# Arimoto-Blahut Algorithms for Computing Channel Capacity and Rate-Distortion with Side-Information

Frédéric Dupuis and Wei Yu
University of Toronto

Frans Willems
Technische Universiteit Eindhoven

June 29, 2004

ISIT 2004, Chicago

# Outline

- Goal: Numerical computation of channel capacity and rate distortion with side-information
- Main results: Arimoto-Blahut algorithms for
  - Channel capacity with side-information
  - Source coding with side-information
  - Broadcast Channels
- Main technique: Shannon strategies
- Conclusions

# Channel capacity computation

$$X \longrightarrow \boxed{\; p(y|x) \;} \longrightarrow Y$$

The goal: solve the optimization problem:

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} \sum_{x,y} p(x)p(y|x) \log \frac{p(x|y)}{p(x)}$$

Arimoto-Blahut algorithm: Treat $p(x)$ and $p(x|y)$ as independent variables.

$$p(x) \to q(x)$$

$$p(x|y) \to Q(x|y)$$

# Arimoto-Blahut algorithm for channel capacity

$$C = \max_{q(x), Q(x|y)} \sum_{x,y} q(x) p(y|x) \log \frac{Q(x|y)}{q(x)}$$

Fix $q(x)$:

$$Q^*(x|y) = \frac{q(x)p(y|x)}{\sum_{x'} q(x')p(y|x')}$$

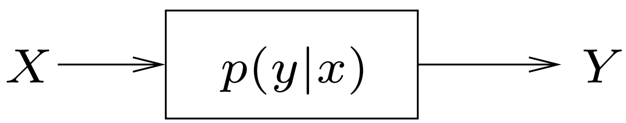
Fix $Q(x|y)$:

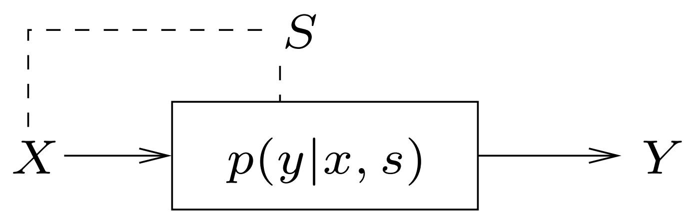
$$q^*(x) = \frac{\prod_y Q(x|y)^{p(y|x)}}{\sum_{x'} \prod_y Q(x'|y)^{p(y|x')}}$$

The algorithm:

$$q^{(1)}(x) \to Q^{(1)}(x|y) \to q^{(2)}(x) \to Q^{(2)}(x|y) \to \cdots$$

# Channels with Non-Causal Side-Information




$$C = \max_{p(x)} I(X;Y)$$

$$C = \max_{p(u|s)p(x|u,s)} I(U;Y) - I(U;S)$$

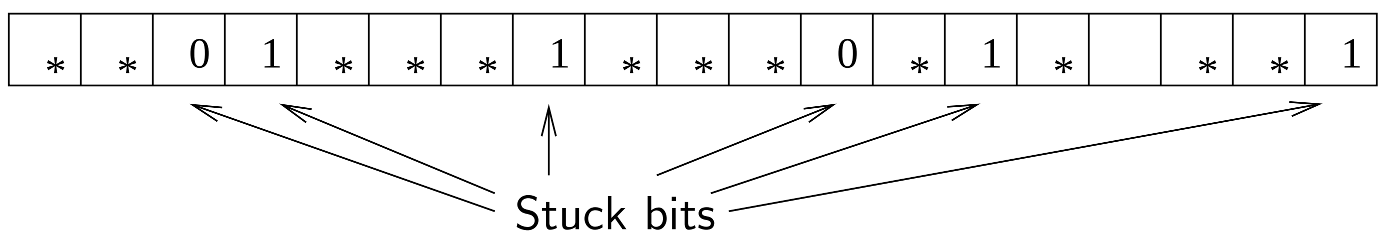
$$|\mathcal{U}| \leq |\mathcal{X}| + |\mathcal{S}|$$

(Gel'fand and Pinsker, 1979)

Question: How to compute capacity for channels with side-information?

# Example: Defective Memory

- Most bits act as a BSC, but some bits are stuck at either 0 or 1.
- Encoder knows which bits are bad, but not the decoder.



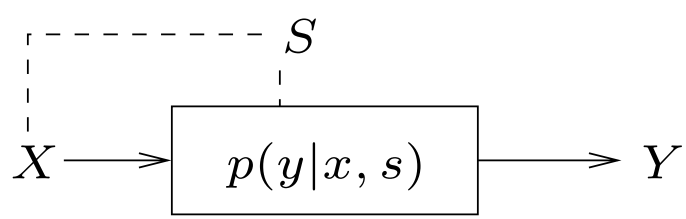
# Arimoto-Blahut with Side-Information

$$C = \max_{p(u|s)p(x|u,s)} \sum_{x,u,s,y} p(s)p(u|s)p(x|u,s)p(y|x,s) \log \frac{p(u|y)}{p(u|s)}$$

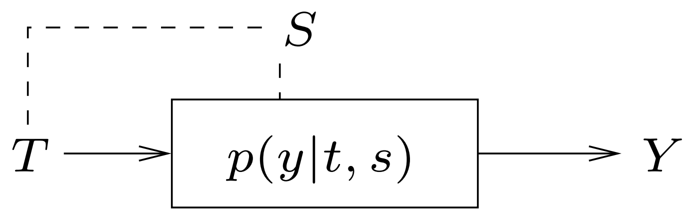
- Need to optimize over $p(u|s)$ and $p(x|u,s)$
- Mutual information is concave in $p(u|s)$ but convex in $p(x|u,s)$, thus difficult to optimize directly.

## Shannon strategies

- Because $C$ is convex in $p(x|u,s)$, the optimal $p(x|u,s)$ is a deterministic function
- So: $U = T$, with $T \in \mathcal{T}$, $\mathcal{T} = \{t : \mathcal{S} \to \mathcal{X}\}$.
- Now, $x = t(s)$ and $p(y|x,s) \to p(y|t,s)$
- Expansion of alphabet: $|\mathcal{T}| = |\mathcal{X}|^{|\mathcal{S}|}$



# Shannon strategies




Causal

$$C = \max_{p(t)} I(T;Y)$$

Non-causal

$$C = \max_{p(t|s)} I(T;Y) - I(T;S)$$

# Arimoto-Blahut with side-information

$$C = \max_{q(t|s),Q(t|y)} \sum_{t,s,y} p(s)q(t|s)p(y|t,s) \log \frac{Q(t|y)}{q(t|s)}$$

Fix $q(t|s)$:

$$Q^*(t|y) = \frac{\sum_s p(s)q(t|s)p(y|t,s)}{\sum_{s,t'} p(s)q(t'|s)p(y|t',s)}$$

Fix $Q(t|y)$:

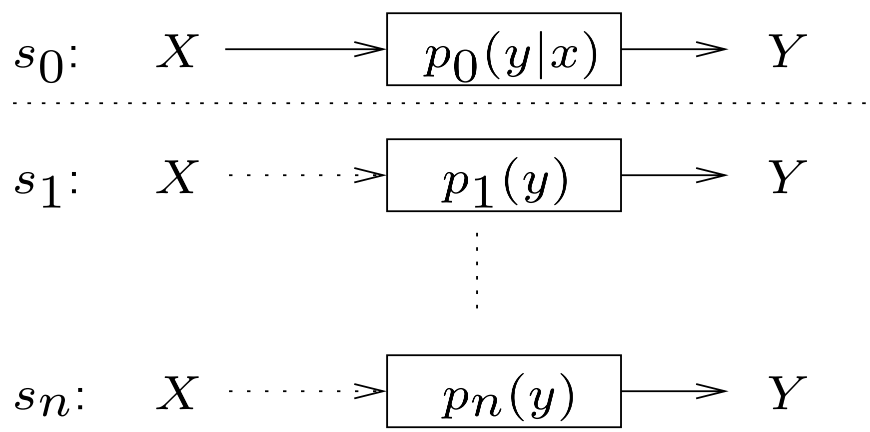
$$q^*(t|s) = \frac{\prod_y Q(t|y)^{p(y|t,s)}}{\sum_{t'} \prod_y Q(t'|y)^{p(y|t',s)}}$$

The algorithm:

$$q^{(1)}(t|s) \to Q^{(1)}(t|y) \to q^{(2)}(t|s) \to Q^{(2)}(t|y) \to \cdots$$

# Special case: “Stuck-at” channels

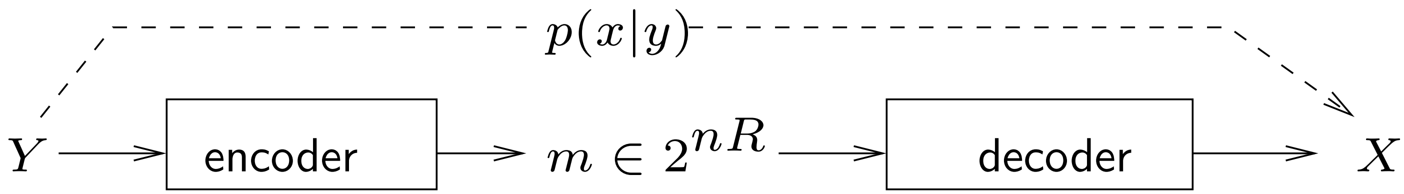
The output depends on the input only in one state $s_0$:




- Optimal $U$ is $X$.
- Reason: We can let $t(s) = t(s_0)$ for all $t$ and $s$.

# Source Coding

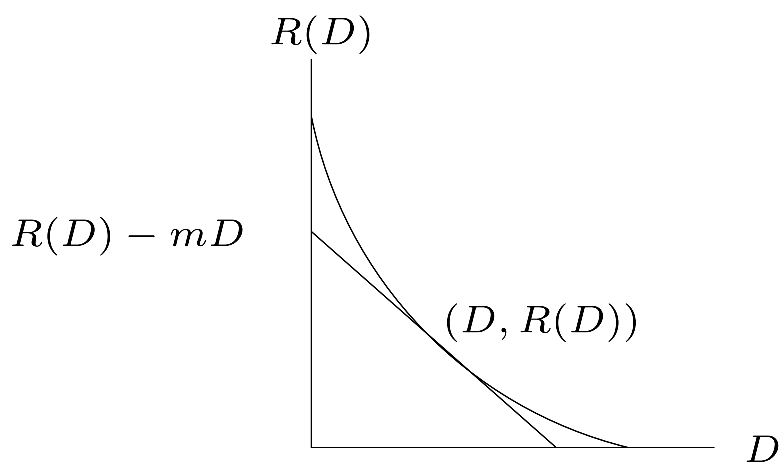
$$R(D) = \min_{p(x|y),E[d(X,Y)\leq D]} I(X;Y)$$



# Evaluating the Rate Distortion Function for Source Coding

It is easier to evaluate $R(D)$ at a given slope $m$.

$$R(D) - mD = \min_{p(x|y)} I(X;Y) - mE[d(X,Y)]$$



# Arimoto-Blahut for Source Coding

$$R(D) - mD = \min_{p(x|y)} I(X;Y) - mE[d(X,Y)]$$

Regard $q(x)$ to be independent of $Q(x|y)$ and perform alternating minimization:

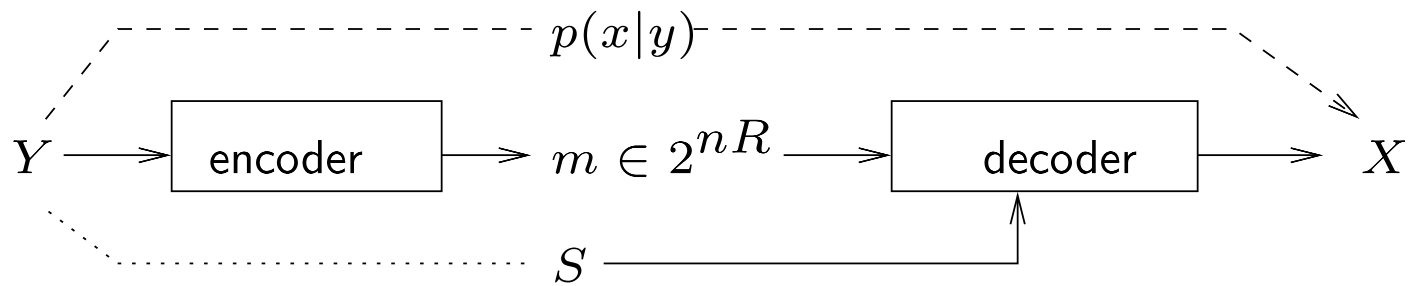
$$q^*(x) = \sum_y p(y)Q(x|y)$$

$$Q^*(x|y) = \frac{q(x)e^{md(x,y)}}{\sum_{x'} q(x')e^{md(x',y)}}$$

The algorithm: $Q^{(1)}(x|y) \to q^{(1)}(x) \to Q^{(2)}(x|y) \to q^{(2)}(x) \to \cdots$

# Source coding with side-information

The decoder has access to side-information that is correlated with the source.




$$R(D) = \min_{p(u|y), x=f(u,s), E[d(X,Y) \leq D]} I(U;Y) - I(U;S) = \min I(U;Y|S)$$

$$|\mathcal{U}| \leq |\mathcal{Y}| + 1$$

(Wyner and Ziv, 1976)

## Arimoto-Blahut for source coding with side-information

$$R(D) = \min_{p(u|y), x=f(u,s), E[d(X,Y) \leq D]} I(U;Y) - I(U;S)$$

- Difficulty: To find the optimal $f(u,s)$
- The trick: Use Shannon strategies: $x = t(s)$, where $t \in \mathcal{T} = \{t : \mathcal{S} \to \mathcal{X}\}$.
- Again, $|\mathcal{T}| = |\mathcal{X}|^{|\mathcal{S}|}$

$$R(D) = \min_{p(t|y), E[d(T(S),Y) \leq D]} I(T;Y) - I(T;S)$$

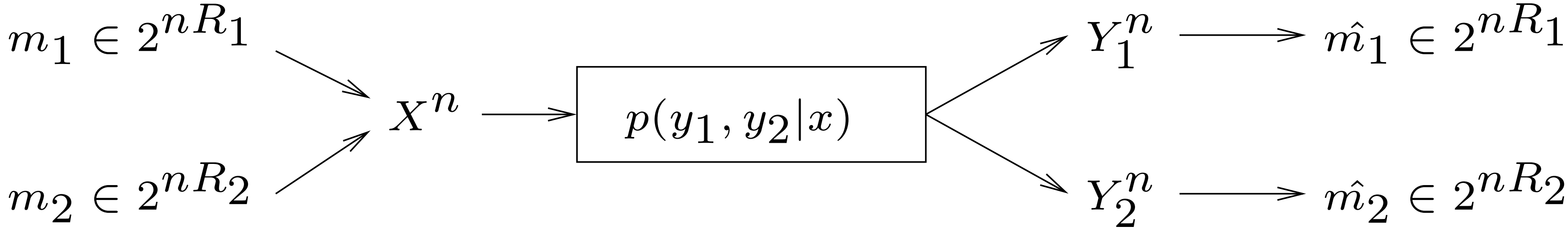
## Arimoto-Blahut for source coding with side-information

$$R(D) - mD = \min_{q(t|s),Q(t|y)} \sum_{x,y,s,t} p(y,s)Q(t|y)\log\frac{Q(t|y)}{q(t|s)} - mp(y,s)Q(t|y)d(t(s),y)$$

- Treat $q(t|s)$ as independent from $Q(t|y)$ and alternately optimize them.
- Result: $q^*(t|s)$ is the marginal given $Q(t|y)$, $Q^*(t|y)$ is a complex expression.
- Related work: Computation of $R(D)$ via geometric programming (Chiang'02).

# Broadcast Channels

One transmitter sends messages to several receivers at once:




Achievable rate region (Marton, 1979):

$$
\begin{aligned}
R_1 &\leq I(U_1; Y_1) \\
R_2 &\leq I(U_2; Y_2) \\
R_1 + R_2 &\leq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2)
\end{aligned}
$$

Note: there are no known cardinality bounds on $U_1$ and $U_2$.

# Arimoto-Blahut for Broadcast Channels

Write sum rate $R_1 + R_2$ as

$$\max_{q, q', Q_1, Q_2} \sum_{u_1, u_2, x, y_1, y_2} q(u_1, u_2) q'(x|u_1, u_2) p(y_1, y_2|x) \log \frac{Q_1(u_1|y_1) Q_2(u_2|y_2)}{q(u_1, u_2)}$$

- The expression is convex in $q'(x|u_1, u_2) \Rightarrow q'$ is a deterministic function.
- The expression is concave in $q$, $Q_1$ and $Q_2$.
- The trick: Perform alternate optimization over $q$, $Q_1$ and $Q_2$ for all possible functions $q'$.

## Arimoto-Blahut for Broadcast Channels

The optimal functions:

$$q^*(u_1,u_2) = \frac{\prod_{x,y_1,y_2}(Q_1(u_1|y_1)Q_2(u_2|y_2))^{q'(x|u_1,u_2)p(y_1,y_2|x)}}{\sum_{u_1,u_2}\prod_{x,y_1,y_2}(Q_1(u_1|y_1)Q_2(u_2|y_2))^{q'(x|u_1,u_2)p(y_1,y_2|x)}}$$

$$\begin{aligned}
Q_1^*(u_1|y_1) &= \frac{\sum_{u_2,x,y_2} q(u_1,u_2)q'(x|u_1,u_2)p(y_1,y_2|x)}{\sum_{u_2,x,y_2,u_1} q(u_1,u_2)q'(x|u_1,u_2)p(y_1,y_2|x)} \\
Q_2^*(u_2|y_2) &= \frac{\sum_{u_1,x,y_1} q(u_1,u_2)q'(x|u_1,u_2)p(y_1,y_2|x)}{\sum_{u_1,x,y_1,u_2} q(u_1,u_2)q'(x|u_1,u_2)p(y_1,y_2|x)}
\end{aligned}$$

# Conclusions

- Shannon strategies can simplify computations involving side-information.
- The Arimoto-Blahut algorithms can be generalized to channels with side-information and source coding with side-information.
- A similar strategy can be used to compute the achievable sum rate for discrete memoryless broadcast channels.