Pooling Commentaries on Legal Texts from Lawyers and Experts in a Collaborative Work

El-Hassan Bezzazi
University of Lille 2
59000 Lille, France
bezzazi@univ-lille2.fr

ABSTRACT
This short paper presents a methodological framework for commentaries from various sources on the same texts allowing them to be assembeled in an intelligent and coherent way.

Categories and Subject Descriptors
I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems; H.4 [**Information Systems Applications**]: Miscellaneous

General Terms
Documentation, Human Factors

Keywords
collaborative authoring, document commenting

1. INTRODUCTION
It often happens that legal texts tackle very technical subjects which require technical knowledge that the lawyers do not necessarily apprehend with all their subtleties. This implies then, for the clearness as well as for the coherence of the texts considered, the implication of qualified people in the technical field of interest. These people in their turn do not necessarily perceive the legal meaning of some terms and expressions of these texts. The object of this article is to investigate mathematical properties for a methodological framework that accommodates and guides these activities on the texts being evaluated.

2. DEFINITIONS AND FACTS
Let us consider the case of a given team E studying a given document in order to model the work of a team. A document is a text T which we will identify to the set of natural numbers $[1, N]$ where N is the length of T i.e. the number of lexical units composing it. A segment of T is a closed interval $[x, y]$ on T and represents a passage of T to be commented on. Let $\mathcal{S}(T)$ be the set of all segments on T. A support S on T is some subset of $\mathcal{S}(T)$ such that for any s, s' in S, $s \cap s' = \emptyset$. The work of a team is defined by a couple (S, κ) where S is the support of the work and κ, the commenting function, is a map from S to the set of all commentaries C.

We define two binary relations $\prec$ on S and $\ll$ on Σ as follows:

definition 1. Let $s = [x, y]$ and $s' = [x', y']$ be two segments in $\mathcal{S}(T)$. $s \prec s'$ if and only if $y < y'$ or ($y = y'$ and $x < x'$). It is the lexical ordering from right to left. Let $S \in \Sigma$ and $S' \in \Sigma$ be two supports on T. $S \ll S'$ if and only if $\forall s \in S. \exists s' \in S'. s \subseteq s'$.

It is easy to check that the relations $\prec$ and $\ll$ define respectively a total ordering on $\mathcal{S}(T)$ and a partial order on Σ. On the other hand, note that the relation $\equiv$ defined on Σ by $S \equiv S'$ if and only if $\cup S = \cup S'$ is an equivalence relation. The aquivalence class of a given support S w.r.t $\equiv$ possesses a minimum we shall denote $\hat{S}$. The proof of the existence and the construction of this minimum which is merely technical will not be given here. Roughly speaking, it consists in ordering the segments of all the supports in the equivalence class of S using the ordering $\prec$ and merging the segments that form a chain of intersections into one segment. The support $\hat{S}$ will be called the normal form for the supports of its equivalence class.

Let S be a support on T, we define the set $sup(S)$ to be the set of the maximums of the segments of S: $sup(S) = \{y \in N : [x, y] \in S\}$

PROPOSITION 1. *Let S and S' be two supports on T. $S = S'$ if and only if $\cup S = \cup S'$ and $sup(S) = sup(S')$.*

definition 2. Let $\hat{\Sigma}$ be the set of normal supports in Σ. We define a binary operation $+$ on $\hat{\Sigma}$ as: $\hat{S} + \hat{S}'$ is the normal support on T such that $\cup(S+S') = (\cup S)\cup(\cup S')$ and $\forall s \in S \cup S'.\exists t \in S+S'.s \subseteq t$.

PROPOSITION 2. *The support $S+S'$ exists and is unique and the operation $+$ is commutative and associative,* i.e. $S+S' = S'+S$ and $(S+S')+S'' = S+(S'+S'')$.

3. ALTERNATIVE OPERATIONS

The synthesis of two works using the operation $+$ does not take into account the fact that some passages might have more importance than others because of the fact that they drew the attention of several commentators. Therefore, we will define a second operation on the supports noted $S \cdot S'$ which will be used for the synthesis of two works. This operation will select the segments of S and S' which intersect each other and will retain their unions as being the only segments of the new support $S \cdot S'$. In order to define this operation, we need the following proposition which extends the definition of normal supports to the class $\mathcal{S}(T)$.

PROPOSITION 3. *Let U be some subset of $\mathcal{S}(T)$. There exists a unique normal support $\hat{U}$ in Σ s.t. $\cup U = \cup \hat{U}$.*

definition 3. Let S and S' be two supports on T. The consensus synthesis is defined by:
$S \cdot S' = \{s \cup s' : s \in \widehat{S, s' \in} S', s \cap s' \neq \emptyset\}$.

The fact that we retain in this definition the union of the segments which are intersected is justified by the construction of the commentary to be associated to it. This commentary will be the union of the commentaries. This is easily defined whereas the intersection is not when the textual contents of the comments are freely expressed as this is in our case. Unfortunately this operation is commutative but not associative. To see this, consider the following supports: $S = \{s\}, S' = \{s'\}, S'' = \{s''\}$ such that $s \cap s' = \emptyset$, $s \cap s'' \neq \emptyset$ and $s' \cap s" \neq \emptyset$. We have: $S \cdot (S' \cdot S") = S \cdot \{s' \cup s"\} = \{s \cup s' \cup s"\}$ and $(S \cdot S') \cdot S" = \emptyset \cdot \{s"\} = \emptyset$.

The fact that this operation is not associative leads us to redefine it according to the set of the commenting teams. If the number of teams is m: $E_1 ..., E_m$, one defines the n-ary operation on their respective supports:

$\sigma(S_1, ..., S_m) = \{\cup_{i\in[1,m]}\{s_i\} : \widehat{\forall i \in [1,m]} s_i \in S_i \textit{ and } \cap_{i\in[1,m]}\{s_i\} \neq \emptyset\}$.

Note that this defintion is too strong as it retains only the segments having a common part for all the teams. It can be made less rigid if we parameterize it with an integer p:

$\sigma_p(S_1, ..., S_m) = \{\cup_{i\in I}\{s_i\} : \widehat{\forall i \in I}.s_i \in S_i \textit{ and } \cap_{i\in I}\{s_i\} \neq \emptyset \textit{ for some } I \subseteq [1,m] \textit{ s.t. } card(I) = p\}$.

It is easy to check the following properties for the class $\{\sigma_p\}_{p\in[1,m]}$:
(i) $\forall i, j \in [1,m].i \leq j \Rightarrow \sigma_j(S_1, ..., S_m) \subseteq \sigma_i(S_1, ..., S_m)$, i.e. $\{\sigma_p\}_{p\in[1,m]}$ forms a decreasing chain.
(ii) $\sigma_1(S_1, ..., S_m) = S_1 + ... + S_m$ and $\sigma_m(S_1, ..., S_m) = \sigma(S_1, ..., S_m)$.

4. COLOURED SUPPORTS

Consider the set of colours $\mathcal{C} = \{blue, yellow, green\}$ used for example to mark the pieces of text being the subject of one or more comments: blue for a lawyer commentary, yellow for an expert commentary and green for a commentary from both a lawyer and an expert.

Consider now the commutative binary operation $+$ defined on $\mathcal{C}$ by:

+	blue	yellow	green
blue	blue	green	green
yellow	green	yellow	green
green	green	green	green

This operation is associative. A coloured support is a couple (S, c) where $S \in \Sigma$ and $c : S \longrightarrow \mathcal{C}$ Let (S, c) and (S', c') be two coloured supports, we define $(S, c) + (S', c')$ as being the coloured support $(S+S', c+c')$ where $c + c' : S + S' \longrightarrow C$ is defined by the following :
$c + c'(s) = c(s)$ if $s \in S$ and $s \cap S' = \emptyset$,
$c + c'(s) = c'(s)$ if $s \in S'$ and $s \cup S = \emptyset$,
$c + c'(s) = c(s) + c'(s)$ if $s \cap S \neq \emptyset$ and $s \cap S' \neq \emptyset$.
This operation is commutative and associative. This is simply the consequence of the same properties for addition over S and addition over C.

5. CONCLUSION

Different strategies can be adopted to select and synthesise the relevant commentaries depending on the degree of consensus we wish to have in the choice of relevant passages. In [1], information merging is investigated with respect to multiple documents. Techniques from text summarization [2, 3] may be investigated to be applied to synthesized commentaries to make them more structured and more intelligible.

6. REFERENCES

[1] R. Barzilay, K. R. McKeown, and M. Elhadad. Information fusion in the context of multi-document summarization. In *Proceedings of the 37th Annual Meeting of the ACL*, pages 550–557. Association for Computational Linguistics, 1999.

[2] M. A. Hearst. Texttiling: A quantitative approach to discourse segmentation. *Computational Linguistics*, 23(1):33–64, Mar. 1997.

[3] A. Hoffmann and S. B. Pham. Towards topic-based summarization for interactive document viewing. In *Proceedings of K-CAP'03*, pages 28–35.