

نظریه اطلاعات کلاسیک - بخش دوم

وحید کریمی پور - دانشکده فیزیک - دانشگاه صنعتی شریف

۹ آذر ۱۳۹۷

۱ مقدمه

درفصل های قبل دیدیم که می توان پیام های یک منبع را که از کلمات $X = \{x_1, x_2, \dots, x_{2^n}\}$ تشکیل شده است را می توان با کدگذاری مناسب یعنی کدگذاری بلوکی چنان مخابره کرد که بجای n بیت به ازای هر کلمه، $nH(X) \leq n$ بیت مخابره کنیم. این نتایج البته در حد n های بزرگ دقیق هستند. تمام این حرف ها در صورتی بود که از نوفه داخل کانال صرف نظر کنیم و فرض کنیم که پیام بدون هیچ نوع تغییری به مقصد می رسد. مطالب بالا محتوی قضیه شانون را درباره کدگذاری بدون نوفه تشکیل می دهند. در این درس می خواهیم درباره کدگذاری در حضور نوفه بحث کنیم و نهایتا قضیه شانون درباره کدگذاری در حضور نوفه را بیان کنیم. در این درس هم چنین با مفهوم مهم ظرفیت کانال آشنا خواهیم شد و در حالت های خیلی ساده ای مقدار آن را محاسبه می کنیم.

۲ موازنه بین پایین آوردن خطا و بالا بردن نرخ انتقال اطلاعات

از این به بعد فرض می کنیم که کلمات منبع خود را به طریق بهینه ای کدگذاری کرده ایم و تنها می خواهیم کد کلمه هارا که رشته هایی از 0 و 1 هستند مخابره کنیم. فرض می کنیم که کانالی که پیام های خود را از طریق آن ارسال می کنیم تحت تاثیر نوفه قرار می گیرد و هر علامت 0 یا

1 که ارسال می کنیم با احتمال q تبدیل به 1 یا 0 شده و با احتمال $1 - q$ دست نخورده باقی می ماند. می خواهیم کاری کنیم که گیرنده پیام ها همچنان بتواند پیام های صحیح را از پیام های دریافت شده استخراج کند.

۱.۲ یک تذکره مهم

آنچه که در مثال بالا گفته شد یک نوع کد موسوم به کد تکرار سه تایی برای تشخیص و تصحیح خطاست و با کدگذاری کلمات فرق می کند. از این به بعد در این درس هر وقت که ما از کلمه کدگذاری استفاده می کنیم منظورمان این نوع کدگذاری است. به عبارت دیگر فرض می کنیم که کلمات منبع X که می توانند حروف یک زبان مثل فارسی یا انگلیسی یا علائم ریاضی باشند به نحو بهینه ای تبدیل به رشته های 0 و 1 شده اند و هدف ما مخابره این رشته ها از درون کانال است. در انتهای کانال رشته های دریافت شده نخست برای تعیین و تصحیح خطاها بررسی شده و پس از اصلاح به کلمات زبان مورد نظر بازگشایی می شوند.

راهی که برای تعیین و تصحیح خطا باید پیش گرفت آن است عنصر تکرار را به نوعی وارد پیام های خود کنیم. به عنوان مثال می توانیم بجای مخابره یک 0 سه تا 0 و بجای یک 1 سه تا 1 مخابره کنیم و از گیرنده بخواهیم که با استفاده از قانون اکثریت تصمیم بگیرد که یک رشته سه تایی که دریافت کرده است در واقع چه رشته ای بوده است (جدول 2.1).

در واقع احتمال خطا که قبلاً برابر بود با q اینک کمتر شده است، زیرا احتمال وقوع خطا اینک برابر است با احتمال وقوع دو برگردان و یا سه برگردان در کد سه تایی که اولی برابر است با $3q^2(1 - q)$ و دومی برابر است با q^3 . در نتیجه احتمال وقوع خطا برابر خواهد بود با $q^3 + 3q^2(1 - q)$ که برای q های کوچک از مرتبه q^2 است. البته بهایی برای این کاهش خطا پرداخت کرده ایم و آن این است که نرخ مخابره اطلاعات را پایین آورده ایم و از سه بیت برای مخابره یک بیت استفاده کرده ایم. در این حالت نرخ مخابره اطلاعات یعنی R برابر است با $\frac{1}{3}$. در حالت کلی اگر از n بیت برای مخابره 2^k پیام استفاده کنیم (که در نبود نوبه می توانست برای مخابره 2^n پیام مورد استفاده قرار گیرد) می گوئیم که نرخ مخابره اطلاعات برابر است

$$R := \frac{k}{n}. \quad (1)$$

مسلم است که در این نوع کدگذاری می توان باز هم احتمال خطا را کاهش داد و در حد ایده آل آن را به صفر نزدیک کرد ولی در این صورت نرخ R نیز به سمت صفر میل خواهد کرد.

■ تمرین: فرض کنید که کد تکرار n تایی را به کار می بریم که در آن n یک عدد فرد است. نرخ مخابره اطلاعات برابر است با $R = \frac{1}{n}$.

...	0
۱ ۱ ۱	1

جدول ۱: کد تکرار سه تایی

در این کد احتمال خطا را بر حسب n حساب کنید. سپس منحنی نرخ بر حسب احتمال خطا را رسم کنید.

■ تمرین: در کد هامینگ که با مشخصات $[7, 4, 3]$ تعیین می شود نرخ مخابره اطلاعات و هم چنین احتمال خطا را تعیین کنید.

آنچه که مهم است این است که ما نمی خواهیم اطلاعات را با احتمال خطای کم مخابره یا در جایی ذخیره کنیم. هدف ما آن است که این کار را با نرخ خطای صفر انجام دهیم یا به عبارت دقیق تر مطلوب ما آن است که این نرخ خطا را هر چقدر که می خواهیم بتوانیم به صفر نزدیک کنیم. حال سوالی که با آن مواجه هستیم آن است که آیا اصولاً می توان از یک کانال نوفه دار با نرخ محدود و غیر صفر اطلاعاتی را عبور داد به نحوی که احتمال وقوع خطا لا اقل در حد مجانبی (برای رشته های بزرگ) به سمت صفر میل کند؟ آیا یک نوع کدگذاری وجود دارد که این امکان را برای ما فراهم آورد؟ آیا نرخ بیشینه ای وجود دارد که هر نوع مخابره اطلاعات با بیش از آن نرخ بدون خطا غیر ممکن باشد؟ این ها سوالاتی است که در قضیه دوم شانون پاسخ داده شده است.

۳ قضیه دوم شانون و ظرفیت کانال های کلاسیک

شکل (۱) یک کانال کلاسیک را به طور کامل نشان می دهد. این کانال دارای پنج قسمت مهم است:

■ یک - کد گذاری منبع^۱.

در این جا خروجی یک منبع که به صورت رشته ای از بیت ها ست نخست فشرده می شود. این همان کاری است که موضوع قضیه اول شانون است و در درس قبل به آن پرداخته ایم. مطابق با آنچه که در درس پیشین گفتیم این فشرده سازی مبتنی بر کد کردن تنها رشته های نمونه است. با بزرگ گرفتن طول رشته ها همواره می توانیم کاری کنیم که خطایی که در اثر کد نکردن رشته های غیر نمونه مرتکب می شویم به سمت صفر میل کند. هم چنین در آن درس یاد گرفتیم که رشته های نمونه تقریباً هم احتمال هستند. یعنی تابع احتمال روی زیر مجموعه رشته های نمونه تقریباً تخت است. خروجی این قسمت را با رشته های به طول k و به صورت $\mathbf{m} = m_1 m_2 \dots m_k$ نشان می دهیم. الفبای m_i معمولاً همان بیت های 0 و 1 است.

■ دو - کد گذاری کانال

^۲. در این قسمت پیام های m طوری کد می شوند که خطای ایجاد شده در آنها را بتوانیم آشکار و تصحیح کنیم. این کار را به صورت زیر نشان می دهیم:

$$\mathbf{m}^k \longrightarrow \mathbf{x}^n \quad (۲)$$

یا

$$m_1 m_2 \dots m_k \longrightarrow x_1 x_2 \dots x_n. \quad (۳)$$

در این مرحله تعداد 2^k رشته را که همه هم احتمال بودند (نگاه کنید به شکل (۲))، به فضایی می نگاریم که تعداد 2^n رشته دارد. طبیعی است که $n > k$ است، به این معنا که کد رشته های $\mathbf{m}$ را به فضایی بزرگتر نگاشته ایم چنانکه فاصله کدرشته های جدید از هم به اندازه ای زیاد باشد که در اثر خطا با هم اشتباه نشوند و بتوانیم خطاهای احتمالی ایجاد شده را آشکار کرده و تصحیح کنیم. اباید دقت کنیم که این کد گذاری برای فشرده کردن اطلاعات منبع اولیه نیست بلکه برای کاستن خطای کانال است. انتروپی رشته های تولید شده را با $H(X)$ نشان می دهیم. هرگاه نحوه کدگذاری خود را عوض کنیم $H(X)$ نیز عوض خواهد شد.

■ سه - کانال کلاسیک

^۳. کانال کلاسیک اگر بدون خطا می بود می بایست رشته ورودی $\mathbf{x}$ را به صورت رشته خروجی $\mathbf{x}$ تحویل می داد، اما از آنجا که این کانال دارای

^۱Coding Source

^۲Coding Channel

^۳Channel Classical

خطاست ممکن است رشته ورودی x را به صورت رشته خروجی y تحویل دهد. هر دو رشته x و y در یک فضا قرار دارند.

■ چهار - کد گشایی کانال،

^۴ در این قسمت که مهم ترین قسمت برای بحث ما در این درس است، رشته های خروجی y خوانده می شوند. این رشته ها همان رشته های ورودی نیستند چرا که در آن ها خطا رخ داده است. می بایست خطای آنها آشکار شده و به رشته های اولیه x نگاشته شوند.

■ پنج - کد گشایی منبع

کد رشته x به رشته اولیه یعنی m نگاشته می شود.

■ تعریف ظرفیت کانال:

ظرفیت کانال فوق به صورت زیر تعریف می شود:

$$C := \lim_{n \rightarrow \infty} \max_X \frac{I(X : Y)}{n} \quad (۴)$$

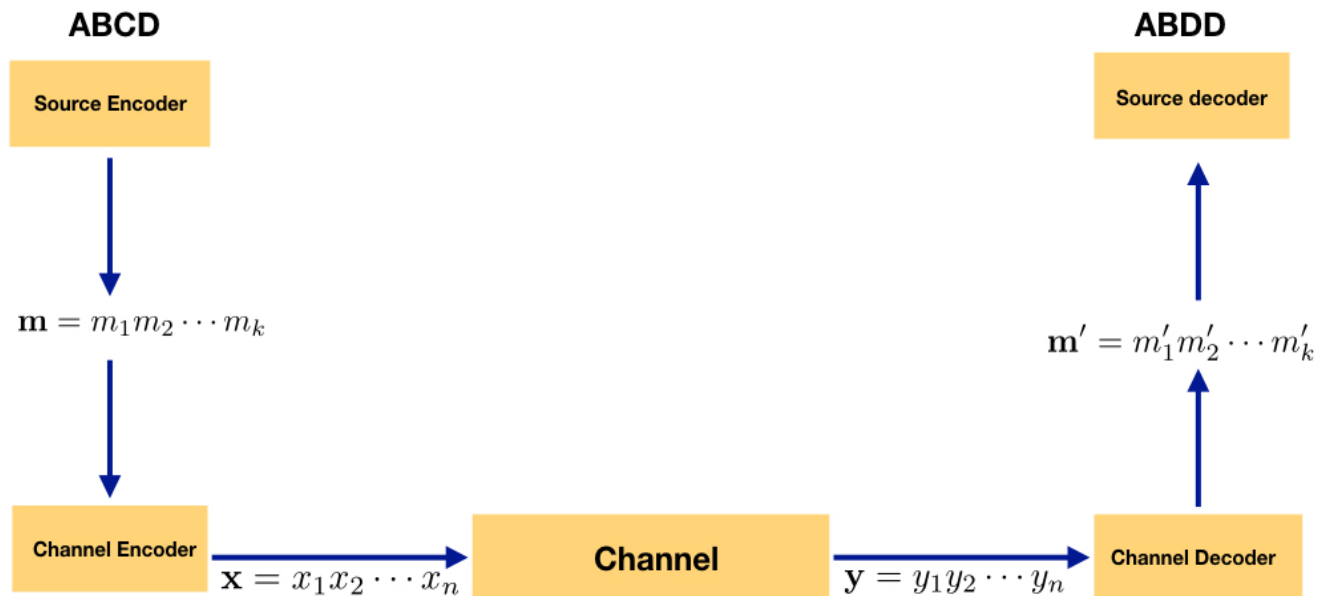
که در آن منظور از $\max_X$ این است که روی تمام انواع کدگذاری های X مقدار بیشینه را حساب کنیم.

■ قضیه دوم شانون: همواره می توانیم با افزایش n نرخ مخابره از یک کانال کلاسیک یعنی $R \equiv \frac{k}{n}$ را تا حد امکان به ظرفیت کانال نزدیک

کرد و در عین حال خطا را به هر اندازه که می خواهیم کوچک کنیم.

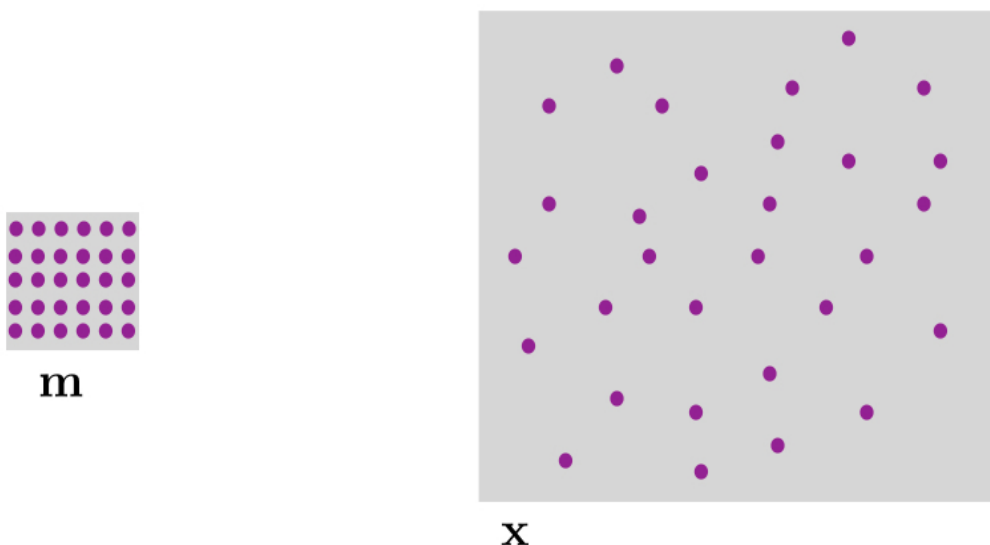
برای اینکه معنای این قضیه و اثبات آن را بفهمیم به ترتیب زیر عمل می کنیم. یک پیام m را در نظر بگیرید. این پیام به یک رشته n بیتی x کد می شود. مجموعه X یک مجموعه خیلی بزرگ است که دارای 2^n رشته است اما در کدگذاری ای که ما انجام می دهیم تنها 2^k تا از این رشته ها انتخاب می شود و این مجموعه رشته ها دارای آنتروپی $H(X)$ است. وقتی که رشته ای مثل x را به درون کانال می فرستیم ممکن است دچار خطا شده و تبدیل به یک رشته نزدیک مثل y شود. در مقصد وقتی که رشته y را دریافت می کنیم، می بایست آن را به رشته اولیه x

Decoding Channel^۵



شکل ۱: قسمت های مختلف یک کانال کلاسیک.

برگردانیم و خطایی که در آن رخ داده است را تشخیص و تصحیح کنیم. این تصحیح به این صورت انجام می شود که یک ناحیه حول رشته $\mathbf{y}$ در نظر می گیریم که محتمل ترین $\mathbf{x}$ هایی که می توانسته اند در اثر خطا به $\mathbf{y}$ تبدیل شوند، در آن قرار دارد. در شکل ۳ این ناحیه با رنگ آبی کم رنگ مشخص شده است. در این ناحیه به دنبال یک رشته معتبر (یعنی یک رشته $\mathbf{x}$ که در مجموعه کد رشته های ما قرار دارد می گردیم. آن رشته همان رشته ای است که در اثر خطا تبدیل به $\mathbf{y}$ شده است. به این ترتیب خطا را اصلاح می کنیم. در واقع اگر همیشه این طور بود می توانستیم بگوییم که هیچ خطایی رخ نمی دهد زیرا هر خطایی که رخ داده است به این روش اصلاح می کنیم. اما گاهی اتفاق می افتد که ناحیه آبی رنگ دارای دو کد رشته یا بیشتر است. در این شرایط خطا رخ می دهد زیرا نمی دانیم رشته دریافت شده را به کدام یک از کد رشته ها برگردانیم. می خواهیم تخمینی از این خطا بدست آوریم. احتمال این که یک کد رشته به طور اتفاقی درون این ناحیه بیفتد چقدر است؟ این احتمال را با P



شکل ۲: مجموعه پیام های m می بایست به مجموعه پیام های x چنان نگاشته شوند که هنگام عبور از کانال اگر خطایی در آنها ایجاد شود، بتوان این خطاها را تصحیح کرد.

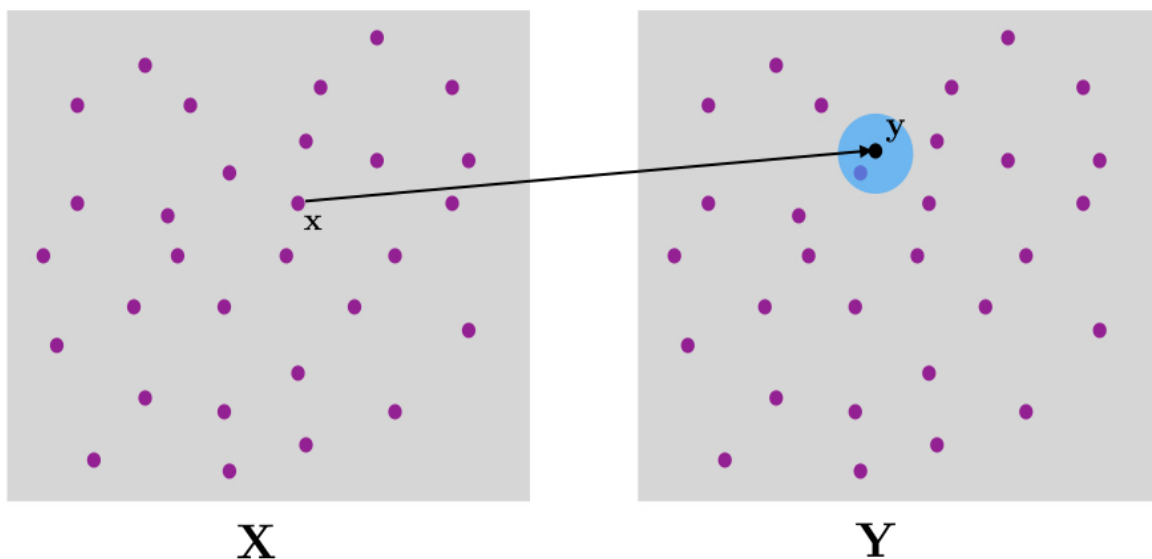
نشان می دهیم. اگر حجم ناحیه آبی رنگ (یا به عبارت دیگر تعداد نقاط ناحیه آبی رنگ را با v) نشان دهیم این احتمال برابر است با:

$$P = 2^k \times \frac{v}{2^{nH(X)}} \quad (5)$$

این احتمال در واقع برابر است با تعداد کل کد رشته های پیام یعنی x ها که تعداد آنها 2^k تاست ضرب در احتمال اینکه یک کد رشته (یک نقطه در شکل ()) داخل ناحیه آبی رنگ بیفتد.

اما v به طور متوسط چقدر است؟ اگر یک کد رشته مثل $y = y_1 y_2 \dots y_n$ در نظر بگیریم این کد رشته می توانسته به خطا، هنگام عبور کد رشته $x = x_1 x_2 \dots x_n$ از کانال رخ داده باشد. احتمال اینکه y دریافت شده از یک x حاصل شده باشد، برابر است با:

$$P(\mathbf{x}|\mathbf{y}) = P(x_1|y_1)P(x_2|y_2)P(x_3|y_3) \dots P(x_n|y_n).$$

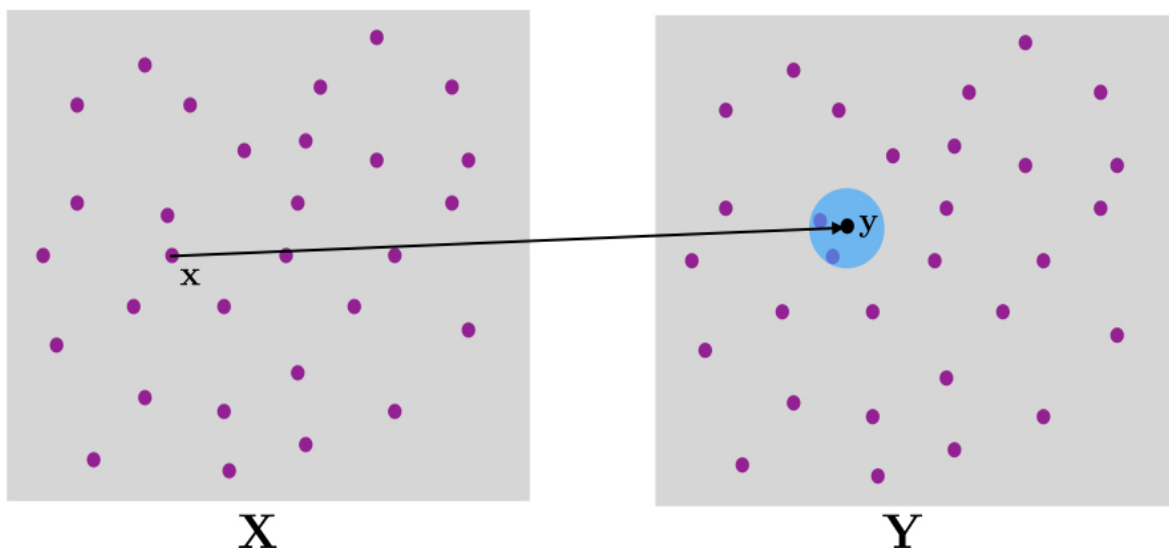


شکل ۳: رشته x وقتی از کانال عبور می کند در اثر خطا تبدیل به رشته y می شود. اگر مجموعه نقاط نزدیک y (دایره آبی رنگ در شکل سمت راست) دارای تنها یک کدرشته معتبر مثل x باشد، رشته دریافتی y را به همان رشته تصحیح می کنیم.

تعداد چنین رشته هایی در حد n های بزرگ چندتااست؟ برای سادگی فرض کنید که x_i ها و y_i ها صفر و یک هستند و احتمال خطا نیز به صورت زیر است:

$$P(0 \rightarrow 1) = P(1 \rightarrow 0) = q$$

در این صورت در یک رشته n تایی معمولاً nq تا خطا رخ می دهد. این خطاها می توانند در جاهای مختلفی از رشته رخ دهند. تعداد متوسط رشته هایی که با این خطاها می توانند به یک رشته تبدیل شوند برابر است با $\binom{n}{nq} \approx 2^{nH(q)}$. حال فرض کنید که الفبای x ها و y ها صفر و یک نیست بلکه 0, 1, 2 است. فرض کنید که رشته ای مثل $y = 1100120120120....120001120220112$ دریافت کرده باشیم، که دارای n_0 تا 0، n_1 تا 1 و n_2 تا 2 است. این رشته از چه رشته هایی می توانسته تولید شده باشد؟ برای پاسخ به این سوال



شکل ۴: خطا وقتی رخ می دهد که دایره آبی رنگ دارای بیش از یک رشته معتبر باشد. رشته اضافی می توانسته از هر جایی از نقاط x آمده باشد.

فقط به جایگاه ۰ ها توجه می کنیم:

$$y_0 = ..00..0..0..0.....000...0..0...$$

تعداد صفرها را در این رشته بخصوص با $n_0(y)$ نشان می دهیم.

این رشته می توانسته از هر رشته دیگری که رقم های نقطه چین اش (یعنی جایگاه ۱ ها و ۲ ها) تغییر نکرده ولی در این جایگاه های خاص

ارقام ۰ ، ۱ و ۲ در اثر خطا به رقم ۰ تبدیل شده اند بوجود آمده باشد. به عنوان مثال یکی از این رشته این است:

$$x = ..01..0..1..0.....000...2..0...$$

در چنین رشته هایی به طور متوسط تعداد 0 ها، تعداد 1 ها و تعداد 2 ها با تقریب خیلی خوب برابر است با:

$$r_0 \approx n_0(\mathbf{y})P(x=0|y=0) \quad r_1 \approx n_0(\mathbf{y})P(x=1|y=0) \quad r_2 \approx n_0(\mathbf{y})P(x=2|y=0). \quad (۶)$$

تعداد چنین رشته هایی با تقریب خیلی خوب برابر است با:

$$Q_0(\mathbf{y}) \approx \binom{n_0(\mathbf{y})}{r_0, r_1, r_2} = \frac{n_0(\mathbf{y})!}{r_0!r_1!r_2!} \quad (۷)$$

و یا

$$Q_0(\mathbf{y}) \approx \frac{n_0(\mathbf{y})!}{(n_0(\mathbf{y})P(x=0|y=0))!(n_0(\mathbf{y})P(x=1|y=0))!(n_0(\mathbf{y})P(x=2|y=0))!} \quad (۸)$$

و یا با استفاده از تقریب استرلینگ

$$Q_0(\mathbf{y}) \approx 2^{n_0(\mathbf{y})H(X|y=0)}. \quad (۹)$$

این کار را می شد برای زیر رشته ای که فقط شامل یک ها یا زیر رشته ای که فقط شامل ۲ هاست نیز انجام داد و بدست آورد:

$$Q_1(\mathbf{y}) \approx 2^{n_1(\mathbf{y})H(X|y=1)} \quad Q_2(\mathbf{y}) \approx 2^{n_2(\mathbf{y})H(X|y=2)} \quad (۱۰)$$

اما تعداد کل رشته هایی که می توانسته اند به این رشته تبدیل شوند برابر است با:

$$\begin{aligned} Q(\mathbf{y}) &= Q_0(\mathbf{y})Q_1(\mathbf{y})Q_2(\mathbf{y}) \\ &= 2^{\sum_y n(y)H(X|y)} = 2^{\sum_y nP(y)H(X|y)} = 2^{nH(X|Y)}. \end{aligned} \quad (۱۱)$$

بنابراین اگر ما حجم v را کمی بیش از این در نظر بگیریم یعنی قرار دهیم $v = 2^{nH(X|Y)+n\delta}$ احتمال خطا برابر است با:

$$P_{error} = 2^k \times \frac{2^{nH(X|Y)+n\delta}}{2^{nH(X)}} \quad (۱۲)$$

و با توجه به این که $R = \frac{k}{n}$ و $H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y) = I(X; Y)$ بدست می آوریم:

$$P_{error} = 2^{n(R-I(X;Y)-\delta)}. \quad (۱۳)$$

معنای این حرف این است که با افزایش طول رشته ها می توانیم هرچقدر که می خواهیم نرخ خطا را به سمت صفر میل دهیم مشروط بر آنکه نرخ R از $I(X, Y)$ کمتر باشد. بنابراین ظرفیت کانال مطابق با رابطه زیر تعریف می شود:

$$C := \sup_X I(X, Y). \quad (۱۴)$$