPRISE.
g. ** NOT EVERYONE'S REFUSAL OF OUR OFFER WAS A SURPRISE.

As we stated above this series does not tell us anything about the fuzziness of categories or of their degrees of grammaticality. All that is demonstrated by such exemplaric tableaus of different degrees of grammaticality is that at least seven (in our case!) sentential and/or nounphrase categories are in play.

In order to demonstrate the existence of a continuum of categories we have to provide for a set of sentences whose degrees of grammaticality is dense in the interval [0,1] (or some subinterval) and argue that it is methodologically meaningful to take the completion of that set! As is known the completion of a dense set is a continuum.

My claim is that starting from one of the sentences of Ross's list, say (6.e) we could get a set of sentences whose degrees of grammaticality is dense in some subinterval of [0,1]. The idea is to systematically weaken and strengthen the NOT occurring in all these sentences (6).

(7) 1. NOT
   2. NOT QUITE BUT NEARLY NOT
   3. LESS THAN NOT QUITE BUT MORE THAN NEARLY NOT
   ⋮ ⋮

The fact is now that such series of expressions fill in the different degrees of grammaticality between

(8a) ✓ EVERYONE'S REFUSING OUR OFFER WAS EXPECTED.

and

(8b) ?* NOT EVERYONE'S REFUSING OUR OFFER WAS EXPECTED.

3. Heavy noun phrase constructions provide a further argument for the continuum of categories. It is of a slightly other sort: the grammaticality of certain sentences depends on the length of certain NP s. The length of NP s is a discrete entity. Nevertheless it is reasonable to postulate an underlying continuum, because of a simplicity argument. (Another way to

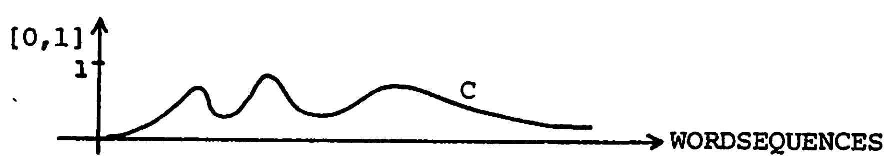
connect these facts with the existence of a continuum would be the following. That (heavy) NPs are discrete is an abstraction. Remembering that sounds are underlying we could try to establish a connection with the grammaticality of the sentences in question and the duration of uttering these NPs. Then the continuum of time inherits the continuum of degrees of grammaticality. In fact, if you test a couple of examples you find convincing evidence that grammaticality is strongly correlated with the duration of speech.)

To support the claim based on the heavy NP construction let us consider the following sequence of sentences:

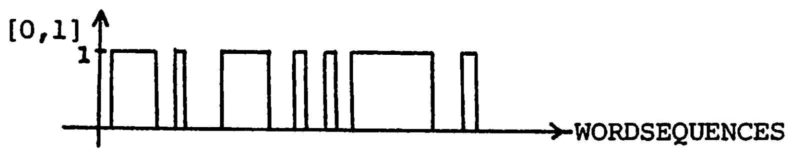
(9) *I FOUND DISGRACEFUL RON.
??*I FOUND DISGRACEFUL RON LYING.
?*I FOUND DISGRACEFUL RON LYING TO US.
??I FOUND DISGRACEFUL RON LYING TO US LIKE THAT.
%?I FOUND DISGRACEFUL RON LYING TO US LIKE THAT AGAIN.
?I FOUND DISGRACEFUL RON LYING TO US LIKE THAT AGAIN AND AGAIN.

IV. We have now demonstrated, and as I hope persuasively enough, that we need continuous grammatical categories. This being so, every theory of grammar based on the assumption that there are few, finitely many categories is bound to fail. Of course we hope that there are few basic categories--or at least a number of basic categories which is finitely characterizable--and that the rest is accounted for by taking different degrees of these basic categories. This is in fact what we assume henceforth until the contrary is demonstrated.

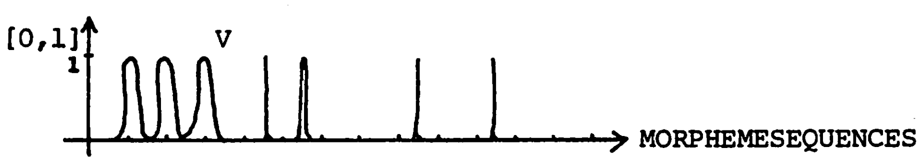
IV a. Grammatical categories C are taken to be functions from the set of morpheme-sequences (wordsequences) to the interval [0,1], the degrees of grammaticality of these wordsequences as basic categories:

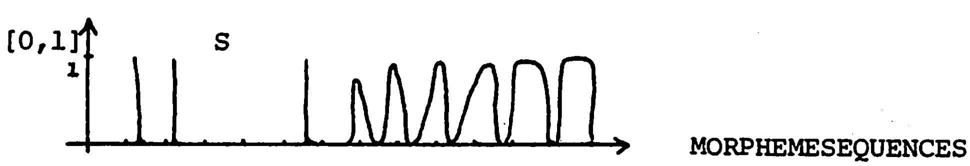


Classically, grammatical categories are characteristic functions, i.e. they take only the values 0 and 1. No (nontrivial) continuous categories are expressible that way. Moreover in the classical case the basic categories are thought to be disjoint, e.g. no V is an NP. This, of course, is a mere theoretical constraint, because it can be obtained by appropriate renaming of the categories. An example of a classical category would be:



There are one-morpheme sequences, two-morpheme sequences, . . . , n-morpheme sequences. NP , V, ADJ, ADV are categories which are truer of shorter morpheme sequences, S is a category which is, in general, truer of longer sequences. But this is, of course, only a tendency. This tendency, however, is true also for fuzzy categories. If we arrange wordsequences according to their length, the following are two typical functions:



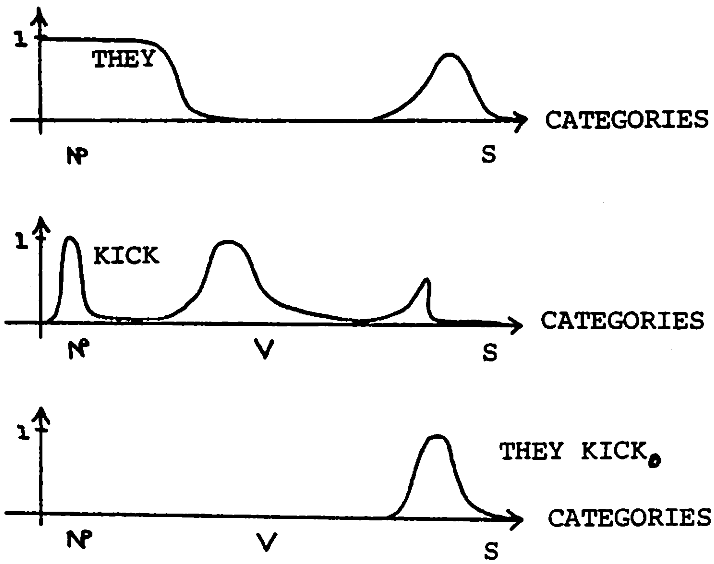


As is seen there is a continuum of possible functions from morphemesequences to the interval [0,1], i.e. a continuum of possible categories.

IV b. Another picture of the fuzziness of categories arises, if specific morphemesequences are considered and the question asked to what degree is it a given category, included fuzzy categories. As such a morphemesequence is a function from the set of categories to the interval [0,1].

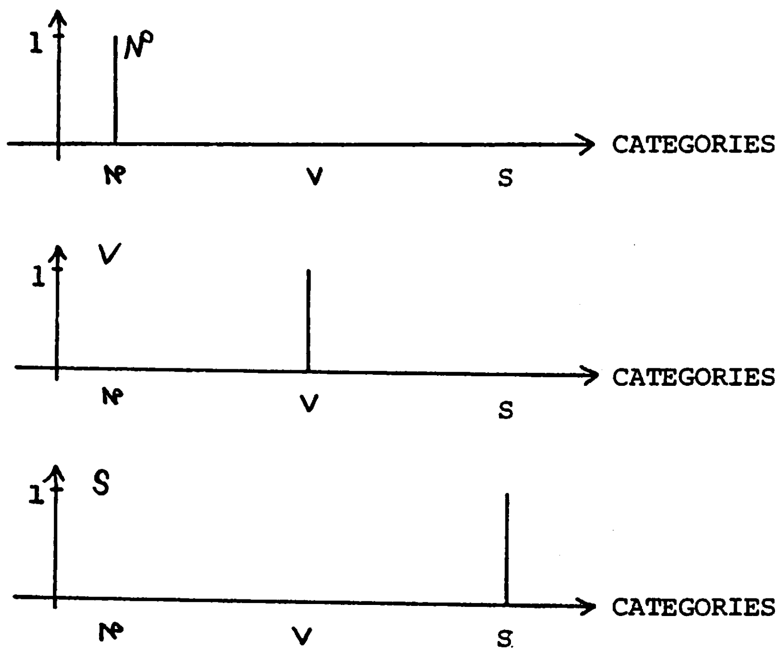
Take as rough examples:

(10)



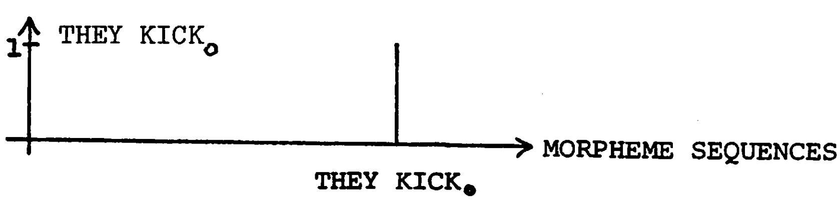
The problem is now to determine the category of the punctuation sign $_o$! Before we do that, let us point out that categories may be considered as special morpheme (word) sequences, e.g., NP , V and S are the following functions.

(11)



As a matter of fact the converse is also true. Morpheme (word) sequences are special categories, as the following function graph suggests:

(12)



The functions represented in (11) and (12) are called δ-functions. Their value is 0 for every argument except one, where their value is 1. Let us denote the δ-function which is 1 at the argument C and 0 elsewhere δ_C. Hence

(13) $$\delta_C(X) = \begin{cases} 1 & X = C \\ 0 & X \neq C \end{cases}$$

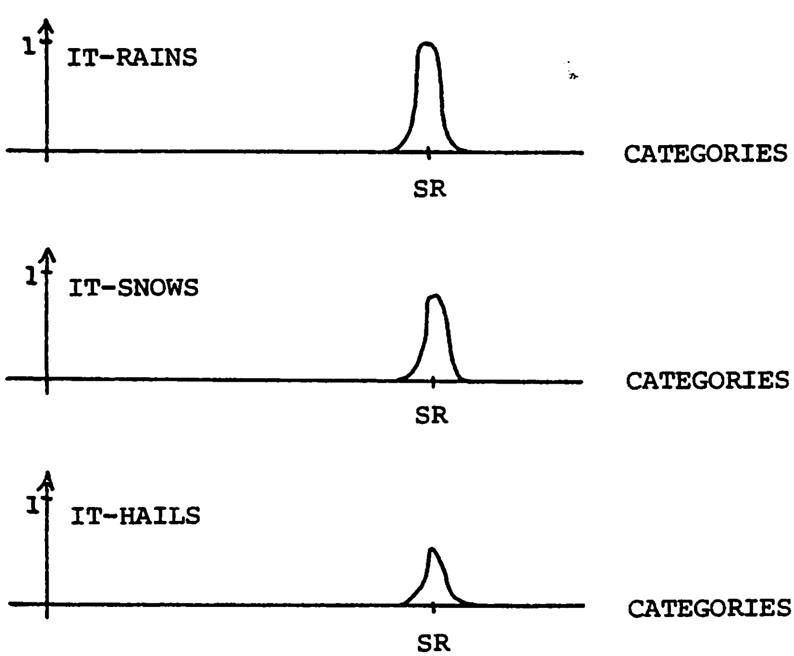
V. Considerations for the Appropriate Categories of Punctuation

In order to find out about the category of punctuations let us consider, as an example, a two-word sentence, consisting of a propositional expression and a full-stop, e.g.

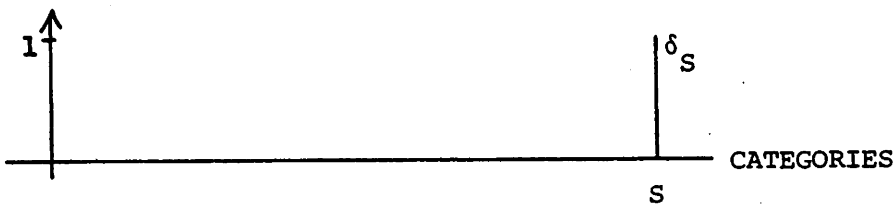
(11) IT-RAINS. (English dialect of philosophers and logicians).
?IT-SNOWS.
??IT-HAILS.

Because the different propositional expressions are sentences to more or less degrees, the following are reasonable approximations to their category functions, where SR stands for the category sentence radical.

(12)



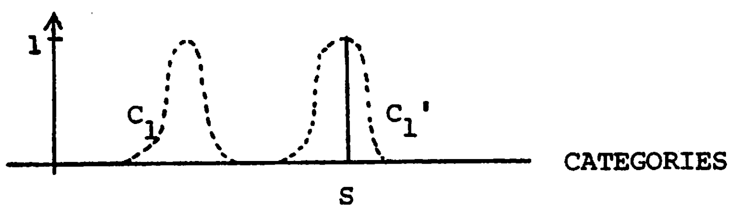
Let these category functions of IT-RAINS, IT-SNOWS, IT-HAILS be abbreviated by C_1, C_2, C_3 respectively. The category function for an ideal sentence is a delta function at the argument of S, i.e. a function which is zero everywhere, with the exception of S, where it is 1.



How can we state the fact that a perfect sentence radical gives rise to a perfect sentence and a crummy sentence radical gives rise to a crummy sentence only?

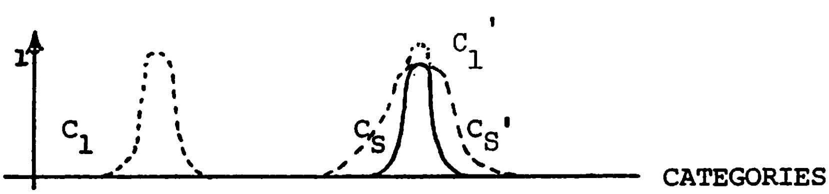
The basic idea is to shift the category functions of the sentence radical of, say, IT-SNOWS on top of the category function of the perfect sentence and compare the two category functions appropriately:

(13)



In (13) three category functions are represented simultaneously: C_1, the shifted C_1, which is called C_1', and δ_S. The comparison of C_1' and S should result in a function C_S' which is as broad as C_1' and which is as high as C_1'. In fact C_S' should be equal to C_1', because the filter category function is a perfect sentence, i.e. no fuzziness is contributed from <u>it</u>. In other words the full stop itself does not, in this case, add to the fuzziness of grammaticality. If it would, a case which is illustrated by the following figure,

(14)



would arise. The category function C_S' is then both depressed and broadened by the amalgamating the function C_S and C_1'.

Our problem of finding the effect of a possibly fuzzy punctuation on the possibly fuzzily categorized words it accepts from the left can be formulated as follows:

1. How is the shifting of the category functions of the words (morphemes) formulated formally?

2. How are the amalgamation of the category function of the words (morphemes) and the filter category function of the punctuation formulated formally?

3. How is one to proceed when dealing with more complex sentence patterns?

A short reflection shows that the amalgamation cannot be a simple logical conjunction, which would be simply taking the minimum value of the two functions at every argument category. In (14), C_S and C_1' are to be amalgamated to C_S'. In contrast to the amalgamation of δ_S with C_S, where the full height of C_1' is kept for C_S' (cf. (13)), an additional depression occurs in the case (14), which originates from C_S. C_S represents thereby the fuzziness stemming from the punctuation. In case (14) the conjunction of C_S and C_1' is equal to C_S, and no additional depression occurs. Hence the inadequacy is apparent.

Fuzzy Logic and Manyvalued Logic

What has to be done is what Zadeh (1972) calls support fuzzification. In order to introduce his idea let us very briefly summarize how fuzzy sets are represented as scalar products or more generally integrals, and some essentials of fuzzy logic. For this reason we present the following translational characterization and interpretation for the logical language L_{FUZZ}. The function of the degree of truthfulness μ is hereby a function from propositions and worlds to values $\varrho\alpha,i$ (or σ) of the interval $[0,1]$.

(15) $h(\ulcorner\alpha\urcorner, i) \equiv (\mu(\ulcorner\alpha\urcorner, i) = \rho_{\alpha,i})$
$\mu(\ulcorner\pi\urcorner, i) \equiv \rho_{\pi,i}$
$\mu(\ulcorner\neg\alpha\urcorner, i) \equiv (1-\mu(\ulcorner\alpha\urcorner, i))$
$\mu(\ulcorner\alpha\wedge\beta\urcorner, i) \equiv \min(\mu(\ulcorner\alpha\urcorner, i), \mu(\ulcorner\beta\urcorner, i))$
$\mu(\ulcorner\alpha\vee\beta\urcorner, i) \equiv \max(\mu(\ulcorner\alpha\urcorner, i), \mu(\ulcorner\beta\urcorner, i))$
$\mu(\ulcorner\alpha\rightarrow\beta\urcorner, i) \equiv \iota\rho((\mu(\ulcorner\alpha\urcorner, i) \leq \mu(\ulcorner\beta\urcorner, i) \rightarrow \rho=1) \wedge (\mu(\ulcorner\beta\urcorner, i) < \mu(\ulcorner\alpha\urcorner, i) \rightarrow \rho=0))$
$\mu(\ulcorner\Pi(\gamma_1, \ldots, \gamma_n)\urcorner, i) \equiv \rho_{\Pi,\gamma_1,\ldots\gamma_n,i}$

This translational interpretation h is two-valued. $h(\ulcorner\alpha\urcorner, i)$ is true just in case the degree $\mu(\ulcorner\alpha\urcorner, i)$ truthfulness of $\ulcorner\alpha\urcorner$ in i has a determinate value σ, which may depend on $\ulcorner\alpha\urcorner$ and i, and which is denoted therefore $\rho\alpha,i$. That this is <u>no more circular</u> than any other modeltheoretic truth-condition consider the following example:

$$h(\ulcorner\pi_1\wedge\pi_2\urcorner, i) \equiv (\mu(\ulcorner\pi_1\wedge\pi_2\urcorner, i) = \rho_{\pi_1\wedge\pi_2,i})$$
$$\equiv (\min(\mu(\ulcorner\pi\urcorner_1, i), \mu(\ulcorner\pi\urcorner_2, i)) = \rho_{\pi_1\wedge\pi_2,i})$$
$$\equiv (\min(\rho_{\pi}, _i, \rho_{\pi_2,i}) = \rho_{\pi_1\wedge\pi_2,i})$$

This states that $h(\ulcorner\pi_1\wedge\pi_2\urcorner, i)$ is true just in case min $(\rho_{\pi_1,i}, \rho_{\pi_2,i}) = \rho_{\pi_1\wedge\pi_2,i}$, a sensible result. The translational interpretation h can be given a multivalue flavor, however: $h(\ulcorner\alpha\urcorner, i)$ is true to the degree σ iff $\mu(\ulcorner\alpha\urcorner, i) = \sigma$. This is what we shall do henceforth. We therefore identify fuzzy logic with a special kind of multivalued logic!

A translational characterization of a logical language L_{MVAL} corresponding to L_{FUZZ} is the following: h in this case maps expressions and world in the whole interval [0,1]. This is not anymore an intepretation by translation in predicate calculus. As we have seen for L_{FUZZ}, a two valued interpretation is conceivable in principle, but/as I take it, is less intuitive and slightly redundant.

(16) $h(\ulcorner\pi\urcorner, i) \equiv \rho_{\pi,i} \qquad \epsilon[0,1]$
$h(\ulcorner\neg\alpha\urcorner, i) \equiv 1 - h(\ulcorner\alpha\urcorner, i)$

$$h(\ulcorner\alpha\wedge\beta\urcorner,i) \equiv \min(h(\ulcorner\alpha\urcorner,i),h(\ulcorner\beta\urcorner,i))$$
$$h(\ulcorner\alpha\vee\beta\urcorner,i) \equiv \max(h(\ulcorner\alpha\urcorner,i),h(\ulcorner\beta\urcorner,i))$$
$$h(\ulcorner\alpha\rightarrow\beta\urcorner,i) \equiv \iota\varrho((h(\ulcorner\alpha\urcorner,i)\leq h(\ulcorner\beta\urcorner,i)\rightarrow\varrho=1)\wedge(h(\ulcorner\beta\urcorner,i)<h(\ulcorner\alpha\urcorner,i)\rightarrow\varrho=0))$$
$$h(\ulcorner\Pi(\gamma_1,\ldots,\gamma_n)\urcorner,i) \equiv \varrho_{\Pi,\gamma_1,\ldots,\gamma_n,i}$$

Our main interest is now to have a handy way to represent the functions $\varrho_{\pi,i}$ and $\varrho_{\Pi,\gamma_1,\ldots,\gamma_n,i}$. As is seen the <u>world index</u> i occurs parallel to individuals γ_ν. In the following treatment, no special treatment of that index will be given, because all that is said for individuals applies as a special case to <u>it</u>.

Thus we are left with the problem of finding a handy way to represent a σ-valued function ϱ whose values have the form $\varrho_{\Pi,\gamma_1,\ldots,\gamma_n}$. The standard way to do this is, of course, to use λ-calculus. Hence ϱ, in the case described, is

(17) $\lambda\gamma_1\ldots\lambda\gamma_n\,\varrho_{\Pi,\gamma_1,\ldots,\gamma_n}$

This is an adequate way to represent a fuzzy relation. In the special case of one individual argument only, this reduces to

(17') $\lambda\gamma\varrho_{\Pi,\gamma}$

Because $\varrho_{\Pi,\gamma}$ for specific individuals, say $\gamma_1,\gamma_2,\ldots$, takes a value between 0 and 1, (17') is an appropriate way to represent a fuzzy set. To simplify matters, let us omit the function symbol $\ulcorner\varrho\urcorner$, because it occurs invariantly in all the formulas in question now! Then (17) and (17') read respectively:

(18) $\lambda\gamma_1\ldots\lambda\gamma_n\Pi(\gamma_1,\ldots,\gamma_n)$ (λ - notation)
(18') $\lambda\gamma\Pi(\gamma)$

As is seen, no overt syntactical distinction exists between fuzzy and non-fuzzy relations or sets! This is a very nice fact, because it allows us to take nonfuzzy logics and look at them as fuzzy ones. It is just a matter of interpretation. The first and the last clause of (16) can be changed appropriately:

$$(16^*)\ h(\pi,i) = \pi(i) \qquad \varepsilon\,[0,1]$$

$$h(\Pi(\gamma_1,\ldots,\gamma_n),i) \equiv \Pi(\gamma_1,\ldots,\gamma_n,i) \qquad \varepsilon\,[0,1]$$

If the (relational) predicate Π is marked for the number of places of its arguments, the following equivalence, which reflects extensionality, may be introduced:

$$(19)\ \lambda\gamma_1\ldots\lambda\gamma_n\ \Pi^n(\gamma_1,\ldots,\gamma_n) = \Pi^n \qquad \text{(intensive notation)}$$

and especially

$$(19')\ \lambda\gamma\Pi^1(\gamma) = \Pi^1 \qquad \text{(intensive notation)}$$

Let us come back to the representation (18') of a fuzzy set. (18') is a very succinct way to present the fuzzy set in question. We cannot read off explicitly to what degree a <u>special individual</u> occurs in the set, however. An explicit way to do this would be to say, γ_1 is to degree σ_1 in Π, γ_2 is to degree σ_2 in Π, and so on, or formally:

$$(20)\ \lambda\gamma\Pi\gamma = \{\sigma_1\gamma_1,\sigma_2\gamma_2,\sigma_3\gamma_3,\ldots\} \qquad \text{(extensive notation)}$$

If the σ_ν are just 0 or 1, the explicit form of a non-fuzzy set would stand on the right hand side of (20). There is a small difference to classical sets, however, in that the non-occurring individuals are listed also in this representation, namely with a coefficient 0. As a matter of fact σ_1 equals $\Pi\gamma_1$, σ_2 equals $\Pi\gamma_2$ and so forth. Hence the following equation holds:

$$(21)\ \lambda\gamma\Pi\gamma = \{\langle\Pi,\gamma_1\rangle\gamma_1,\langle\Pi,\gamma_2\rangle\gamma_2,\ldots.\} \qquad \text{(extensive notation)}$$

We have used the bracket notation $\langle\Pi,\gamma_\nu\rangle$ instead of simple juxtaposition just because of reasons of perspicuity.

A notational variant (which we shall dub $\int$-representation) for the two representations of (fuzzy) sets in (20/21) are the following:

(22a) $\lambda\gamma\Pi\gamma \;:\Leftrightarrow:\; \begin{cases} \sum_{\gamma} \Pi(\gamma) & \text{for a discrete universum U of individuals} \\ \int \Pi(\gamma)\, d\gamma & \text{for a discrete or continuous Universum U of individuals} \end{cases}$

(22b) $\{\sigma_1\gamma_1, \sigma_2\gamma_2, \ldots\} \quad :\Leftrightarrow: \quad \sigma_1\gamma_1 + \sigma_2\gamma_2 + \ldots$
$\{\langle\Pi,\gamma_1\rangle\gamma_1, \ldots\} \quad :\Leftrightarrow: \quad \langle\Pi,\gamma_1\rangle\gamma_1 + \langle\Pi,\gamma_2\rangle\gamma_2 + \ldots$

This notational variant will be made use of very much from now on. For several reasons which will become clear subsequently, the view of λ-abstraction as integration (or sum) has a formidable intuitive appeal. Both for the implicit and explicit representation of (fuzzy) it will prove useful to change the notational conventions. Many consequences cannot be pursued in this paper, however, the topic of which is to provide for fuzzy grammatical categories after all. A more thorough study will be undertaken elsewhere.

The change in notational conventions can easily be generalized to relational (fuzzy) predicates. Multiple integration and summation will occur then.

The right hand side of (20) can be rewritten as

(23) $\{\sigma\gamma_1\} \cup \{\sigma\gamma_2\} \cup \{\sigma\gamma_3\} \cup \ldots$

A comparison with (24)

(24) $\sigma_1\gamma_1 + \sigma_2\gamma_2 + \sigma_3\gamma_3 + \ldots$

shows that the γ_ν in (24) correspond to singletons in (23) and that the plus operations correspond to unions and hence to taking the maximum! A singleton $\{\gamma\}$ is represented as $1{:}\gamma$.

Representation (22b) allows us to introduce the notion of <u>support-fuzzification</u> (or s-fuzzification (Zadeh, 1972)). At the base of these approaches is the process of <u>individual fuzzification</u> which transforms a singleton $1\cdot\gamma$ into a fuzzy set $\underset{\sim}{\gamma}$ which may be concentrated around γ or which may be shifted away from γ. To place in evidence the dependency of $\underset{\sim}{\gamma}$ of γ, $\underset{\sim}{\gamma}$ will be written as $K(\gamma)$. The (fuzzy) set K will be referred to as the <u>kernel</u> of the fuzzification. We introduce the following special types of kernels):

(25a) K_M^L: $K_M^L(\gamma) \begin{cases} =M \text{ for } \gamma=L \\ \approx M \text{ for } \gamma\approx L \\ \dots \end{cases}$ shifter (shifts from L to M)

(25b) K_M^M: $K_M^M(\gamma) \begin{cases} =M \text{ for } \gamma=M \\ \approx M \text{ for } \gamma\approx M \\ \dots \end{cases}$ squeezer (deformator)

(25c) I_M: $I_M(\gamma) \begin{cases} =M \text{ for } \gamma=M \\ =0 \text{ for } \gamma\neq M \end{cases}$ δ-kernel

(25d) I_M^L: $I_M(\gamma) \begin{cases} =M \text{ for } \gamma=L \\ =0 \text{ for } \gamma\neq L \end{cases}$ δ-shifter

Fuzzification $\underset{\sim}{\Pi}$ of a set Π is now easily defined:

(26) $\Pi = \sigma_1 K(\gamma_1) + \sigma_2 K(\gamma_2) + \dots$ if $\Pi = \sigma_1\gamma_1 + \sigma_2\gamma_2 + \dots$

$\Pi = \int \Pi(\gamma)\,dK(\gamma)$ if $\Pi = \int \Pi(\gamma)\,d\gamma$

Two easy examples are the following:

(27a) $U = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ (U = universe of individuals)

$\Pi = 0.8\cdot\gamma_1 + 0.6\cdot\gamma_2$

$K(\gamma_1) = 1\cdot\gamma_1 + 0.4\cdot\gamma_2$

$K(\gamma_2) = 1\cdot\gamma_2 + 0.4\cdot\gamma_1 + 0.4\cdot\gamma_3$

$$\begin{aligned}\underset{\sim}{\Pi} &= 0.8\cdot(1\cdot\gamma_1 + 0.4\cdot\gamma_2) + 0.6(1\cdot\gamma_2 + 0.4\cdot\gamma_1 + 0.4\cdot\gamma_3) \\ &= 0.8\cdot\gamma_1 + 0.32\cdot\gamma_2 + 0.6\cdot\gamma_2 + 0.24\cdot\gamma_1 + 0.24\cdot\gamma_3 \\ &= 0.8\cdot\gamma_1 + 0.6\cdot\gamma_2 + 0.24\cdot\gamma_3\end{aligned}$$

$$
\begin{aligned}
(27b)\quad U &= \gamma_1 + \gamma_2 + \gamma_3 \\
\Pi &= 0.8\cdot\gamma_1 + 0.6\cdot\gamma_2 + 0.4\cdot\gamma_3 \\
K'(\gamma_1) &= 1\cdot\gamma_2 \\
K'(\gamma_2) &= 1\cdot\gamma_3 \\
K'(\gamma_3) &= 0\cdot\gamma_1
\end{aligned}
$$

$$
\begin{aligned}
\underset{\sim}{\Pi} &= 0.8(1\cdot\gamma_2) + 0.6(1\cdot\gamma_3) + 0.4(0\cdot\gamma_1) \\
&= 0.8\cdot\gamma_2 + 0.6\cdot\gamma_3
\end{aligned}
$$

K' in the second example is a shifter. It shifts the set Π by one individual to the right. We have now an idea how s-fuzzification can be formalized, in the $\int$-representation. The reader should make sure that fuzzification in the λ-representation is much more clumsier.

Calculus, λ-calculus, and Punctuation

Our aim is, as we stated above, to provide for a solution for (1) how to shift category functions, (2) how to amalgamate category functions of words with those of punctuations, and (3) how to proceed for more complex sentence patterns. In order to do this we shall need a little more technical means, which are immediately related to the $\int$-representation of fuzzy sets. We follow hereby somewhat more the notational conventions of calculus. Instead of predicate letters we use function symbols, and we use ⌜x⌝, ⌜y⌝, ⌜z⌝ as variables.

We shall list some obvious equations:

$$
\begin{aligned}
(28):\quad 1 &= \int 1(y)dy && \text{(1-function)} \\
0 &= \int 0(y)dy && \text{(0-function)} \\
c &= \int c(y)dy && \text{(c-constant function)} \\
f &= \int f(y)dy && \text{(arbitrary function f)} \\
\underline{\text{not}}\ f &= 1-\int f(y)dy = \int (1(y)-f(y))dy && \text{(negation)} \\
f\ \underline{\text{and}}\ g &= \int \min(f(y),g(y))dy = \int (f(y)\wedge g(y))dy \equiv f\cdot g && \text{(conjunction)}
\end{aligned}
$$

$f \text{ } \underline{or} \text{ } g = \int \max(f(y),g(y))dy = \int (f(y) \vee g(y))dy \equiv f + g$ (disjunction)

$f \text{ } \underline{implies} \text{ } g = \int \max(1-f(y),g(y))dy = \int (f(y) \to g(y))dy$ (classical implication)

$f \odot g = \int f(y) \cdot g(y)dy$ (multiplication)

$f \oplus g = \int (f(y) + g(y) - f(y) \cdot g(y))dy$ (addition)

$f * g = \iint f(x)g(y)dxdy$ (exterior multiplication)

$\delta_L = \int \delta_L(y)dy \equiv dL$ (δ-function)

$\delta_L \odot g \equiv \int \delta_L(y)g(y)dy = g(L)$

$\int \delta_L(y)dK(y) = dK(L)$

$K = \int dK(y)$ (kernel of function)

$K_M^L = \int dK_M^L(y)$ (shifter)

$K_L^L = \int dK_L^L(y)$ (squeezer)

$f \bullet K = \int f(y)dK(y)$ (combination of a function and a kernel)

$(\int f(y)dy \bullet \int dK(y) = \int f(y)dK(y))$

$\delta_M = \delta_L \bullet K_M^L$ (convolution, or expectation-product)

$(\delta_M = dM = dK_M^L(L) = \int \delta_L(y)dK_M^L(y) = \delta_L \bullet K_M^L$

$(\delta_L * \delta_{L'}) \bullet K_M^{LL'} = \delta_L \bullet (\delta_{L'} \bullet K_M^{LL'})$ (iterated convolution)

$[(\delta_L * \delta_{L'}) \cdot K_M^{LL'} = (\iint \delta_L(y) \delta_{L'}(z)dydz) \bullet K_M^{LL'}$

$= \iint \delta_L(y) \cdot \delta_{L'}(z)dydz \bullet \iint dK_M^{LL'}(y,z)$

$= \iint \delta_L(y) \cdot \delta_{L'}(z) \cdot dK_M^{LL'}(y,z)$

$= \int \delta_L(y) \cdot \int \delta_{L'}(z)dK_M^{LL'}(y,z)$

$= \int \delta_L(y)dK_M^{LL'}(y,L') = \int \delta_L(y)dK_M^L(y) = \delta_L \bullet K_M^L$

$K_M^L = \int dK_M^L(y) = \iint \delta_{L'}(z)dK_M^{LL'}(y,z) = \delta_{L'} \bullet K_M^{LL'}]$

hence the result; moreover the following is true:

$(\delta_L * \delta_{L'}) \bullet K_M^{LL'} = \delta_L \bullet K_M^L = K_M$

$d\int = 1$ (d-$\int$, inverses of each other) : $\frac{d}{dx}\int f(y)\,dy = f(x)$

$[\int f(y)dy](z) = f(z)$ (conversion/extraversion)

$f(x) = f(y) \frac{dy}{dx}$ (substitution)

$\frac{d}{dx} f = f(x)$ (differentiation = application)

$(\frac{d}{dx}f = \frac{d}{dx}\int f(y)dy = d\int f(y) \frac{dy}{dx} = f(x))$

Convolution, generalized to arbitrary functions, allows us to formulate functions which carry a fuzzy argument to a fuzzy value. Kernels represent fuzzy functions of this most general sort. We have thus arrived at a reconstruction of a fully fuzzified λ-calculus.

It is now clear how the fuzzy categories of a categorial (or phrase structure) grammar with punctuation signs are/going to be fuzzified; instead of the general case let us give an example with two words (morphemes) and one punctuation sign. Let the corresponding categories be C_1, C_2, and F (F can be taken to be $C_1\backslash C_2\backslash S$):

$$(29)\quad C_1 + C_2 + F \longrightarrow F\ (C_1, C_2)$$

If these categories would be classical categories, i.e. non-fuzzy categories, then the corresponding category functions would be simply δ-functions, and the kernel corresponding to F would be a δ-kernel:

$$(30)\quad (\delta_{C_1}, \delta_{C_2}, I_S^{C_1C_2}) \longrightarrow (\delta_{C_1} * \delta_{C_2}) \circ I_S^{C_1C_2}$$

There are equivalent formulations because of:

$$(31)\quad (\delta_{C_1} * \delta_{C_2}) \circ I_S^{C_1C_2} = \delta_{C_1} \circ (\delta_{C_2} \circ I_S^{C_1C_2}) = \iint \delta_{C_1}(y)\delta_{C_2}(z)dI_S^{C_1C_2}(y,z)$$

This is as it should be. But (30) can be easily generalized. Instead of δ-functions and δ-kernels simply take general functions and shifters. Hence for fuzzy categories C_1, C_2, and F (30) reads

$$(32)\quad (f, g, K_S^{C_1C_2}) \longrightarrow (f * g) \circ K_S^{C_1C_2}$$

We assigned hereby the words with alleged categories C_1, C_2 (their) category functions f, g. According to

$$(33)\quad (f * g) \circ K_S^{C_1C_2} = f \bullet (g \bullet K_S^{C_1C_2}) = \iint f(y)g(z)dK_S^{C_1C_2}(y,z)$$

we are able to calculate the outcoming category, if the words and the punctuation sign are concatenated.

The general case is easily given by expressions such as

$$(34)\quad (f_1*f_2*f_3*((f_4*f_5) \bullet K_S^{C_1C_2}) * f_6) \bullet K_S^{C_1C_2C_3SC_6}$$

As is seen from (34), punctuation signs can occur at different places of a sentence. They can be embedded. Other punctuations than such with a target category S are conceivable, of course.

Remarks:

1. For a λ-ϰ calculus right and left integrals have to be distinguished: $\int$, $\rfloor$. Hence the following equations have to be distinguished:

$$[\int f(y)dy](z) = f(z)$$
$$(z)[\rfloor f(y)dy] = f(z)$$

Correspondingly right and left differentiation has to be distinguished.

2. The space of categories should be topologically structured or even more it should be locally parametrized. Otherwise the generalization from the classical δ_L case to the fuzzy f_L case is not valid.

3. The treatment presented here allows to account in a straightforward manner for sentences such as:

(35) "IT THE AS" IS NOT GRAMMATICAL.

The reason is simply that the negation of grammaticality is easily statable. The combination of an ungrammatical sentence with ...IS NOT GRAMMATICAL. is grammatical. Thus the kernel corresponding to ...IS NOT GRAMMATICAL is $(1-K_S^{\cdots})$. A more adequate treatment should combine grammatical and semantical statements, however, because a grammatical sentence filled in for ... does not render the whole sentence ungrammatical but simply false!

4. This stage of a fuzzification of the λ-calculus is not the ultimate one. A totally satisfying stage is only reached, when a preferably complete axiomatization relative to appropriate rules of interference is given. Moreover, it should be aimed at providing for a metamathematical treatment which is of the same kind as the system studied, i.e. there should be provided for a fully fuzzified metamathematical foundation of fuzzy logic and fuzzy λ-calculus.

References:

Th. T Ballmer (1974) "Sprachrekonstruktionssysteme" (Language Reconstruction Systems). Scriptor, Kronberg/Ts. (1975)

Th. T Ballmer (1975) "Logik und Wahrscheinlichkeitstheorie" (Logic and Theory of Probability). Mimeo.

J. R. Ross (1972) "Endstation Hauptwort: The category squish" in: P. M. Peranteau et al. CLS 8, pp. 316-328.

J. R. Ross (1973) "Nouniness" mimeo. MIT.

L. A. Zadeh (1972) "A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges." Memorandum No. ERL-235. Electronics Research Laboratory, U.C. Berkeley, and Journ. of Cybernetics, 2, 3, pp. 4-34.

G. Lakoff (1975) "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts," in: Hockney, et al. (eds.)"Contemporary Research in Philosophical Logic and Linguistic Semantics." Reidel, Dordrecht.