Development of a method to evaluate odour quality based on non-expert analysis

Carıou S.[1]*, Chaıgnaud M.[3], Medjkoune M.[1, 2], Harıspe S.[2], Montmaın J.[2] And Fanlo J.L[1].

[1] Ecole des Mines d'Alès, Laboratoire Génie de l'Environnement industriel, 6 av. de Clavières, F-30319 Ales Cedex, France

[2] Ecole des Mines d'Alès, Laboratoire Génie Informatique et Ingénierie de Production, 6 av. de Clavières, F-30319 Ales Cedex, France

[3] Olentica sas, 14 bd Charles Péguy, F-30100 Ales, France

*corresponding author:

e-mail: stephane.cariou@mines-ales.fr

Abstract. Characterizing odour quality is a complex process that consists in identifying a set of descriptors that best synthesizes the olfactory perception. Generally, this characterization results in a limited set of descriptors provided by professionals in sensorial analysis. These experts previously learnt a common language to describe characteristic odour (Odour wheel or Champ des odeurs©). These sensorial analysis sessions cost industrial manufacturers large sums every year. If this characterization is entrusted to neophytes, the number of participants of a sensorial analysis session can be significantly enlarged while reducing costs. However, each individual description is no more related to a set of non-ambiguous descriptors but to a bag of terms in natural language. Two issues are then related to odour characterization. The first one is how translating free natural language descriptions into structured descriptors; the second one is how summarizing a set of individual characterizations into a consistent and synthetic unique characterization for professional purposes. This paper will propose an approach based on natural language Processing and Knowledge Representation based techniques to formalize and automatize both translation of bags of terms into sets of descriptors and summarization of sets of structured descriptors.

Keywords: Odour Quality, Natural Language, Information Fusion, Taxonomy, Semantic Proximity.

1. Introduction and problem

Odours represent a very important issue for societal and industrial perspective and activities due to the intrinsic character of the odour, or to the frequency of the perception (Gostelow & Parsons 2000). In general, the industries' locations or the sale of materials for building and furnishing, household care products depend of their odour acceptability in the neighbourhood or by potential buyers. Because of around 20% of the European population are annoyed by environmental odours, the rules and the regulations have been enhanced in the odour monitoring's field (JORF n°89 2003), (Belgiorno *et al.* 2012).

The characterization of the odour quality consists in the verbalization of the perception that a person makes of his feeling based on his experience and his knowledge. Currently, the evaluation of odours quality is commonly made through controlled linguistic descriptors provided by trained experts. Industrials have developed specific classifications to calibrate the quality of odours. We can mention the Champ des odeurs© for perfumers (Jaubert *et al.* 1995) and the wheel of odours for example for œnology (Noble *et al.* 1984) or drinkwater (Suffet & Rosenfeld 2007). By using a common referential to qualify odours, these methods facilitate understanding, interpreting and results processing. Nevertheless, a learning phase in which valid descriptors have to be learned is required to use such methods. This (i) prevents their use by non-experts, (ii) implies additional training costs, and (iii) limits the number of evaluators and experiments that can be used to evaluate odours.

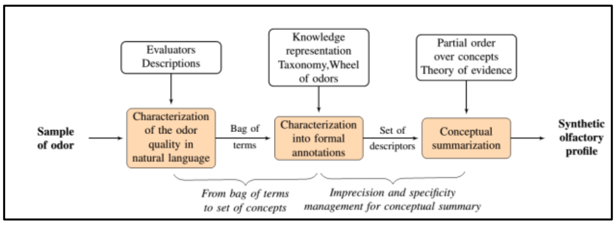

Figure 1: Process of the odour quality evaluation

A natural language (NL) is a language spoken by individuals without prior training. It is a vocabulary that is opposed to the language of experts. The characterization of odour quality by non-expert evaluators is far from a so-called objective description because they used terms referring to sources or pleasantness. On the contrary, experts seem to forget the hedonic tone of the odours when they characterize quality (Sezille *et al.* 2014).

The aim of our project is to propose a method that enables the transcription of free descriptions given in natural language into structured descriptors.

1.1 Introductory example:

Let's consider that we have two non-expert evaluators to describe the odour of a given tested wine, each of their free descriptions is related to a bag of terms in natural language. Defining a method to merge the information expressed by the free descriptions that is useful for professionals necessitates a good deal of knowledge of the domain-specific descriptors and their taxonomical organization. If for example, we consider that we get the following conceptual descriptions d_1 and d_2 after translating the free NL descriptions into structured descriptors:

$$d_1 = \{Asparagus;\ Strawberry;\ Vanilla\}$$
$$d_2 = \{Curry;\ Spinach;\ Raspberry\}$$

The idea is to merge the information expressed by these two conceptual annotations d_1 and d_2 to formally characterize the odour of this wine by a unique synthetic of concepts. Intuitively, abstracting these two descriptions by summarizing the information provided by the descriptors relies on our knowledge of the organization of the descriptors. Because there is no formal definition to what a relevant summary is, several summaries can be proposed, e.g.:

$$Y = \{Red\ fruit;\ Spicy;\ Vegetable\}$$

Indeed, the idea beyond this choice is that $Asparagus$ and $Spinach$ are sorts of $Vegetable$(i.e.$Asparagus \prec Vegetable, Spinach \prec Vegetable$), $Vanilla$ is a sort of $Spice$ and $Red\ fruit$ is a specification of $Fruits$. This reasoning is based on taxonomic knowledge partially ordering concepts (descriptors).

The next section presents the proposed method, we compute the individual conceptual annotation and then a formal method to produce the conceptual summary that best approximates all the annotations.

2. Method:

In this article, we consider that the terms used by non-specialists to describe the quality of an odour are free: it is assumed that the evaluators use their own terms to describe the odour. Therefore, the terms we have to deal with are not part of a controlled vocabulary defining the descriptors commonly used for characterizing odours. The purpose of our work is therefore to identify the descriptors that are the most likely evoked by the natural language descriptions provided by non-experts. The notion of descriptor here refers to the concept used in Knowledge Representation (taxonomy). These concepts are assumed to be partially ordered into taxonomy $\mathcal{O} = (C, \preccurlyeq)$, with C the set of concepts (Harispe, Ranwez, *et al*. 2015). A concept c is subsumed by a concept c', i.e., $c \prec c'$ if c is more specific than c' and inversely c' is more imprecise than c, e.g. *Strawberry* is a sort of *Red fruit*.

The first step of our method is to define a correspondence between terms and concepts of the related taxonomy. This mapping is based on a vectorial word-based representation that allows comparing, then associating terms and concepts. In this way, we compute the degree to which a term evokes a concept of the taxonomy. Then, each term of the NL description is associated the concept the term evokes the most. Finally the free description is transformed into a conceptual annotation. The second step is to synthesize conceptual annotations into a limited set of concepts. This step is all the more necessary when several evaluators provide their annotations to make the collective merged evaluation more synthetic and tractable.

In order to ease the readability of this section, the various definitions which will be used are listed below:

Notations:

- The notion of Information Content (IC) refers to the degree of specificity of concepts. (Seco *et al*. 2004). For any IC function must monotonically decrease from the leaves to the root of the taxonomy such as $\forall (x, y) \in C^2, x \preccurlyeq y \Rightarrow IC(x) \geq IC(y)$.
- We denote by $\mathcal{A}(C')$ and $\mathcal{D}(C')$ respectively the inclusive ancestors and inclusive descendants of the set of concepts $C' \subseteq C$.

$$\mathcal{A}(C') = \cup_{c\in C'} \cup_{c\preccurlyeq c'} \{c'\}, \mathcal{D}(C') = \cup_{c\in C'} \cup_{c'\preccurlyeq c} \{c'\}$$

- We denote by $m: C \rightarrow [0,1]$ the mass function, $m(c)$, $c \in C$ that corresponds to the number of observations of concept c and $\sum_{c\in C} m(c) = 1$. In our application, $m(c)$ represents the number of times concept c has been proposed by evaluators.
- The belief and plausibility functions $bel: C \rightarrow [0,1]$ and $pl: C \rightarrow [0,1]$ proposed in the Dempster–Shafer theory are next defined such as (Harispe, Imoussaten, *et al*. 2015):

$$bel(c) = \sum_{x \preccurlyeq c} m(x)$$
$$pl(c) = \sum_{x\in C, \mathcal{D}(x)\cap\mathcal{D}(c)\neq\emptyset} m(x)$$

Step 1: Individual conceptual annotation

In this section, we try to associate to each term of a NL description, the concept that is the most likely evoked by the term. We intuitively propose to consider that a term evokes concepts (*e.g*. the term strawberry evokes with some degrees the concepts *Red fruits* and *Fruits*). The intensity of evocation a term has with a concept can be expressed as the semantic proximity between the term and the concept denoted by $\sigma_{TC}: T \times C \rightarrow [0.1]$ with T the set of terms that constitute the vocabulary that has been previously established for the vectorial representation of words. We propose to use the distributional semantics models for evaluating term proximity with regard to their meaning (*i.e*. concepts). The distributional models are generally defined to capture the meaning of terms. These models are based on the distributional hypothesis which, in linguistics, states that words that are used and occur in the same contexts tend to purport similar meanings (Harris 1981). The distributional semantic model is built from the distributional analysis of corpora. The first step to obtain such a model is to analyze terms co-occurrences in a large corpus (e.g. Wikipedia). These co-occurrences are then used to derive models. These models can be of various forms, e.g. word-word matrix (Harispe, Ranwez, *et al.* 2015). We next consider that measuring the similarity between a term t and a concept c returns to calculate the similarity between the term and the labels (terms) associated to the concept c. The labels are provided in the input taxonomy. This semantic proximity between t and c can be estimated, for example, as the maximal of similarity between the term t and the labels associated to the concept c: $\sigma_{TC}(t,c) = \max_{l \in L_c} \sigma_{TT}(t,l)$ with L_c the set of terms associated to the concept c and $\sigma_{TT}: T \times T \rightarrow [0.1]$. Numerous measures σ_{TT} for comparing terms have already been proposed in the literature (Harispe, Ranwez, *et al.* 2015) , e.g. cosine measure. The objective of this step is to synthesize the information expressed by the terms of a description T_e provided by evaluator e to characterize the odour through a conceptual annotation. This conceptual annotation X_e can be computed as follows:

$$X_e = \bigcup_{t \in T_e} \{c \in C \mid \max_{c \in C} \sigma_{TC}(t,c)\} \qquad (1)$$

This model considers a simple one-to-one correspondence between terms and concepts.

The next section focuses on the issue of summarizing the individual conceptual annotations in the case where we have many evaluators who have to characterize the same odour.

Step 2: Collective conceptual annotation (summary)

The aim of the study is to summarize the information given by a set of evaluators (E). We consider that each evaluator e provides a set of terms T_e to describe the odour. Using the model proposed in the previous section (Eq. (1)), a conceptual annotation X_e can be associated to each individual bag of terms T_e. The purpose of this section is to provide a synthesis (summary) that summarizes the information provided by evaluators. Indeed to each evaluator $e_n \in E$ is now associated a set of concepts (individual conceptual annotation) denoted $X_n \in \mathcal{P}(C)$. In the following, we propose a method to summarize a set of individual conceptual annotations.

We denote by $\hat{X}=(X_1, X_2, \dots, X_n)$ the sequence of annotations to be summarized and X the set of concepts issued from individual annotations such that: $X = \bigcup_{i=1}^{n} X_i$. We suppose that each synthesis $Y \in \mathcal{P}(C)$ respects the following properties:

1) Summarizing: $|Y| \leq |X|$
2) Fidelity: $\forall y \in Y, \exists x \in X$ such as $x \preccurlyeq y$
3) Non-total-redundancy: $\forall (x,y) \in Y^2, x \nprec y \wedge y \nprec x$,

We denote by $S^X \subseteq \mathcal{P}(C)$ the set of summaries of a sequence of annotations $\hat{X}$, each summary $Y \in S^X$ respects the three properties defined above. Based on these definitions, we formally define by S the function summarizing a sequence of n individual conceptual annotations by a single summary from S^X:

$S: \mathcal{P}(C)^n \rightarrow S^X$, with $S(\hat{X}) \in S^X$.

We define the problem of summarizing a sequence of annotations $\hat{X}$ by finding $Y \in S^X$, the best summary for $\hat{X}$. The following section introduces the model that we propose for defining a function S.

In order to define the best summary, we propose to define some notions that could be used to evaluate the relevance of a summary. The search of best summary $Y \in S^X$ for any $\hat{X} \in \mathcal{P}(C)^n$ can be expressed as an optimization problem. The objective function is defined as:

$$S(\hat{X}) = arg\max_{Y \in S^X}(\psi(Y,\hat{X}) - \mathcal{L}(Y,\hat{X}))$$
$$\mathcal{L}(Y,\hat{X}) = \Delta(Y,\hat{X}) + \lambda(Y) + \gamma(Y,\hat{X})$$

The function $\psi(Y,\hat{X})$ models the amount of information from $\hat{X}$ covered by Y and $\mathcal{L}(Y,\hat{X})$ is the penalty associated to the abstraction of $\hat{X}$ by Y. The function $\psi(Y,\hat{X})$ is used to estimate the amount of conceptual information conveyed by $\hat{X}$ which is summarized by Y. In this context, we are interested to measure the common abstract notions which are mentioned by $\hat{X}$ and Y. That is to say, we search to evaluate the common information provided by the ancestors of $\hat{X}$ and Y. Intuitively this quantity could be defined as follows:

$$\psi(Y,\hat{X}) = \sum_{c \in \mathcal{A}(X) \cap \mathcal{A}(Y)} w(c) \times IC(c)$$

The function $w: C \rightarrow [0,1]$ is introduced to weigh the importance of concepts. The function $\psi(Y,\hat{X})$ is illustrated in figure 2.

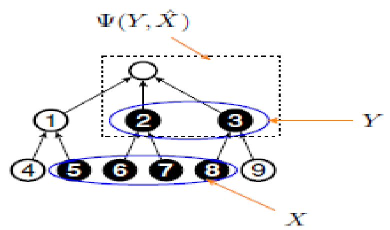

Figure 2: The dashed frame represents the set of concepts that directly contribute to compute $\psi(Y,\hat{X})$.

The function $\mathcal{L}(Y,\hat{X})$ represents the penalty associated to the abstraction of $\hat{X}$ with:

- $\Delta(Y,\hat{X})$: function of penalties regarding loss, addition and distortion of information.
- $\lambda(Y)$: function evaluating the conciseness of the summary.
- $\gamma(Y,\hat{X})$: function that can be used to express additional constraints over Y (e.g. retaining only the covering summaries).

In the following, we will detail the different factors of penalty. We define the penalty of abstraction $\Delta(Y, \hat{X})$ by:

$$\Delta(Y, \hat{X}) = f(\Delta^{E-}, \Delta^{P+}, \Delta^{P-}, \Delta^{D})$$

Δ^{E-} is the penalty associated to the deletion of the exact information. It models the amount of exact information conveyed by $\hat{X}$ which is not conveyed by Y (e.g. concepts 1, 5, 6, 7, 8 in figure 3).

$$\Delta^{E-}(Y, \hat{X}) = f(\mathcal{A}(X) \backslash \mathcal{A}(Y)) = \sum_{c \in \mathcal{A}(X) \backslash \mathcal{A}(Y)} bel(c) \times IC(c)$$

Δ^{P+} models the amount of plausible information conveyed by Y which is not conveyed by $\hat{X}$ (penalty regarding addition of plausible information) (e.g. concept 9 in figure 3).

$$\Delta^{P+}(Y, \hat{X}) = f(\mathcal{D}(Y) \backslash \{\mathcal{D}(X) \cup \mathcal{A}(X)\})$$

$$\Delta^{P+}(Y, \hat{X}) = \sum_{y \in \mathcal{D}(Y) \backslash \{\mathcal{D}(X) \cup \mathcal{A}(X)\}} pl(y) \times IC(y)$$

Δ^{P-} models the amount of plausible information conveyed by $\hat{X}$ which is not conveyed by Y (penalty regarding loss of plausible information) (e.g. descendants of concept 5 in figure 3).

$$\Delta^{P-}(Y, \hat{X}) = f(\mathcal{D}(X) \backslash \mathcal{D}(Y))$$

$$\Delta^{P-}(Y, \hat{X}) = \sum_{y \in \mathcal{D}(X) \backslash \mathcal{D}(Y)} pl(y) \times IC(y)$$

Finally, the aim of the penalty Δ^{D} is to penalize the distortion which is made considering a specific choice among partially covering summaries. This penalty should be a function of $X \backslash \mathcal{D}(Y)$ i.e., all the elements of X that have not been summarized (e.g. concept 5 in figure 3). We propose the following model to estimate the distortion:

$$\Delta^{D}(Y, \hat{X}) = \tau \sum_{x \in X \backslash \mathcal{D}(Y)} \sum_{x' \in \mathcal{A}(\{x\}) \backslash \mathcal{A}(Y)} (bel(x') . IC(x'))$$

The parameter τ is used to weigh the importance of a specific uncovering (i.e. the concepts of X that are not covered by Y). Finally, the penalty of abstraction $\Delta(Y, \hat{X})$ is defined as follow:

$$\Delta(Y, \hat{X}) = \delta_{E-} \Delta^{E-} + \delta_{P+} \Delta^{P+} + \delta_{P-} \Delta^{P-} + \delta_{D} \Delta^{D}$$

With δ_{E-}, δ_{P+}, δ_{P-}, δ_{D} input parameters used to set the importance of each abstraction penalty factor. These penalties are illustrated in figure 3.

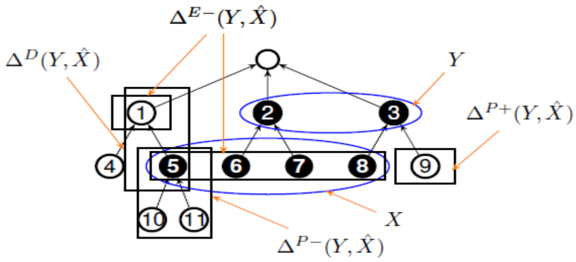

Figure 3: Illustration of the different factors of the penalty. Each frame associated to a penalty represents the set of notions or concepts involved in the computation of the penalty.

$\lambda(Y)$ is a penalty used to evaluate the conciseness of the summary Y by penalizing redundant information implicitly conveyed by a summary.

$$\lambda(Y) = \epsilon \sum_{y' \in \mathcal{A}(Y)} ((|\{y \in Y | y' \in \mathcal{A}(y)\}| - 1) \times IC(y'))$$

The penalization is designed such as each abstracted notion that is repeated more than once will be penalized the number of times the redundant information appears. The parameter ϵ can therefore be used to control the number of descriptors composing a summary. The more important ϵ, the more abstracted the summary.

Finally, the function $\gamma(Y, \hat{X})$ can be used to express additional constraints over Y. This constraint can be used to apply specific restrictions on the type of solutions we are interested in. As an example, the requirement may be to keep only the covering summaries Y ($\forall x \in X, \exists y \in Y$ such as $x \preccurlyeq y$).

Harispe *et al*. have proposed algorithms enabling to use the model for searching for relevant summaries and discuss interesting properties of the search space S^X- details on the performances of these algorithms are provided in (Harispe *et al*. 2017).

3. Experimental framework

In this example, fifteen members of a sensorial jury evaluated odour quality of wine. Each member described his perception with his own words.

For example, a member described the tested wine by the following words: black fruit, prune, strawberry, compote, cinnamon, peach, exotic fruit.

With the description all of the members, a bag of fifty seven different terms is obtained. When combining with our method, a summary of these terms in five concepts is represented in figure 4. The weight of each concept is figured on the axis. It reveals the importance of the concept in the bag of terms and, in fact, in the primary description of the perception.

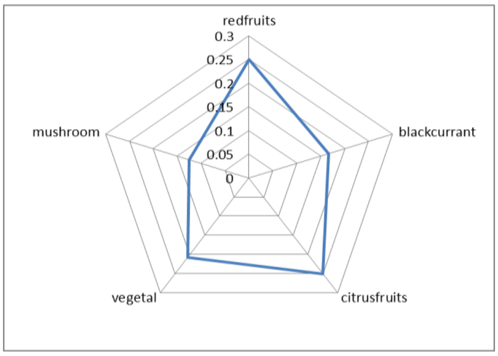

Figure 4: Schematic representation of concepts from odour quality evaluation

The five concepts chosen to characterize odour quality are the best summary to explain the perception of all the panellists without forgetting crucial information.

These representations are very useful to qualify and compare olfactory perception for several products of a same production.

4. Conclusion

This new method is able to combine evaluation of odour quality from non-experts and to sort out a summary of concept. Works will be done on the validation of this method on a large scale of products and different perception. The evaluation of a taxonomy based on a multi-dimensional sensorial space might be useful to be as close as possible of human evaluation of odour quality.

References

Belgiorno, V., Naddeo, V. & Zarra, T., 2012. *Odour Impact Assessment Handbook* , V. Belgiorno, V. Naddeo, & T. Zarra, eds., John Wiley and Sons, Ltd.

Gostelow, P. & Parsons, S.A., 2000. Sewage treatment works odour measurement. *Water science and technology*, 41(6), pp.33–40.

Harispe, S., Imoussaten, A., *et al.*, 2015. On the consideration of a bring-to-mind model for computing the Information Content of concepts defined into ontologies. *IEEE International Conference on Fuzzy Systems (FUZZ IEEE), Istanbul, Turkey*.

Harispe, S., Ranwez, S., *et al.*, 2015. Semantic Similarity from Natural Language and Ontology Analysis. *Synthesis Lectures on Human Language Technologies*, 8(1), pp.1–254.

Harispe, S., Montmain, J. & Medjkoune, M., 2017. Summarizing Conceptual Descriptions using Knowledge Representations. *IEEE Symposium Series on Computational Intelligence,*.

Harris, Z.S., 1981. Distributional structure. *In Papers on syntax. Springer Netherlands*, pp.3–22.

Jaubert, J.-N., Tapiero, C. & Dore, J.-C., 1995. The field of odors: toward a universal language for odor relationships. *Perfumer & flavorist*, 20(3), pp.1–16.

JORF n°89, 2003. Arrêté du 12 / 02 / 03 relatif aux prescriptions applicables aux installations classées soumises à autorisation sous la rubrique 2730 (traitement des cadavres , des déchets ou des sous-produits d' origine animale à l'exclusion des activités visées par. *Journal Officiel de la Republique française*, 89.

Noble, C. *et al.*, 1984. Progress Towards a Standardized System of Wine Aroma Terminology. *American Journal of Enology and Viticulture*, 35(2), pp.107–109.

Seco, N., Veale, T. & Hayes, J., 2004. An intrinsic information content metric for semantic similarity in wordnet. *in ECAI*, 16.

Sezille, C. *et al.*, 2014. Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists, and perfumers. *Frontiers in Psychology*, 5, p.12.

Suffet, I.H. & Rosenfeld, P., 2007. The anatomy of odour wheels for odours of drinking water , wastewater , compost and the urban environment. *Water Science & Technology*, 55(5), pp.335–344.