# THE BLAHUT ALGORITHM FOR THE COMPUTATION OF THE RATE-DISTORTION FUNCTION

William A. Pearlman

2002

Reference: R. Blahut – IEEE Trans. Inform. Thy., July 1972

The development here is somewhat different than that given in the above paper. The notation is chosen to conform to that of the paper and subsequent computer programs.

Assume a discrete memoryless source (DMS) with letter probabilities $\{p_j\}$ and distortion measure $\rho_{jk}$ between the $j$th source letter and $k$th reproduction letter. The rate-distortion function is

$$R(D) = \min_{Q \mathcal{E} Q_D} \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j}}{\sum_{j'} p_{j'} Q_{k/j'}}$$

$$Q_E = \{Q_{k/j} : Q_{k/j} \geq 0, \sum_k Q_{k/j} = 1, \text{ and} \sum_j \sum_k p_j Q_{k/j} \rho_{jk} \leq D\}$$

In words, we minimize the average mutual information overall all test channels $\{Q_{k/j}\}$ yielding average distortion less than or equal to $D$. The average mutual information is a function of these test channel transition probabilities $Q = \{Q_{k/j}\}$ for a given source probability distribution $\{p_j\}$. To emphasize this fact, let us denote it by

$$J(Q) \equiv \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j}}{\sum_{j'} p_{j'} Q_{k/j'}}$$

Consider a generalization of $J(Q)$ obtained by substituting a general set of output probabilities $\{q_k\}$ not linked to the test channel and source. Define

$$J(Q, q) = \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j}}{q_k}$$

**Lemma 1:**

For given $Q$, $J(Q,q) \geq J(Q)$ with equality iff $q_k = \sum_j p_j Q_{k/j}$, all $k$

Proof: Express $J(Q,q) - J(Q)$ and use divergence inequality.

¿From the definition of $J(Q,q)$ and the result of the lemma,

$$R(D) = \min_{Q \mathcal{E} Q_D} J(Q) = \min_{Q \mathcal{E} Q_D} \min_q J(Q,q)$$

Since the two minimizations can be performed in any order,

$$R(D) = \min_{Q \mathcal{E} Q_D} \min_q J(Q,q) = \min_q \min_{Q \mathcal{E} Q_D} J(Q,q)$$

Consider now the function

$$J(Q,q) - s(\sum_j \sum_k p_j Q_{k/j} \rho_{jk} - D)$$

By the method of Lagrange multipliers, we choose $s$ to satisfy the average distortion constraint and minimize the above over all $Q$. Therefore,

$$R(D) = sD + \min_q \min_Q [J(Q,q) - s \sum_j \sum_k p_j Q_{k/j} \rho_{jk}]$$

Define the expression enclosed by brackets [ ] above as $F(Q,q)$. Using an analogous notational convention, we note $F(Q,q)$ and $F(Q)$ obey Lemma 1 with the same equality condition since $F(Q,q) - F(Q) = J(Q,q) - J(Q)$.

$$\begin{aligned} F(Q,q) &= J(Q,q) - s \sum_j \sum_k p_j Q_{k/j} \rho_{jk} \\ &= \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j} e^{-s\rho_{jk}}}{q_k} \end{aligned}$$

The last equality is obtained by substituting the definition of $J(Q,q)$ and collecting terms. That brings us to the following lemma.

**Lemma 2**

For fixed $q$,

$$F(Q,q) \geq -\sum_j p_j \log(\sum_k q_k e^{s\rho_{jk}})$$

with equality iff

$$Q_{k/j} = \frac{q_k e^{s\rho_{jk}}}{\sum_{k'} q_{k'} e^{s\rho_{jk'}}}, \quad s \leq 0$$

<u>Proof:</u>

$$F(Q,q) = \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j}(\sum_{k'} q_{k'} e^{s\rho_{jk'}})}{q_k e^{s\rho_{jk}}(\sum_{k'} q_{k'} e^{s\rho_{jk'}})}$$

$$f_{k/j} = \frac{q_k e^{s\rho_{jk}}}{\sum_{k'} q_{k'} e^{s\rho_{jk'}}} \text{ is a probability distribution over } k \text{ for } s \leq 0 \text{ and any } j.$$

$$F(Q,q) = \sum_j \sum_k p_j Q_{k/j} \log \frac{Q_{k/j}}{f_{k/j}} + \sum_j \sum_k p_j Q_{k/j} \log \lambda_j$$

$$\lambda_j = (\sum_{k'} q_{k'} e^{s\rho_{jk'}})^{-1}$$

By the divergence inequality, the first term $\geq 0$ and equals 0 iff $f_{k/j} = Q_{k/j}$. Therefore,

$$\begin{aligned} F(Q,q) &\geq \sum_j \sum_k p_j Q_{k/j} \log \lambda_j = \sum_j p_j \log \lambda_j \\ &= -\sum_j p_j \log(\sum_k q_k e^{s\rho_{jk}}) \end{aligned}$$

with equality iff $Q_{k/j} = q_k e^{s\rho_{jk}} \Big/ \sum_{k'} q_{k'} e^{s\rho_{jk'}}$

By the above lemma;

$$R(D) = sD + \min_q [-\sum_j p_j \log(\sum_k q_k e^{s\rho_{jk}})]$$

where $D = \sum_j \sum_k p_j Q_{k/j} \rho_{jk}$ with the $Q_{k/j}$ given in the equality statement of Lemma 2. Note that for any $q$, we have an upper bound for $R(D)$ given by

$$R(D) \leq sD - \sum_j p_j \log(\sum_k q_k e^{s\rho_{jk}}).$$

Now we have the basis for an algorithm, for we have expressed $R(D)$ as

$$R(D) = sD + \min_q \min_Q F(Q, q)$$

and found the solution to $\min_Q F(Q, q)$ for fixed $q$ and $\min_q F(Q, q)$ for fixed $Q$.

Because the double minimization can be taken in any order, we initially fix $q$ and find the minimizing $Q$ by Lemma 2. Then fix to this $Q$ and find the minimizing $q$ by Lemma 1 and then repeat. The explicit steps are as follows:

1. Set some $s \leq 0$.
2. Set iteration index $\ell = 0$, initialize $q = q^{(0)}$.
3. Let $q = q^{(\ell)}$, the minimizing $Q$ is

$$Q^{(\ell)}_{k/j} = \frac{q^{(\ell)}_k e^{s\rho_{jk}}}{\sum_{k'} q^{(\ell)}_{k'} e^{s\rho_{jk'}}} \text{ all } j, k.$$

4. For $Q$ given in 3., the minimizing $q$ is

$$q^{(\ell+1)}_k = \sum_j p_j Q^{(\ell)}_{k/j} \text{ all } k$$

5. Let $\ell \to \ell + 1$ and go to 3.

As proved in Lemma 1, before and after Step 4, $F(Q^{(\ell)}, q^{(\ell+1)}) \leq F(Q^{(\ell)}, q^{(\ell)})$.

As proved in Lemma 2, proceeding through Step 3, $F(Q^{(\ell+1)}, q^{(\ell+1)}) \leq F(Q^{(\ell)}, q^{(\ell+1)})$.

The value of $F$ is monotone non-increasing through the given iteration and eventually reaches a minimum. (The proof will be suggested, but not explicitly given.) We seek now a stopping criterion applied after Step 4 and before incrementing the iteration index. To this end, we develop upper and lower bounds to $R(D)$, which we calculate after Step 4. First note that Steps 3 and 4 can be combined as follows:

$$
\begin{aligned}
q_k^{(\ell+1)} &= \sum_j p_j \frac{q_k^{(\ell)} e^{s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}} = q_k^{(\ell)} c_k^{(\ell)} \\
c_k^{(\ell)} &= \frac{\sum_j p_j e^{s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}}
\end{aligned}
$$

$c_k^{(\ell)}$ is an updating coefficient for the output probability $q_k^{(\ell)}$ giving the new $q_k^{(\ell+1)}$.

Let us now turn toward the upper and lower bounds for $R(D)$.

The upper bound is readily apparent, since every step of the iteration must yield an average mutual information no smaller than $R(D)$. From this fact, we prove the following lemma:

**Lemma 3:**

$$
R(D) \leq sD - \sum_j p_j (\sum_k q_k^{(\ell)} e^{s\rho_{jk}}) - \sum_k q_k^{(\ell)} \log c_k^{(\ell)}
$$

Proof: As any step gives an upper bound, after Step 4, we have

$$
\begin{aligned}
R(D) \leq sD + F(Q^{(\ell)}, q^{(\ell)}) &= I(Q^{(\ell)}, q^{(\ell+1)}) \\
&= \sum_j \sum_k p_j Q_{k/j}^{(\ell)} \log \frac{Q^{(\ell)}}{q_k^{(\ell+1)}} \\
&= \sum_j \sum_k p_j Q_{k/j}^{(\ell)} \log \frac{q_k^{(\ell)} e^{s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}} \times \frac{1}{q_k^{(\ell+1)}} \\
&= \sum_k p_j Q_{k/j}^{(\ell)} [\log \frac{1}{c_k^{(\ell)}} + \log \frac{e^{s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}}] \\
&= -\sum_k q_k^{(\ell)} \log c_k^{(\ell)} + sD - \sum_j p_j \log (\sum_k q_k^{(\ell)} e^{s\rho_{jk}})
\end{aligned}
$$

The lower bound is more difficult to obtain. We state it in the following lemma:

**Lemma 4:**

$$R(D) \geq sD - \sum_j p_j \log(\sum_k q_k^{(\ell)} e^{s\rho_{jk}}) - \max_k \log c_k^{(\ell)}$$

Proof: Suppose $Q^*_{k/j}$ achieves $R(D)$, i.e.,

$$R(D) = sD + \sum_j \sum_k p_j Q^*_{k/j} \log \frac{Q^*_{k/j} e^{-s\rho_{jk}}}{q_k^*}$$

Denote $P^*_{j/k} = \frac{Q^*_{k/j} p_j}{q_k^*}$, the optimum backward test channel.

$$R(D) = sD + \sum_j \sum_k p_j Q^*_{k/j} \log \left[ \frac{P^*_{j/k} e^{-s\rho_{jk}}}{p_j} \times \frac{P^{(\ell)}_{j/k}}{P^{(\ell)}_{j/k}} \right]$$

$$= sD + \sum_j \sum_k q_k^* P^*_{j/k} \log \frac{P^*_{j/k}}{P^{(\ell)}_{j/k}} + \sum_j \sum_k p_j Q^*_{k/j} \log \frac{P^{(\ell)}_{j/k} e^{-s\rho_{jk}}}{p_j}$$

$$P^{(\ell)}_{j/k} = \frac{Q^{(\ell)}_{k/j} p_j}{\sum_{j'} Q^{(\ell)}_{k/j'} p_{j'}} = \frac{Q^{(\ell)}_{k/j} p_j}{q_k^{(\ell+1)}} \text{ is the } \ell\text{th backward test channel}$$

By the divergence inequality, the second term $\geq 0$. Substituting,

$$R(D) \geq sD + \sum_j \sum_k p_j Q^*_{k/j} \log \frac{Q^{(\ell)}_{k/j} e^{-sP_{jk}}}{q_k^{(\ell+1)}}$$

$$= sD + \sum_j \sum_k p_j Q^*_{k/j} \log[\frac{1}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}} \times \frac{q_k^{(\ell)}}{q_k^{(\ell+1)}}]$$

$$= sD + \sum_j p_j \log(\sum_k q_k^{(\ell)} e^{s\rho_{jk}}) - \sum_k q_k^* \log c_k^{(\ell)}$$

$$R(D) \geq sD + \sum_j p_j \log(\sum_k q_k^{(\ell)} e^{s\rho_{jk}}) - \max_k \log c_k^{(\ell)}$$

The upper bound of Lemma 3 is equal to the average mutual information after Step 4. So we stop the algorithm when, for some $\epsilon > 0$,

$$sD + F(Q^{(\ell)}, q^{(\ell+1)}) = I(Q^{(\ell)}, q^{(\ell+1)}) < R_L(D) + \epsilon$$

where $R_L(D)$ is the lower bound in Lemma 4. Substituting these quantities, we stop when

$$\max_k \log c_k^{(\ell)} - \sum_k q_k^{(\ell)} \log c_k^{(\ell)} < \epsilon$$

and guarantee that $I(Q^{(\ell)}, q^{(\ell+1)})$ is no more than $\epsilon$ greater than $R(D)$, since $R(D) \geq R_L(D)$.

We remark now on the convergence. By the convex $\cap$ property of log, we see that

$$\sum_k q_k^{(\ell)} \log c_k^{(\ell)} \leq \log(\sum_k q_k^{(\ell)} c_k^{(\ell)})$$
$$\max_k \log c_k^{(\ell)} - \log(\sum_k q_k^{(\ell)} c_k^{(\ell)}) < \epsilon \text{ at stopping time.}$$

or

$$0 < \log\left(\frac{\max_k c_k^{(\ell)}}{\sum_{k'} q_{k'}^{(\ell)} c_{k'}^{(\ell)}}\right) < \epsilon \text{ (maximum greater than the average)}$$

So

$$\max_k c_k^{(\ell)} \longrightarrow \overline{c_k^{(\ell)}} \text{ (maximum approaching average)}$$

and so the $c_k^{(\ell)}$'s, where $q_k^{(\ell)} \neq 0$ are approaching equality in the limit. In fact, they all must be approaching 1, because in the limit no updating is required.

Let us now restate the algorithm with Steps 3 and 4 combined and the stopping criterion included.

1. Input $s \leq 0$.
2. Set iteration index $\ell = 0$, initialize $q = q^{(0)}$.

3. With $q = q^{(\ell)}$, let

$$q_k^{(\ell+1)} = q_k^{(\ell)} c_k^{(\ell)}, c_k^{(\ell)} = \frac{\sum_j p_j e^{-s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{s\rho_{jk'}}}$$

4. If $[\max_k c_k^{(\ell)} - \sum_k q_k^{(\ell)} \log c_k^{(\ell)}] < \epsilon$, go to 6.

5. $\ell \to \ell + 1$ and go to 3.

6.

$$\begin{aligned} R(D) &= sD - \sum_j p_j (\sum_k q_k^{(\ell)} e^{s\rho_{jk}}) - \sum_k q_k^{(\ell)} \log c_k^{(\ell)} \\ D &= \sum_j \sum_k p_j Q_{k/j}^{(\ell)} \rho_{jk} \\ Q_{k/j}^{(\ell)} &= \frac{q_k^{(\ell)} e^{s\rho_{jk}}}{\sum_{k'} q_{k'}^{(\ell)} e^{-s\rho_{jk'}}} \end{aligned}$$

This algorithm calculates one point of $R(D)$. Return to 1 and input another $s \leq 0$ for another point.