A Note on the Economics and Evaluation of Automatic Retrieval of Communication Goods

Richard Bergmair

Information & Attention

“[. . .] in an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it [. . .]”

(Simon, 1971)

Traditional Retrieval Evaluation

$$\mathbb{1}\{\text{Prec}\} = \frac{|\text{Ret} \cap \text{Rel}|}{|\text{Ret}|}$$

$$\mathbb{1}\{\text{Rec}\} = \frac{|\text{Ret} \cap \text{Rel}|}{|\text{Rel}|}$$

Economic Retrieval Evaluation

$$\mathrm{Prec} = \frac{\sum_{d|\hat{\mathrm{P}}(d)\leq \mathrm{P}(d)} \hat{\mathrm{P}}(d)}{\sum_d \hat{\mathrm{P}}(d)},$$

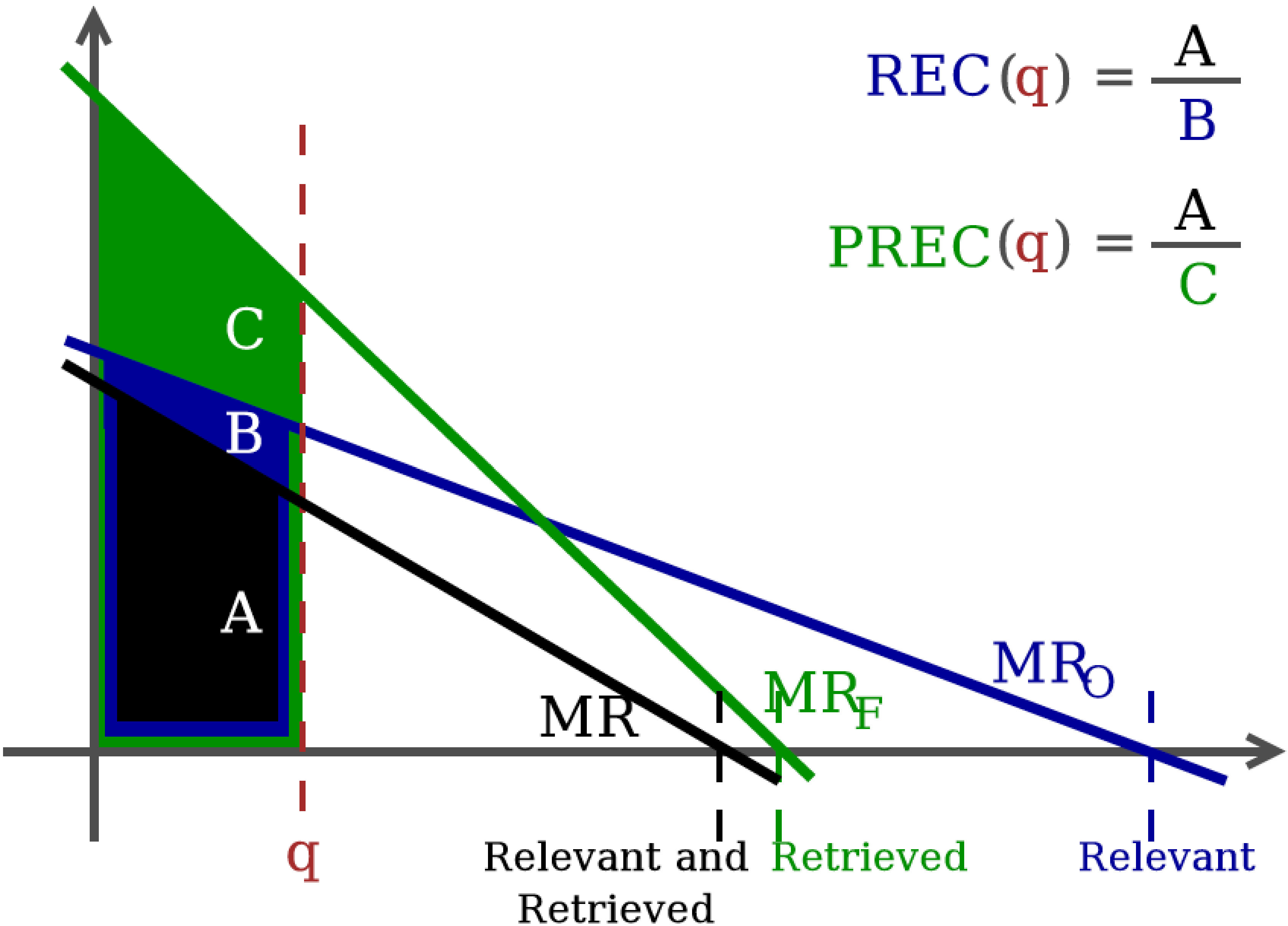
$$\mathrm{Rec} = \frac{\sum_{d|\hat{\mathrm{P}}(d)\leq \mathrm{P}(d)} \hat{\mathrm{P}}(d)}{\sum_d \mathrm{P}(d)}.$$

Retrieval Evaluation

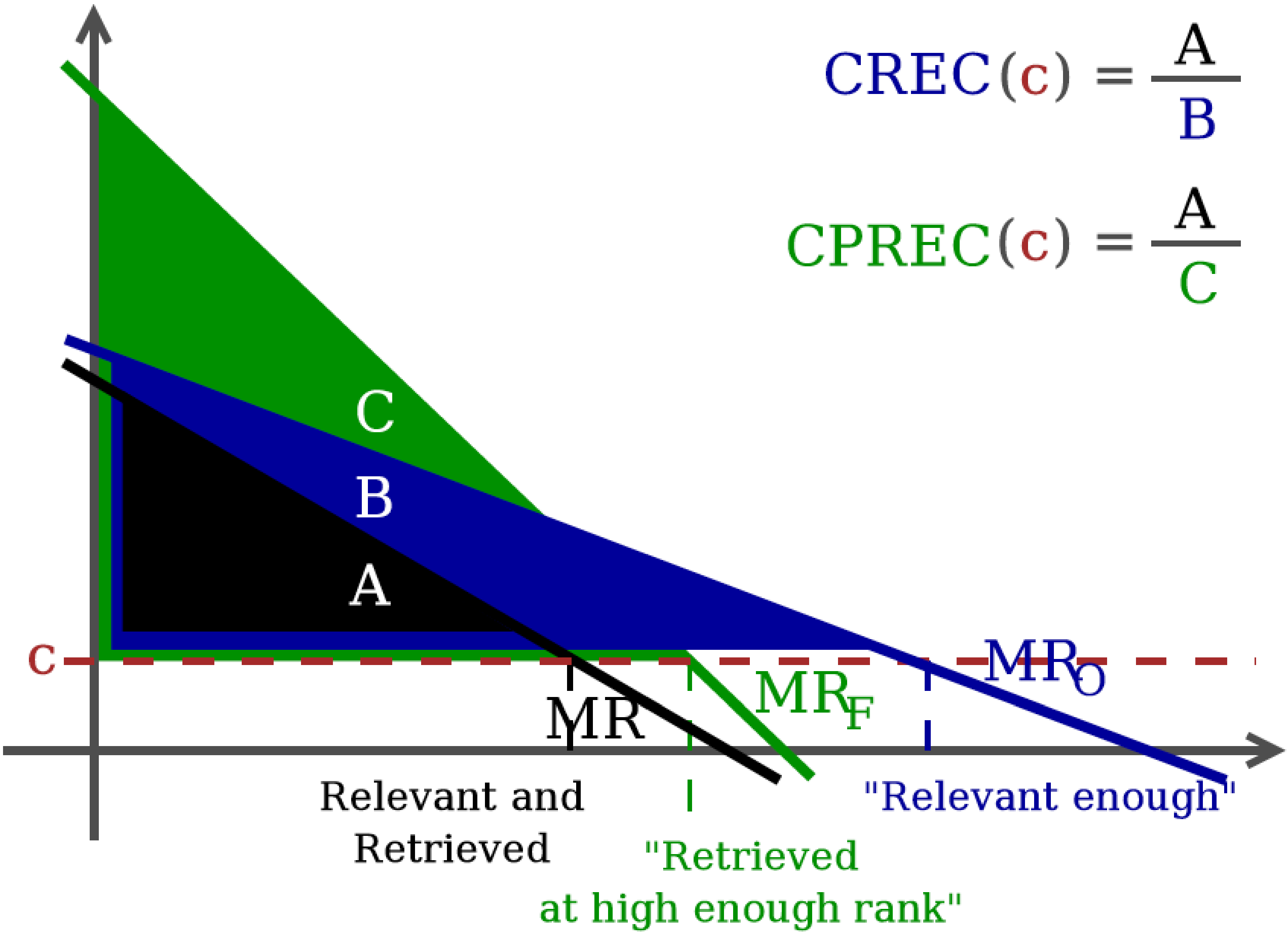
$$\mathbb{1}\{\mathrm{Prec}\} = \frac{\sum_{d|\mathbb{1}\{\hat{\mathrm{P}}(d)\}\leq\mathbb{1}\{\mathrm{P}(d)\}} \mathbb{1}\{\hat{\mathrm{P}}(d)\}}{\sum_d \mathbb{1}\{\hat{\mathrm{P}}(d)\}},$$

$$\mathbb{1}\{\mathrm{Rec}\} = \frac{\sum_{d|\mathbb{1}\{\hat{\mathrm{P}}(d)\}\leq\mathbb{1}\{\mathrm{P}(d)\}} \mathbb{1}\{\hat{\mathrm{P}}(d)\}}{\sum_d \mathbb{1}\{\mathrm{P}(d)\}}$$

Consumption Budget



Costs of Consumption



Economics of Communication (1)

$$\mathrm{P}(c,d) \stackrel{\text{def}}{=} \mathrm{P_I}(c,d) - \mathrm{P_A}(c,d),$$

$$\mathrm{Mp_P}(c,d) \stackrel{\text{def}}{=} \mathrm{Mr_A}(c,d),$$

$$\mathrm{Mp_C}(c,d) \stackrel{\text{def}}{=} \mathrm{Mr_I}(c,d) - \mathrm{Mc_A}(c,d),$$

$$\mathrm{Fc_P}(d) \stackrel{\text{def}}{=} \mathrm{Fc_I}(d).$$

Economics of Communication (2)

$$\mathrm{Tp}_{\mathrm{C}}(c) = \sum_{d \mid \mathrm{produce}(d,\mathrm{P}) \wedge \mathrm{trade}(c,d,\mathrm{P})} -\mathrm{P}(c,d) + \mathrm{Mp}_{\mathrm{C}}(c,d)$$

$$\mathrm{Tp}_{\mathrm{P}}(d) = -\mathrm{Fc}_{\mathrm{P}}(d) + \sum_{c \mid \mathrm{trade}(c,d,\mathrm{P})} \mathrm{P}(c,d) + \mathrm{Mp}_{\mathrm{P}}(c,d).$$

$$\mathrm{trade}(c,d,\mathrm{P}) \stackrel{\mathrm{def}}{\equiv} \mathrm{Mp}_{\mathrm{C}}(c,d) \geq \mathrm{P}(c,d) \geq -\mathrm{Mp}_{\mathrm{P}}(c,d)$$

$$\mathrm{produce}(d,\mathrm{P}) \stackrel{\mathrm{def}}{\equiv} \mathrm{Tp}_{\mathrm{P}}(d) \geq 0$$

IR Eval. & Economics (1)

$$= \frac{\sum_{d|\mathrm{produce}(d,\mathrm{P})} -\mathrm{Fc}_{\mathrm{I}}(d) + \sum_{c|\mathrm{trade}(c,d,\mathrm{P})} +\mathrm{Mr}_{\mathrm{A}}(c,d) + \mathrm{Mr}_{\mathrm{I}}(c,d) - \mathrm{Mc}_{\mathrm{A}}(c,d).}{\sum_d \max\left(\ 0,\ \ -\mathrm{Fc}_{\mathrm{I}}(d) + \sum_c \max\left(\ 0,\ \ +\mathrm{Mr}_{\mathrm{A}}(c,d) + \mathrm{Mr}_{\mathrm{I}}(c,d) - \mathrm{Mc}_{\mathrm{A}}(c,d)\ \right)\ \right)},$$

$$\mathrm{P}(c,d) = \mathrm{rsv}(c,d)$$

$$\mathrm{Fc}_{\mathrm{I}}(d) = 0$$

$$\mathrm{Mr}_{\mathrm{A}}(c,d) = 0$$

$$\mathrm{Mr}_{\mathrm{I}}(c,d) = \mathrm{r}(c,d)$$

IR Eval. & Economics (2)

"If your information need is *c*, how much would you be willing to pay for the option of reading *d*, or how much would you demand in compensation if you were forced to read *d*."

willing to pay -> positive r(d,c)
claim for compensation -> negative r(d,c)

$$\frac{\sum_{d,c|\mathrm{r}(d,c)\geq\mathrm{rsv}(d,c)\geq 0}\mathrm{r}(d,c)}{\sum_{d,c}\max(0,\mathrm{r}(d,c))}$$

Conclusions

characterized

- function of a retrieval engine
- evaluation model

offered an approach towards

- precision/recall tradeoff
- filtering with cost-based cutoffs