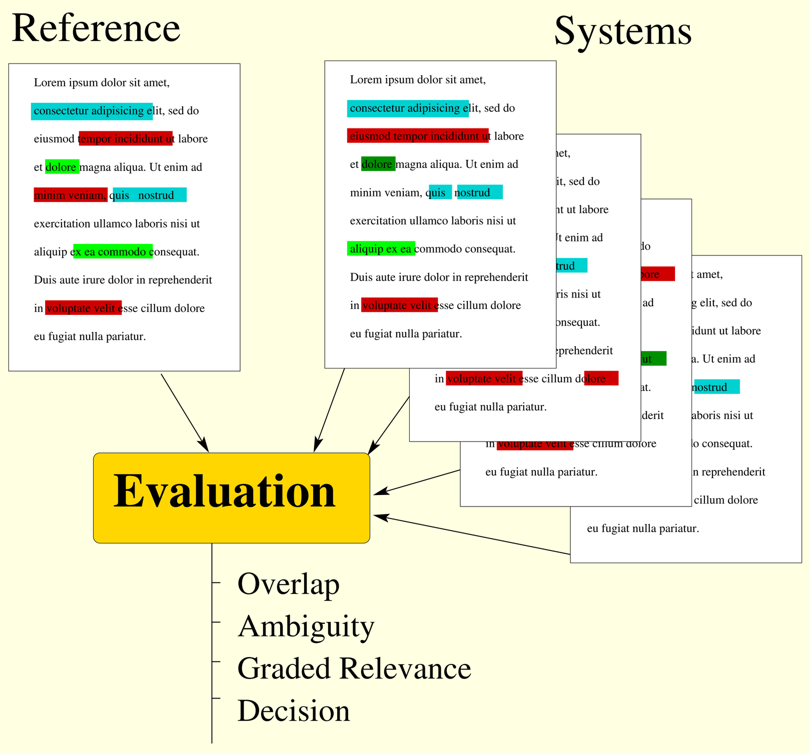
A Rough Set Formalization of Quantitative Evaluation with Ambiguity

Patrick Paroubek and Xavier Tannier, LIMSI-CNRS, {pap,xtannier}@limsi.fr

Context



Evaluation :

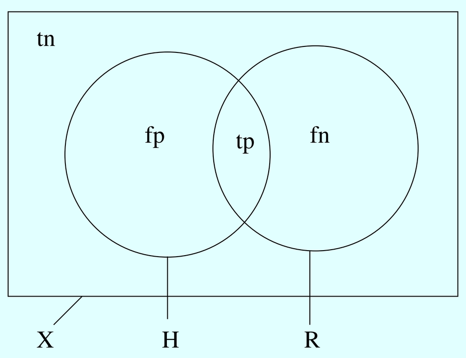
of Technology,
Objective,
Quantitative,
Black-Box.

Objectives

– No formal framework exists for studying the **evaluation paradigm** ; we propose to lay the foundation for such model based on the mathematical notion of "**rough sets**".

– We propose to consider the notion of **potential performance space**, for describing the performance variations corresponding to the **ambiguity** present in the hypothesis data.

Evaluation



$$acc = \frac{|H \cap R| \cup (X \setminus |H \cup R|)}{|X|}$$

$$err = \frac{|(H \cup R) \setminus (H \cap R)|}{|X|}$$

$$j = \frac{|H \cap R}{|H \cup R|}$$

$$f = \frac{1}{\frac{\alpha}{p} + \frac{(1-\alpha)}{r}}, 0 < \alpha < 1, \mathbf{p} = \frac{|\mathbf{H} \cap \mathbf{R}|}{|\mathbf{H}|}, r = \frac{|H \cap R|}{|R|}$$

$$\forall\, protocols, performance = \phi(|TP|, |FP|, |FN|, |TN|)$$

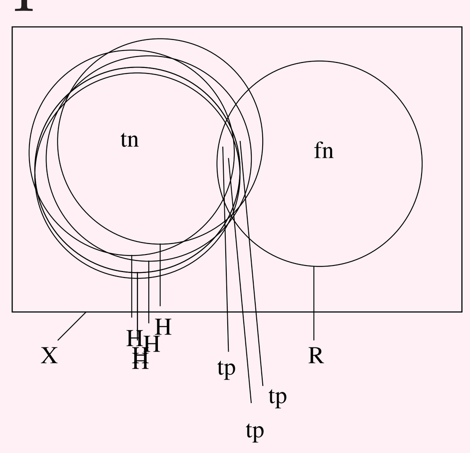
A rough set model of ambiguity

Let $\mathcal{A} = (U; A)$ be an *information system* ($\mathcal{A} \subseteq U \times A$) and let $B \subseteq A$ and $X \subseteq U$.

Lower and upper approximation of X are : $\underline{B}X = \{x/[x]_B \subseteq X\}$ $\quad \overline{B}X = \{x/[x]_B \cap X \neq \emptyset\}$

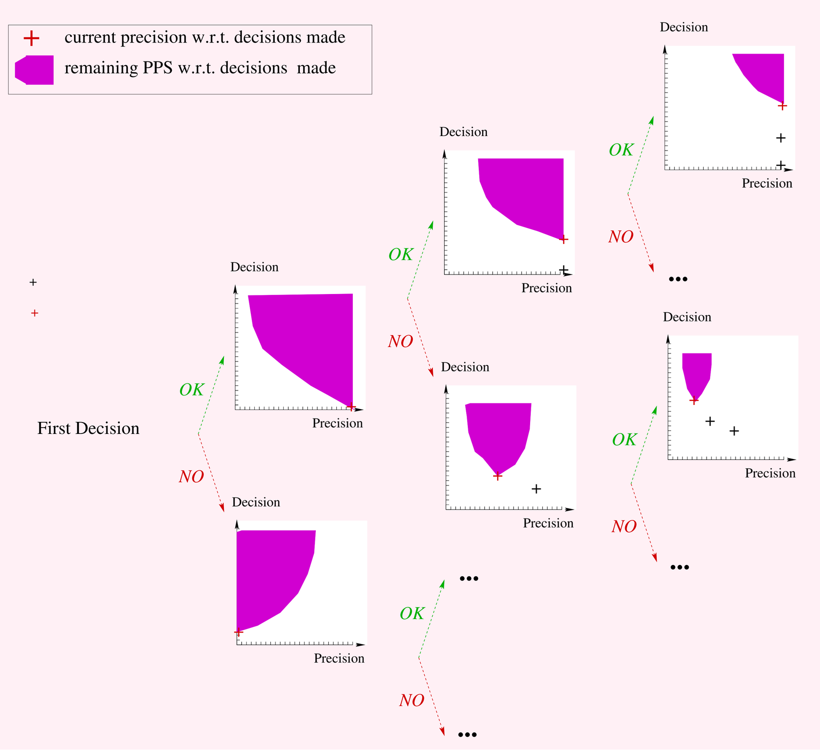
$\alpha_B(X) = \frac{|\underline{B}X|}{|\overline{B}X|}$ is the accuracy approximation coeff.

If we consider an equivalence relation $\approx$ instead of $=$



α can quantify the amount of change, e.g. for **precision** : $p.(1 - \alpha_{\approx}(H)) \leq p_{\approx} \leq p.(1 + \alpha_{\approx}(H))$

Potentential Performance Space



If ambiguity is allowed in the hypothesis data, one can ask what is the limit performance range defined by failures or successes while disambiguating the remaining (partially) undecided annotations. The potential variation defines what we call the **potential performance space**.

The measure of *decision* gauge the level of annotation disambiguation

$$D = \frac{|\{x/|[x]_{\approx}|=1\}|}{|H/_{\approx}|} = \frac{|\{x/[x]_{\approx}=\{x\}\}|}{|H/_{\approx}|}$$

The amount of performance variability due to partially disambiguated hypothesis data can be quantified with : $\alpha_R(tp) = \frac{|\underline{R}tp|}{|\overline{R}tp|}$

Example : PASSAGE parsing annotations

$$\rho = \bigcup_{j=1}^{m} \rho_j,\ m \in \mathbb{N}$$

$$\rho_1 = \bigcup_{k=1}^{q} \{r_l / l \in \mathbb{N}, r_l \subset \mathcal{P}(S^k \times A)\}$$

$$\rho_i = \bigcup_{k=1}^{u} \{r \subset \mathcal{P}((S \cup \rho_{x_1}) \times (S \cup \rho_{x_2}) \ldots \times (S \cup \rho_{x_k}) \times A), 1 \leq x_k < i\}$$

S : segmentation into words units
A : chunk and relation labels
ρ_1 : word chunk label associations, word word or word chunk relations
ρ_2 : relations between chunks only
ρ : $\rho_1 \cup \rho_2$

$H \cap R =$ continuous line relations
$H\ R =$ dashed lines relations
Decision for relations $= \frac{2}{4}$
$\alpha_R(tp) = \frac{4}{7}$