Progress Document for Bounding Box Problem

June 12, 2015

1 Problem Definition

We defined 2 parameters, α and β where,

$$\alpha = \frac{\mathcal{A}(BB \cap T)}{\mathcal{A}(T)}$$
$$\beta = \frac{\mathcal{A}(BB \cap T)}{\mathcal{A}(BB)}$$

α and β correspond to the recall and precision (respectively) w.r.t. the true bounding box T, when a worker proposes the bounding box BB. $\mathcal{A}$ is a mapping from a region to its area. We also have that $0 \leq \alpha, \beta \leq 1$ holds.

For all workers, we define a single probability density function each, for both α and β. That is, we define a single $p_\alpha(\alpha)$ and $p_\beta(\beta)$ for all workers. The probability of drawing a bounding box BB becomes,

$$p(BB|T) = p_\alpha(\alpha|T) \cdot p_\beta(\beta|T)$$

We'll omit writing "$|T$" below. Assume we have N bounding boxes given to us, each by 1 annotator. BB_i indicates the bounding box given by the i^{th} worker, and in general i is used to index annotators. We can write the likelihood function and try to maximize it w.r.t. a choice of T,

$$\mathcal{L}(T) = \prod_{i=1}^{N} p(BB_i)$$
$$= \prod_{i=1}^{N} p_\alpha(\alpha_i) \cdot p_\beta(\beta_i)$$
$$\log \mathcal{L}(T) = \sum_{i=1}^{N} [\log p_\alpha(\alpha_i) + \log p_\beta(\beta_i)]$$

2 Progress

2.1 Construction of M-L estimate by adding complete sub-regions

To construct the M-L estimate, we can add sub-regions 1 by 1, based on whether the likelihood increases or not. There can be up-to 2^N sub-regions. Each sub-region R can be identified by which k annotators proposed that region (of a total of N).

If the initial value of the likelihood function is $\mathcal{L}^{(1)}(T^{(1)})$, and adding a region R (given by k of N annotators) makes it $\mathcal{L}^{(2)}(T^{(2)})$, then we can write the following condition,

$$
\begin{aligned}
&\frac{\mathcal{L}^{(2)}}{\mathcal{L}^{(1)}} > 1 \\
\Rightarrow &\prod_{i=1}^{N} \frac{p_\alpha(\alpha_i^{(2)})}{p_\alpha(\alpha_i^{(1)})} \cdot \frac{p_\beta(\beta_i^{(2)})}{p_\beta(\beta_i^{(1)})} > 1 \\
\Rightarrow &\prod_{i=1}^{k} \underbrace{\frac{p_\alpha(\alpha_i^{(2)})}{p_\alpha(\alpha_i^{(1)})}}_{A} \cdot \underbrace{\frac{p_\beta(\beta_i^{(2)})}{p_\beta(\beta_i^{(1)})}}_{B} \cdot \prod_{i=k+1}^{N} \underbrace{\frac{p_\alpha(\alpha_i^{(2)})}{p_\alpha(\alpha_i^{(1)})}}_{C} \cdot \underbrace{\frac{p_\beta(\beta_i^{(2)})}{p_\beta(\beta_i^{(1)})}}_{D} > 1
\end{aligned}
$$

We note the following for terms A, B, C, D,

$$\alpha_i^{(1)} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(T^{(1)})} \leq \frac{\mathcal{A}(BB_i \cap T^{(2)})}{\mathcal{A}(T^{(2)})} = \frac{\mathcal{A}(BB_i \cap T^{(1)}) + \mathcal{A}(R)}{\mathcal{A}(T^{(1)}) + \mathcal{A}(R)} = \alpha_i^{(2)} \text{ for } i \in \{1...k\} \tag{A}$$

$$\beta_i^{(1)} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(BB_i)} < \frac{\mathcal{A}(BB_i \cap T^{(2)})}{\mathcal{A}(BB_i)} = \frac{\mathcal{A}(BB_i \cap T^{(1)}) + \mathcal{A}(R)}{\mathcal{A}(BB_i)} = \beta_i^{(2)} \text{ for } i \in \{1...k\} \tag{B}$$

$$\alpha_i^{(1)} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(T^{(1)})} > \frac{\mathcal{A}(BB_i \cap T^{(2)})}{\mathcal{A}(T^{(2)})} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(T^{(1)}) + \mathcal{A}(R)} = \alpha_i^{(2)} \text{ for } i \in \{(k+1)...N\} \tag{C}$$

$$\beta_i^{(1)} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(BB_i)} = \frac{\mathcal{A}(BB_i \cap T^{(2)})}{\mathcal{A}(BB_i)} = \frac{\mathcal{A}(BB_i \cap T^{(1)})}{\mathcal{A}(BB_i)} = \beta_i^{(2)} \text{ for } i \in \{(k+1)...N\} \tag{D}$$

Assume that both $p_\alpha(\alpha)$ and $p_\beta(\beta)$ are strictly monotonic, *i.e.*

$$
\begin{aligned}
p_\alpha(\alpha_2) > p_\alpha(\alpha_1) \text{ if } \alpha_2 > \alpha_1 \\
p_\beta(\beta_2) > p_\beta(\beta_1) \text{ if } \beta_2 > \beta_1
\end{aligned}
$$

We then get that $A \geq 1$, $B > 1$, $C < 1$ and $D = 1$. On adding region R, we can increase the likelihood by balancing C using A and B. 2 results follow immediately.

Result 1. If $k = N$ for R, *i.e.* R was proposed by all N annotators, then it must lie in the M-L solution.

Proof. Term C will cease to exist, so $\mathcal{L}$ will always increase on adding R. $\square$

Result 2. If $k = 0$ for R, *i.e.* R was proposed by none of the N annotators, then it will never lie in the M-L solution.

Proof. Terms A and B will cease to exist, so $\mathcal{L}$ will always decrease on adding R. $\square$

Result 3. If $k > \frac{N}{2}$ for R, *i.e.* R was proposed by atleast $\frac{N}{2}$ annotators, and we have $p_\alpha(\alpha) = p_\beta(\beta) = p(x)$ along with $p(x) \cdot p(y) = p(xy)$ then R lies in the M-L solution.

Proof. We can rewrite the condition on $\frac{\mathcal{L}^{(2)}}{\mathcal{L}^{(1)}}$ as,

$$\prod_{i=1}^{k} \underbrace{\frac{p_\alpha(\alpha_i^{(2)})}{p_\alpha(\alpha_i^{(1)})}}_{\geq 1} \cdot \underbrace{\frac{p_\beta(\beta_i^{(2)})}{p_\beta(\beta_i^{(1)})}}_{>1} > \prod_{i=k+1}^{N} \underbrace{\frac{p_\alpha(\alpha_i^{(1)})}{p_\alpha(\alpha_i^{(2)})}}_{>1}$$

$$\prod_{i=1}^{k} \underbrace{\frac{p(\alpha_i^{(2)} \cdot \beta_i^{(2)})}{p(\alpha_i^{(1)} \cdot \beta_i^{(1)})}}_{k \text{ terms}} > \prod_{i=k+1}^{N} \underbrace{\frac{p(\alpha_i^{(1)})}{p(\alpha_i^{(2)})}}_{N-k \text{ terms}}$$

Since $k > N - k$, one way for the condition to hold would be if we can find $N - k$ pairs of indices (x, y), where $x \in \{1...k\}$ and $y \in \{(k+1)...N\}$ such that, $\frac{p(\alpha_x^{(2)} \cdot \beta_x^{(2)})}{p(\alpha_x^{(1)} \cdot \beta_x^{(1)})} \geq \frac{p(\alpha_y^{(1)})}{p(\alpha_y^{(2)})}$. Since each term on the left side is > 1, the unpaired terms on the left side (at least 1 would exist) would strengthen the inequality and make it strict (if it is not). Simplifying the condition for valid pairs,

$$\begin{aligned} & p(\alpha_x^{(2)} \cdot \beta_x^{(2)}) \cdot p(\alpha_y^{(2)}) \geq p(\alpha_y^{(1)}) \cdot p(\alpha_x^{(1)} \cdot \beta_x^{(1)}) \\ \Rightarrow\ & p(\alpha_x^{(2)} \cdot \beta_x^{(2)} \cdot \alpha_y^{(2)}) \geq p(\alpha_y^{(1)} \cdot \alpha_x^{(1)} \cdot \beta_x^{(1)}) \\ \Rightarrow\ & \alpha_x^{(2)} \cdot \beta_x^{(2)} \cdot \alpha_y^{(2)} \geq \alpha_y^{(1)} \cdot \alpha_x^{(1)} \cdot \beta_x^{(1)} \end{aligned}$$

$$\Rightarrow \frac{\mathcal{A}(BB_x \cap T^{(1)}) + \mathcal{A}(R)}{\mathcal{A}(T^{(1)}) + \mathcal{A}(R)} \cdot \frac{\mathcal{A}(BB_x \cap T^{(1)}) + \mathcal{A}(R)}{\mathcal{A}(BB_x)} \cdot \frac{\mathcal{A}(BB_y \cap T^{(1)})}{\mathcal{A}(T^{(1)}) + \mathcal{A}(R)} \geq \frac{\mathcal{A}(BB_y \cap T^{(1)})}{\mathcal{A}(T^{(1)})} \cdot \frac{\mathcal{A}(BB_x \cap T^{(1)})}{\mathcal{A}(T^{(1)})} \cdot \frac{\mathcal{A}(BB_x \cap T^{(1)})}{\mathcal{A}(BB_x)}$$

$$\Rightarrow \frac{[\mathcal{A}(BB_x \cap T^{(1)}) + \mathcal{A}(R)]^2}{[\mathcal{A}(T^{(1)}) + \mathcal{A}(R)]^2} \geq \frac{[\mathcal{A}(BB_x \cap T^{(1)})]^2}{[\mathcal{A}(T^{(1)})]^2}$$

$$\therefore \mathcal{A}(T^{(1)}) \geq \mathcal{A}(BB_x \cap T^{(1)})$$

Since $BB_x \cap T^{(1)} \subseteq T^{(1)}$, this will always hold. Thus, we can always identify $N - k$ such pairs, which further means that $\mathcal{L}$ increases whenever we choose R such that $k > \frac{N}{2}$. As a side note, the multiplicative property $p(x) \cdot p(y) = p(xy)$ admits polynomial functions of the form $p(x) = (d + 1) \cdot x^d$. $\square$

2.2 Solving for fractional sub-regions

Instead of adding sub-regions piece by piece, we define a fraction γ_j of each sub-region that lies in the M-L solution. Define $X_1, ..., X_M$ as the area of the sub-regions (non-overlapping) that the given bounding boxes can be divided into ($M < 2^N$). Let γ_j be the fraction ($0 \leq \gamma_j \leq 1$) of each sub-region ($j \in \{1...M\}$) that lies in any estimate of the true bounding box. Let $\mathcal{R}_i$ denote the set of indices $\{j | X_j \subseteq BB_i\}$ *i.e.* $\mathcal{R}_i$ points to all those sub-regions that lie inside BB_i. Then, we have

$$\begin{aligned}
\mathcal{A}(T) &= \sum_{j=1}^{M} \gamma_j X_j \\
\mathcal{A}(BB_i \cap T) &= \sum_{j \in \mathcal{R}_i} \gamma_j X_j \\
\mathcal{A}(BB_i) &= \sum_{j \in \mathcal{R}_i} X_j
\end{aligned}$$

We can now write,

$$\begin{aligned}
\log \mathcal{L}(T) &= \sum_{i=1}^{N} [\log p_\alpha(\alpha_i) + \log p_\beta(\beta_i)] \\
&= \sum_{i=1}^{N} \left[\underbrace{\log p_\alpha \left(\frac{\sum_{j \in \mathcal{R}_i} \gamma_j X_j}{\sum_{j=1}^{M} \gamma_j X_j} \right)}_{quasi-concave} + \underbrace{\log p_\beta \left(\frac{\sum_{j \in \mathcal{R}_i} \gamma_j X_j}{\sum_{j \in \mathcal{R}_i} X_j} \right)}_{concave} \right]
\end{aligned}$$

The first term is quasi-concave if $p_\alpha(\cdot)$ is monotonic, since quasi-concavity is preserved under functional composition with a monotonic function. The second is term is concave if $p_\beta(\cdot)$ is concave and non-decreasing.

It is possible to solve for global maxima of quasi-concave functions; however the sum of quasi-concave functions may or may not be quasi-concave, which means that we cannot directly run a solver for the M-L estimates for γ_j's which guarantees global maxima.

2.3 Fixing the area of the true bounding box and finding M-L estimate

In the previous sub-section, if we fix the denominator of the quasi-concave term (which corresponds to setting $\mathcal{A}(T) = C$), then it becomes concave (with conditions on $p_\alpha(\cdot)$). The log-likelihood becomes overall concave (since the sum of concave terms is also concave), which guarantees the optimum assignment for γ_j's for this particular value of $\mathcal{A}(T)$.

The problem changes to testing the maximum value of the likelihood function at each value of $\mathcal{A}(T)$. The variation of $\mathcal{L}$ with $\mathcal{A}(T)$ probably has several local maximas, so a binary-search like procedure would not directly work. However, we can restrict the admissible values of area to $\mathcal{A}(\cap_{i=1}^{N} BB_i) \leq \mathcal{A}(T) \leq \mathcal{A}(\cup_{i=1}^{N} BB_i)$. Since the area can be changed incrementally with a pixel by pixel addition, there are a finite number of assignments to $\mathcal{A}(T)$ that can be exhaustively covered.

2.4 Fixing the area of the true bounding box by a heuristic

In the previous section, we can pick the area of the true bounding box ($\mathcal{A}(T)$) to be the mean or median (among other options) of the given bounding boxes. Since given bounding box areas are likely to be close to each other, we would most likely be close to the M-L estimate, if we solve for γ_j's.

We can carry out an experiment to determine how close any heuristic may be to the M-L estimate (by performing the exhaustive search outlined above). By comparing the value of the likelihood at the value of $\mathcal{A}(T)$ suggested by the heuristic to the global maximum of likelihood, we can get an idea of how well the heuristic is working.

2.5 Regions with agreement by a subset of annotators

Result 1. Given a region R_S, which lies in the bounding boxes of k_S annotators, no region R_s (or any fraction of R_s) which lies in the bounding boxes of k_s annotators where $k_s \subseteq k_S$ will lie in the M-L solution, unless R_S is completely added to the M-L solution.

Proof. Consider regions $R_S < k_S, X_S, \gamma_S >$ and $R_s < k_s, X_s, \gamma_s >$ where k_j are the set of annotators, X_j are the areas and γ_j are the area fractions included in the M-L estimate for regions R_S and R_s. $k_s \subseteq k_S$ *i.e.* k_s is a subset of the annotators in k_S.

Suppose that $\gamma_s^{(1)} = \gamma_S^{(1)} = \gamma$. Let the value of the likelihood function in this case be $\mathcal{L}^{(1)}$. Now suppose that in constructing the optimal M-L estimate, we add a fractional area $\frac{A}{X_s}$ to $\gamma_s^{(1)}$ *i.e.* we have $\gamma_s^{(2)} = \gamma + \frac{A}{X_s}$ while $\gamma_S^{(2)} = \gamma$. The likelihood in this case is $\mathcal{L}^{(2)}$ and we note that $\mathcal{L}^{(2)} > \mathcal{L}^{(1)}$.

We now claim that it possible to construct a solution having likelihood $\mathcal{L}^{(3)} > \mathcal{L}^{(2)}$ by transferring the area added to R_s, to R_S instead, by either completely transferring this area, or transferring enough that R_S is included completely. More formally, we propose the following changes,

$$\begin{aligned}
&\gamma_S^{(3)} = \gamma + \frac{A'}{X_S} \\
&\gamma_s^{(3)} = \gamma_s^{(2)} - \frac{A'}{X_s} = \gamma + \frac{(A - A')}{X_s} \\
&A' \leq A
\end{aligned}$$

Also note that,

$$\begin{aligned}
&\gamma X_S + \gamma X_s + A = \gamma X_S + A' + \gamma X_s + A - A' \\
\Rightarrow\ &\gamma_S^{(2)} X_S + \gamma_s^{(2)} X_s = \gamma_S^{(3)} X_S + \gamma_s^{(3)} X_s
\end{aligned}$$

We will now show that $\mathcal{L}^{(3)} > \mathcal{L}^{(2)}$. Following from an earlier discussion, we need to show that,

$$\prod_{i=1}^{k_s} \underbrace{\frac{p_\alpha(\alpha_i^{(3)})}{p_\alpha(\alpha_i^{(2)})}}_{A} \cdot \underbrace{\frac{p_\beta(\beta_i^{(3)})}{p_\beta(\beta_i^{(2)})}}_{B} \cdot \prod_{i=k_s+1}^{k_S} \underbrace{\frac{p_\alpha(\alpha_i^{(3)})}{p_\alpha(\alpha_i^{(2)})}}_{C} \cdot \underbrace{\frac{p_\beta(\beta_i^{(3)})}{p_\beta(\beta_i^{(2)})}}_{D} \cdot \prod_{i=k_S+1}^{N} \underbrace{\frac{p_\alpha(\alpha_i^{(3)})}{p_\alpha(\alpha_i^{(2)})}}_{E} \cdot \underbrace{\frac{p_\beta(\beta_i^{(3)})}{p_\beta(\beta_i^{(2)})}}_{F} > 1$$

where we have split the likelihood into 3 parts: k_s annotators that are common to both regions, $k_S - k_s$ annotators that lie only in the region with the superset of annotators (R_S), and all other $N - k_S$ annotators *i.e.* those that did not propose either region. Assume that $p_\alpha(\cdot)$ and $p_\beta(\cdot)$ are strictly monotonic. A sufficient condition for the inequality to hold is that each term in $\{A, ..., F\} \geq 1$, and at least 1 term is > 1. Because of the monotonicity assumption, this will also hold if the ratio of arguments in each term follows this same condition of sufficiency. We also note that,

$$\alpha_i = \frac{\sum_{j \in \mathcal{R}_i} \gamma_j X_j}{\sum_{j=1}^{M} \gamma_j X_j}$$
$$\beta_i = \frac{\sum_{j \in \mathcal{R}_i} \gamma_j X_j}{\sum_{j \in \mathcal{R}_i} X_j}$$

Observe first that the denominators for both α_i and β_i are fixed (across all terms) in moving from $\mathcal{L}^{(2)}$ to $\mathcal{L}^{(3)}$, so we need to only consider the numerators (which are identical). Thus if the required condition holds for the ratio of α_i's for some i's it will also hold for the ratio of β_i's for the same i's.

- Both $A = 1$ and $B = 1$ hold, since for the set of k_s annotators, there is no change in $\sum_{j \in \mathcal{R}_i} \gamma_j X_j$, as $\{s, S\} \in \mathcal{R}_i$ holds *i.e.* the summation always includes $\gamma_s X_s + \gamma_S X_S$ which remains constant (and all other terms are unchanged) in moving from $\mathcal{L}^{(2)}$ to $\mathcal{L}^{(3)}$.

- Both $E = 1$ and $F = 1$ hold, since for the set of $N - k_S$ annotators, there is no change in regions concerning these annotators.

- Both $C > 1$ and $D > 1$ hold, since for the set of $k_S - k_s$ annotators, $\sum_{j \in \mathcal{R}_i} \gamma_j X_j$ has $\{S\} \in \mathcal{R}_i$ and $\{s\} \notin \mathcal{R}_i$. Thus, the summation grows on transferring A' to R_S.

Using the above, we see that $\mathcal{L}^{(3)} > \mathcal{L}^{(2)}$ holds which proves the result.

Thus, if a particular region R_S is not completely included in the M-L estimate (i.e. $\gamma_S < 1$), then no region R_s (labeled by a subset of annotators of R_S) can ever lie in the M-L estimate. This gives us a strategy to prune away candidate regions from being part of the M-L estimate.

$\square$

Result 1 (Simpler Restatement). A region proposed by k annotators cannot lie in the M-L solution unless all regions proposed by K annotators (such that $k \subset K$) are part of the M-L solution.

Proof. Assume on the contrary that the M-L solution T_{ML} has a region R_k proposed by k annotators, and there exists a region R_K proposed by K annotators ($k \subset K$) which is not a part of the M-L solution.

Consider a region $T^{'}$, which is identical to T_{ML}, except for the fact that R_K is completely added before R_k starts being added in, in such a way that the total area remains the same *i.e.* $\mathcal{A}(T') = \mathcal{A}(T_{ML})$. We show that the likelihood of $T^{'}$, $\mathcal{L}(T^{'})$, is more than the likelihood of T_{ML}, $\mathcal{L}(T_{ML})$, contradicting the fact that T_{ML} is the M-L solution.

The likelihood of a bounding box estimate is the product (over all annotators) of probabilities of that box, given all annotations. Note that $T^{'}$ and T_{ML} have identical areas. Also note that, the bounding boxes provided by the workers are fixed and so, their area is also fixed. Therefore, the denominators in $\alpha_i = \frac{\mathcal{A}(T \cap BB_i)}{\mathcal{A}(T)}$ and $\beta_i = \frac{\mathcal{A}(T \cap BB_i)}{\mathcal{A}(BB_i)}$ remain the same for $T = \{T^{'}, T_{ML}\}$. Further, assume that $p_\alpha(\cdot)$ and $p_\beta(\cdot)$ are strictly monotonic. We may note that,

- For annotators in k, $\mathcal{A}(T^{'} \cap BB_i) = \mathcal{A}(T_{ML} \cap BB_i)$. This is because these annotators propose both R_k and R_K. Therefore the likelihoods of $T^{'}$ and T_{ML} are the same for these annotators.

- For annotators in $K \backslash k$, $\mathcal{A}(T^{'} \cap BB_i) > \mathcal{A}(T_{ML} \cap BB_i)$. This is because these annotators propose R_K, but not R_k. The likelihood of $T^{'}$ is greater than that of T_{ML} for these annotators.

- For all other annotators not in K, $\mathcal{A}(T^{'} \cap BB_i) = \mathcal{A}(T_{ML} \cap BB_i)$ since these annotators propose neither regions R_k nor R_K. Therefore the likelihoods of $T^{'}$ and T_{ML} are the same for these annotators.

The above observations together imply that $\mathcal{L}(T) > \mathcal{L}(T_{ML})$, thereby contradicting the fact that T_{ML} is the M-L solution. $\square$

2.6 Attempts to prove NP-Hardness of likelihood maximization

To begin, we restrict ourselves to the specific case where the probability density function is multiplicative, *i.e.* $p(x) \cdot p(y) = p(xy)$ and $p_\alpha(\alpha) = p_\beta(\beta)$. We get the following simplified form of the likelihood,

$$\begin{aligned}
\mathcal{L} &= p\left(\prod_{i=1}^{N} \alpha_i \cdot \beta_i\right) \\
&= p\left(\prod_{i=1}^{N} \frac{\sum_{j=1}^{M} \gamma_j X_j \mathcal{R}_j(i)}{\sum_{j=1}^{M} \gamma_j X_j} \cdot \frac{\sum_{j=1}^{M} \gamma_j X_j \mathcal{R}_j(i)}{\sum_{j=1}^{M} X_j \mathcal{R}_j(i)}\right)
\end{aligned}$$

Maximizing this expression is identical to maximizing,

$$\prod_{i=1}^{N} \frac{\left[\sum_{j=1}^{M} \gamma_j X_j \mathcal{R}_j(i)\right]^2}{\sum_{j=1}^{M} \gamma_j X_j}$$

We consider the case where $\gamma_j \in \{0, 1\}$ for this problem.

2.6.1 Fixing the value of the true bounding box

If we fix the value of $\sum_{j=1}^{M} \gamma_j X_j$, then the problem simplifies to maximizing,

$$\prod_{i=1}^{N} \left[\sum_{j=1}^{M} \gamma_j X_j \mathcal{R}_j(i) \right]$$

However, proving NP-Hardness for this optimization problem would not automatically lead to our original problem being NP-Hard, since our original problem is not a generalization of this problem, but is a different problem. Thus, this approach does not lead us to the required result.

2.6.2 Simplifications

The best direction towards proving our problem to be NP-Hard is to take a simpler instance of our problem and prove it to be NP-Hard. A clear choice is to set $X_j = 1$, making each sub-region equal in area. The quantity to maximize (call it $\mathcal{Q}$) then becomes,

$$\mathcal{Q} = \prod_{i=1}^{N} \frac{\left[\sum_{j=1}^{M} \gamma_j \mathcal{R}_j(i) \right]^2}{\sum_{j=1}^{M} \gamma_j}$$

which is slightly easier to work with. Another simplification that seems possible is to pick a small, fixed value for N. A thing to note here is the dependence of M on N *i.e.* $M \leq 2^N$. Thus, it is clearly counterintuitive to prove NP-hardness for small values of N, and in addition, there seem to be polynomial strategies for cases where $X_j = 1$. Thus, it seems it there is no straightforward simplification to be made to the product.

A different approach to simplification is in terms of the structure with which the N bounding boxes overlap. This can be done by picking special values for $\mathcal{R}_j(i)$ that may make the problem slightly easier for reductions. However, the quadratic nature of the numerator still poses a problem in this case. A different way of writing $\mathcal{Q}$ is,

$$\mathcal{Q} = \prod_{i=1}^{N} \left[\frac{\sum_{j=1}^{M} \sum_{k=1}^{M} \gamma_j \gamma_k \mathcal{R}_{jk}(i)}{\sum_{j=1}^{M} \gamma_j} \right]$$

where $\mathcal{R}_{jk}(i)$ can be built by combining $\mathcal{R}_j(i)$ and $\mathcal{R}_k(i)$, which bears a resemblance to the Quadratic Knapsack problem, without a clear way to reduce from it.

A combination of these strategies may help in bringing about the reduction.

2.6.3 Partially ordered set of regions

Using *Result 1* from the previous section, it is clear that there is an order that must be followed when adding in regions to the M-L estimate (equivalent to making $\gamma_i = 1$ for that

Figure 1: Partially ordered set on the inclusion relation

region). This order lends itself to the description of a poset ordered on an inclusion relation as shown in Figure 1. Thus, the intersection of all annotators must be added first, before moving down to the next layer. An element in the poset is accessible only if all its children are already added to the M-L estimate. The width of any level of this poset is $\binom{N}{k}$ which is maximum for $N = \frac{N}{2}$.

The NP-Hard reduction may be carried out on the original problem, since using the poset amounts to a generalization of the original problem (which has all nodes at the same level), thus if the original problem formulation is NP-Hard, this must also be.

2.6.4 Resemblance to the fractional sum-of-ratios problem

A paper by Roland and Jarre ("Solving the Sum-of-Ratios problem by an Interior Point Method") outlines a proof of NP-Hardness for the sum-of-ratios optimization problem, using a reduction from Knapsack. They consider the problem of maximizing,

$$h(x) + \sum_{i=1}^{N} \frac{f_i(x)}{g_i(x)}$$

subject to convex constraints. In addition, $h(x)$ and $f_i(x)$ must be concave, while $g_i(x)$ must be convex. While our problem is a product rather than a sum, our constraints are similar since our numerator is the square of a positive linear function (concave), while the denominator is also a linear function. Their proof is instructive and may be useful in constructing a reduction from Knapsack.

Gao, Mishra and Shi (2010) ("An Extension of Branch-and-Bound Algorithm for Solving Sum-of-Nonlinear-Ratios Problem") include a proof that is similar in spirit to the one above, but is instead a reduction from the Traveling Salesman Problem.

2.6.5 Resemblance to the Knapsack problem

Several versions of the Knapsack problem seem to be relevant; in particular, the Precedence Constraint Knapsack Problem encodes a precedence DAG (exactly similar to our poset), in its problem definition. The Quadratic Knapsack Problem has a quadratic form for the function being maximized similar to our numerator. In addition, a Multi-Dimensional Knapsack approach may also be useful.

2.6.6 NP-Hard Reduction

Consider the following optimization problems,

$$
\begin{aligned}
\mathsf{P}_1: \quad & \underset{\gamma_j, 1\le j\le M}{\text{maximize}} & \mathcal{Q}_1 = \prod_{i=1}^{N} \frac{\left[\sum_{j=1}^{M} \gamma_j X_j \mathcal{R}_{ij}\right]^2}{\sum_{j=1}^{M} \gamma_j X_j} \\
& \text{subject to} & 0 < \sum_{j=1}^{M} \gamma_j X_j \le T \\
& & \gamma_j \in \{0,1\},\ j = 1, \ldots, M. \\
& \text{given} & X_j > 0,\ j = 1, \ldots, M. \\
& & \mathcal{R}_{ij} \in \{0,1\},\ i = 1, \ldots, N,\ j = 1, \ldots, M.
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{P}_2: \quad & \underset{\gamma_j, 1\le j\le M}{\text{maximize}} & \mathcal{Q}_2 = \prod_{i=1}^{N} \frac{\left[\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij}\right]^2}{\sum_{j=1}^{M} \gamma_j} \\
& \text{subject to} & 0 < \sum_{j=1}^{M} \gamma_j \le T \\
& & \gamma_j \in \{0,1\},\ j = 1, \ldots, M. \\
& \text{given} & \mathcal{R}_{ij} \in \{0,1\},\ i = 1, \ldots, N,\ j = 1, \ldots, M.
\end{aligned}
$$

Note that if P_2 is NP-Hard, then so is P_1, since P_1 is a generalization of P_2. Thus, it is sufficient to show that P_2 is NP-Hard. Also note that our original problem, P is a generalization of P_1, as it is defined on a poset over γ_j's. P_1 is the specific case where the poset imposes no order, and thus it is sufficient to show that P_1 is NP-Hard.

We now define the decision form of the problem P_2. Instead of maximizing, we ask if there is a solution such that $\mathcal{Q}_2 > L$ holds, subject to the given constraints. Since the optimization problem is always as hard as the decision problem, we will prove that the decision problem is NP-Hard.

Claim 1. The decision version of P_2 is NP-Hard.

Proof. We show a reduction from the Set Cover problem as proof. Consider the Set Cover problem, with a universal set of elements (that are to be covered) $\mathcal{U}$ and a set $\mathcal{S}$ of candidate covering sets, that are defined as follows,

$$
\begin{aligned}
\mathcal{U} &= \{1, 2, \ldots, N\},\ |\mathcal{U}| = N \\
\mathcal{S} &= \{S_1, S_2, \ldots, S_M\},\ |\mathcal{S}| = M \\
S_j &\subset \mathcal{U},\ j \in \{1, \ldots, M\}
\end{aligned}
$$

We define an indicator function $\mathcal{R}_{ij}$, which denotes the presence of i in S_j *i.e.* $\mathcal{R}_{ij} = 1$ if $i \in S_j$ otherwise $\mathcal{R}_{ij} = 0$. Lastly, we let $\gamma_j \in \{0,1\},\ j = 1, \ldots, M$ denote the presence or absence of S_j in any candidate set cover. We define the problem as checking if there exists a set cover of size at-most T.

Now, in the decision version of P_2, we set $L = 0$,

$$\begin{aligned}
&\prod_{i=1}^{N} \frac{\left[\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij}\right]^2}{\sum_{j=1}^{M} \gamma_j} > 0 \\
&\Rightarrow \prod_{i=1}^{N} \left[\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij}\right] > 0 \\
&\quad \Rightarrow \left[\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij}\right] > 0,\ i = 1, \ldots, N
\end{aligned}$$

Observe now that the Set Cover problem is equivalent to the decision form of P_2. If there exists a set cover that satisfies the given constraints, then each element of $\mathcal{U}$ would be covered *i.e.* for each $i = 1, \ldots, N$ we have picked atleast one $j = 1, \ldots, M$ such that $\mathcal{R}_{ij} = 1$. Thus, for each such i, $\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij} \geq 1 > 0$, which tells us that this is also a satisfying assignment for P_2.

Conversely, if there exists a satisfying assignment for P_2 such that $\sum_{j=1}^{M} \gamma_j \mathcal{R}_{ij} > 0,\ i = 1, \ldots, N$, then for every such i, there is guaranteed to be a $j = 1, \ldots, M$ such that $\gamma_j = 1$ and $\mathcal{R}_{ij} = 1$. This indicates that for every assignment of i, there is at-least one S_j that covers it. Thus this is also a satisfying assignment for the Set Cover problem.

The reduction is in polynomial-time since both problems are equivalent. Therefore, the decision version of P_2 is NP-Hard. □

2.7 Observations about the error model

Our worker error model encapsulates 2 quantities,

$$\begin{aligned}
R\text{ecall: } \alpha &= \frac{\mathcal{A}(BB \cap T)}{\mathcal{A}(T)} \\
P\text{recision: } \beta &= \frac{\mathcal{A}(BB \cap T)}{\mathcal{A}(BB)}
\end{aligned}$$

Observe that we have implicitly modeled our problem in 2 distinct ways:

- *Separate Distributions*: Our attempt to model P and R as separate processes in drawing a bounding box lead us to define $p_\alpha(\alpha)$ and $p_\beta(\beta)$ separately, and carry out M-L estimation using this model.

- *Single Distribution*: The simplification made by making both distributions identical, and assuming a multiplicative probability density function leads to a much simpler form of the likelihood function. However, this simplification can be viewed somewhat differently; we can define a probability density function over the geometric mean of α and β *i.e.* over $p(\sqrt{\alpha\beta})$, and if we make the function multiplicative, we get an equivalent (not identical) likelihood function.

This then leads us to the possibility of defining a distribution over some combination of α and β, which may be mathematically easier to work with. This combination can give equal weight to R and P, or different weights. Ideally, it would be convenient if it defined over $[0, 1]$ and attains its maximum when both R and P are maximum. Some examples of this kind are,

- *F-Measure (Harmonic Mean)*, which is defined as $\frac{2\alpha\beta}{\alpha+\beta} = \frac{2\mathcal{A}(BB \cap T)}{\mathcal{A}(BB)+\mathcal{A}(T)}$. Note that for this definition (under a multiplicative function), the result guaranteeing that regions given by at least half annotators lie in the M-L solution does not hold. However, the optimization problem is easier for proving NP-Hardness on, as it has a slightly easier form (no quadratic numerator) to work with.

- *Arithmetic Mean*, which is defined as $\frac{\alpha+\beta}{2}$.

- *Jaccard Similarity*, which is defined as $\frac{\mathcal{A}(BB \cap T)}{\mathcal{A}(BB \cup T)}$.

A somewhat different but interesting form is $\frac{P}{1-R} = \frac{\beta}{1-\alpha}$, which takes values on $[0, \infty)$, and on which we can choose a probability density function of the form $p(x) \cdot p(y) = p(x + y)$ which lends itself to exponential distributions.