

منع لفظیات : همان طور که اشاره شد، منع لفظیات شامل یاها می‌است که می‌فراهم
از مبدأ - به سمت مقصد ارسال کنیم. غرض منج اظہارات، می‌تواند آنرا گرد یا دیگر باشد.
اگر غرض منج آنرا گذر باشد، با کمک روشهای دیگر کردن پیام، آن را به اظہارات دیگر
تبدیل می‌کنیم. مثلاً می‌توان منج اظہارات می‌تواند پیامهای صوتی، تصویری و ...
را باشد.

هر منج اظہارات، شامل مصروفیات است که برای کار کردن با آن منج، لازم است اظہارات
راجع به این مصروفیات داشته باشیم.

* نرخ اطلاعات منبع : نشان دهنده ی تعداد سیمبل های ترلید شده توسط منبع در هر ثانیه است .

$$r \text{ نرخ اطلاعات منبع} \quad \text{Symbol/Second} \quad \left(\frac{\text{Sym}}{s} \right)$$

* الفبای منبع اطلاعات : شامل مجموعه ی سیمبل های است که یک منبع اطلاعات به عنوان ورودی ترلید می کند . به عنوان مثال سی تران الفبای منبع یا مجموعه ی سیمبل های ترلیدی یک منبع را با مجموعه ای به شکل $\{a_1, a_2, \dots, a_n\}$ نمایش داد .

- در هر نقطه ای از این پیامها به صورت صفای ، در ورودی منبع حاضر می شوند .

۴ خصوصیات آماری منبع اطلاعات

باید به یاد داشت $\{a_1, a_2, \dots, a_n\}$ از الیاس منبع، به صورت تصادفی در خروجی ظاهر می شوند، پس دیگر از خصوصیات منبع، احتمال وقوع هر یک از این پیامها در خروجی منبع است. حتی احتمال اینکه منبع در هر لحظه، هر یک از این پیامها را با چه احتمالی تولید می کند. این خصوصیات آماری را با مجموعه ای احتمالات $\{P_1, P_2, \dots, P_n\}$ نمایش می دهیم که در آن P_i احتمال این است که پیام a_i در خروجی منبع ظاهر شود.

(در این درس، منبع های اطلاعات را بدون حافظه در نظر می گیریم)

* اطلاعات منبع
در این مورد لازم است که مفهم اطلاعات ب منبع در تخری اندازه گیری آن نیز تعریف شود.

در این مورد در درس تخری اطلاعات به صورت کامل ، بحث خلاصه کرد.

به صورت کلی تخری اطلاعات به سؤالات زیر جواب می دهد ،

- اصرار اطلاعات چیست و چگونه می توان اطلاعات ب منبع ایمانیه کرد ؟ (آندوی)

- حد های میزان اطلاعات قابل انتقال در یک کانل چیست ؟ (ظرفیت کانال)

- حد های میزان اتصالاتی که ب منبع به دست می رسد چیست ؟ (۵۰۰ آنندوی منبع)

- حد های تخری سازی اطلاعات ب منبع چیست ؟ (کتاب های تخری سازی ، lossy , lossless)

در ادامه می فرمایم، مفهوم اطلاعات یک منبع را معرفی کنیم.

برای این منظور لازم است که ابتدا مفهوم اطلاعات را برای یک پیام، معرفی کنیم.

به طور صریح، برای یک پیام A هر چه احتمال وقوع پیام کمتر باشد، اطلاعات (پیام) آن بیشتر است. بنابراین اگر اطلاعات پیام A را $I(A)$ نمایش دهیم، خواهیم داشت،

اطلاعات تابع نزولی از احتمال است.
$$1) \quad P(A) < P(B) \Rightarrow I(A) > I(B)$$

اطلاعات یک پیام هرگز مثبت است
$$2) \quad \forall A; \quad I(A) > 0$$

* برای دریافت مستقل A و B ، اگر پیام C را به صورت $C = A \cap B$ تعریف کنیم، خواصیم داشت،

$$3) \quad I(C) = I(A \cap B) = I(A) + I(B) \quad , \quad A \perp B$$

* در رایانشان داده می شود که نتایجی که به دست می آید را برای داده های گذشته نگه می دارند. بنابراین برای پیام A با احتمال وقوع P_A ، اطلاعات پیام به صورت زیر تعریف می شود.

$$I(A) \triangleq \log \frac{1}{P_A} = -\log P_A$$

واحد کمیت اطلاعات براساس پایه گزیده شده در رابطه با اطلاعات مشخص می شود.

$$I(A) = \log_a \frac{1}{P_A} = -\log_a P_A$$

- اگر $a=2$ در نظر گرفته شود، واحد اطلاعات برابر با bit binary digit خواهد بود.

- اگر $a=e$ در نظر گرفته شود، واحد اطلاعات برابر با nat خواهد بود.

natural digit

* باید مقایسه نبندی سازه، سی تران واحد های اطلاعات را به یکدیگر تبدیل کرد، زیرا

$$\log_b P_A = \left(\log_b a \right) \log_a P_A$$

حال که مفهوم اطلاعات برای یک پیام (نمونه A ، به دست آوردیم، می خواهیم مفهوم
اطلاعات را برای یک منبع بیان کنیم.

می دانیم که یک منبع اطلاعات پیام های را از مجموعه $X = \{a_1, a_2, \dots, a_n\}$ با احتمالات
 $\{P_1, P_2, \dots, P_n\}$ تولید می کند. بنابراین برای به دست آوردن اطلاعات یک منبع
باید از اطلاعات پیام های آن منبع امید ریاضی بگیریم (یا میانگین گیری کنیم)

امید ریاضی اطلاعات پیام های منبع = اطلاعات منبع X
 $\Rightarrow$

اطلاعات یک منبع X را با $H(X)$ نمایش می دهیم و به آن آنترپی می گیریم.

$$\Rightarrow H(X) = E\{I(a)\} = E\left\{\log \frac{1}{P_a}\right\}$$

$$= \sum_{i=1}^n P_i \log \frac{1}{P_i} = - \sum_{i=1}^n P_i \log P_i$$

(برای بدینج
نکته ی بدون
حافظه)

* یادآوری : مفهوم امید ریاضی Expectation

برای متغیر تصادفی پیوسته ی X تابع چگالی احتمال $f_X(x)$ امید ریاضی به صورت زیر تعریف می شود

$$E\{g(x)\} = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad \equiv \quad \text{میانگین سادگی که } g(x) \text{ اختیار می کند}$$

برای متغیر تصادفی گسسته ی X تابع چگالی احتمال $P_X(x)$ خواص داشته باشد

$$E\{g(x)\} = \sum_x P_X(x) g(x)$$

$$H(x) \geq 0$$

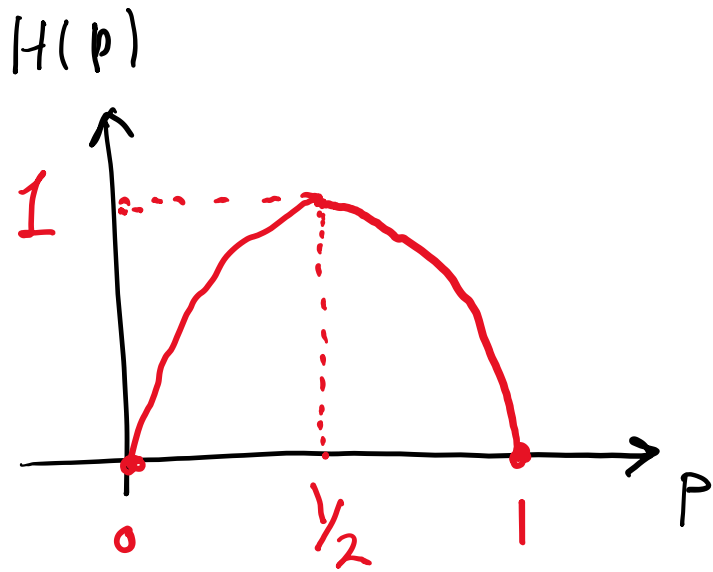
* برای یک منبع کسسه‌ی بدون حافظه، همواره داریم

مثال: برای یک منبع باینری با خصوصیات آماري زیر، آنتروپی را محاسبه کنید.

$$x = \begin{cases} 1 & P \\ 0 & 1-P \end{cases}$$

$$\Rightarrow H(x) = - \sum_{i=1}^n p_i \log p_i$$

$$\Rightarrow H(x) = \underset{\substack{\uparrow \\ \text{باینری}}}{-P \log P} - (1-P) \log (1-P) \equiv H(P)$$



• همان طوری که دیده می شود به ازای $p=0$ و $p=1$ ، آنتروپی (اطلاعات منبع) صفر است. زیرا در این حالت فرد می منبع حالت قطعی (یعنی) دارد و هیچ ابهامی نخواهد داشت. بنابراین اطلاعات منبع برابر صفر است.

• بیشترین مقدار اطلاعات منبع به ازای $p = \frac{1}{2}$ دیده می شود، زیرا در این حالت بیشترین ابهام را در فرد می منبع ضاعف می باشد و در نتیجه بیشترین آنتروپی دیده می شود.

قضیه: برای یک منبع احتمالات $X = \{a_1, a_2, \dots, a_n\}$ با احتمالات $\{P_1, P_2, \dots, P_n\}$ $H(X)$ زیاده‌ترین مقدار را دارد که توزیع احتمالاتی منبع می‌تواند داشته باشد، یعنی داشته باشیم

$$P_1 = P_2 = \dots = P_n = \frac{1}{n} = \frac{1}{|X|}$$

اندازه‌ی مجموعه‌ی X

رضایحتم داشته

$$H_{\max}(X) = \log |X|$$

$$H(X) = -\sum_{i=1}^n P_i \log P_i$$

$$H(x) \leq \log |x|$$

نتیجه: برای هر منبع X حدارده می توان نوشت

با مجموعه پیام های X

مثال: برای منبع X با خصوصیات آماری زیر آنتروپی را محاسبه کنید.

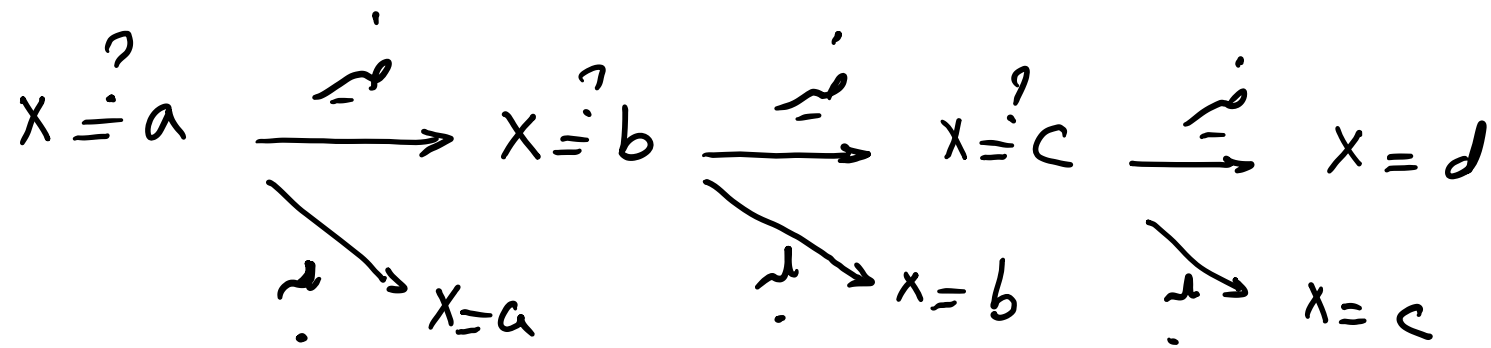
$$X = \begin{cases} a & \frac{1}{2} \\ b & \frac{1}{4} \\ c & \frac{1}{8} \\ d & \frac{1}{8} \end{cases}$$

$$\begin{aligned} \Rightarrow H(X) &= - \sum_{i=1}^4 P_i \log P_i \\ &= \frac{1}{2} \log^1 2 + \frac{1}{4} \log^2 4 + \frac{1}{8} \log^3 8 + \frac{1}{8} \log^3 8 \end{aligned}$$

$$H(X) = \frac{7}{4} \text{ bits}$$

درستوری اطلاعات نشان داده می شود که $H(x)$ برای یک منبع، برابر کمترین تعداد بیت لازم برای که کردن اطلاعات منبع است (که کد میزند)

اگر بخواهیم خودی منبع را با سوالات با پاسخ لمبر غیر، عدس بزنیم، $H(x)$ برابر میماند کمترین تعداد سوالات برای رسیدن به جواب است.

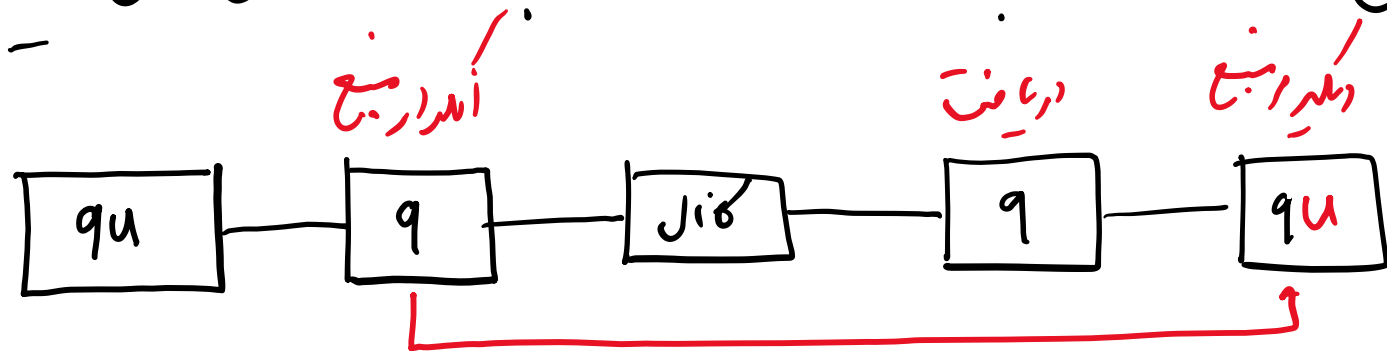


میماند تعداد سوالات برای رسیدن به جواب

$$= \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 3 = \frac{7}{4}$$

در ادامه به بحث گذاری منجی پردازم. در واقع بحث گذاری منجی، این است که برای یک منجی x ، یک ردش گذاری معرفی کنیم به طوری که بتوان انبهای منجی را با بهترین مایلین تعداد بیت گذاری کرد. به این ترتیب از پهنای باند سیستم، به بهترین شکل استفاده می شود. یکی از نکات مهم در طراحی گذاری منجی، این است که دیتا گیرنده، بتواند دیتای ارسالی را با خطای به عنوان دلخواه کم، بازیابی کرد.

به عنوان یک مثال ساده از گذاری منجی، می توان به حذف حرف u بعد از q در متن های انگلیسی اشاره کرد.



در تئوری اطلاعات بحث کنید منبع و فشار در سازی اطلاعات، به منظور استنادی به سراز
پسای باند مطرح می شود در رابطه با قضایای عالم بر این الگو بهم ها بررسی می شود. تلاش طراحان
بر این است که برآیند با پیچیدگی قابل قبول را قابلیت اطمینان بالا، در رتبه های نزدیک منبع اطلاعات
کنند تا بتوان از پسای باند به شکل سبزی استناد کرد.
در روش های که باید منبع را می توان به درستی عمل کرد باید منبع با جدول ثابت را کرد منبع با
جدول متغیر تقسیم بندی کرد. اگر جدول که برای تمام پیام های منبع یکسان باشد، که باید با جدول ثابت است در صورتی
که جدول که برای پیام های مختلف، برآیند متغیر مختلفی باشد، که باید با جدول متغیر کنیم.
این موضوع را باید سال سازه، خواجه نصیر

کتاب منبع اصلی است با انبساط $x = \{a_1, a_2, a_3\}$ راضی است $\{P_1, P_2, P_3\}$ از نظر

می‌گیریم. با فرض $P_1 = \frac{1}{2}$, $P_2 = \frac{1}{4}$, $P_3 = \frac{1}{4}$ می‌فراصیم ردش برای که گران یا نهای
منبع معروض کنیم.

الف) روش کربن-۱۴

خزیت : سارگی پاره سازی اندر در درید، رحصین عملیات

کمر دن سازی
در سطح بیت

$$x = \begin{cases} a_1 & \frac{1}{2} & 00 \\ a_2 & \frac{1}{4} & 10 \\ a_3 & \frac{1}{4} & 01 \end{cases}$$

نمونه ریاضی: $\underbrace{00}_{a_1} \underbrace{01}_{a_3} \underbrace{01}_{a_3} \underbrace{10}_{a_2} \underbrace{00}_{a_1}$

$$H(x) = \frac{1}{2} \log^2 + 2 \times \frac{1}{4} \log^4 = \frac{3}{2} \Rightarrow$$

اشغال روش که سبب باطل ثابت، این است که به حد کمترین میانگین طول کلمات که می رسد
 در رتبه (زیبایی) باشد به فرض اشاره نمی کند.
 dr_haghighi.info / Courses

$$\begin{array}{l} \text{میانگین طول کلمات که} \\ \text{در مثال که سبب} \\ \text{باطل ثابت} \end{array} = \frac{1}{2} \times 12 + \frac{1}{4} \times 2 \times 12 = 2 > \overset{3/2}{\sim} H(x)$$

اما از طریق ساده سازی روشی ساده است.

(ب) روش با جدول معبر

در این روش برای استفاده از بهزیایی بانه، به پیامهای با احتمال بیشتر، طول کد کمتری اختصاص داده می شود.
 به این ترتیب پیامی که بیشترین احتمال وقوع (ارسال ردی) را دارد با کمترین تعداد بیت کد ارسال
 می شود و به این ترتیب از بهزیایی بانه به شکل بهتری استفاده می شود.

پیام	احتمال	0	1
a_1	$\frac{1}{2}$	0	1
a_2	$\frac{1}{4}$	10	11
a_3	$\frac{1}{4}$		

که اینها را می بینیم

→ که

$$\frac{1}{2}(1) + 2 \times \frac{1}{4}(2) = \frac{3}{2} = H(x)$$

طریقه کد

این روش به هر کمترین معاینین طول هاست که می رسد و از بهزیایی بانه به بهترین شکل استفاده می کند.

در عرض پیچیدگی زیاد سازی انگل در دکلر آن بیشتر از رشد با طول ثابت است. همچنین
نگردن سازی در سطح بت (ملات که) پیچیده تر فزاینده بود.

$$\begin{array}{ccccccc} & a_2 & a_3 & a_2 & a_1 & & \\ & \sim & \sim & \sim & \sim & & \\ 00 & 0 & 1 & 1 & 1 & 0 & 0 \\ \sim & \sim & \sim & & & & \\ a_1 & a_1 & a_1 & & & & \end{array}$$

علامه بر این اثر خطای از زمان دکلرید رخ بدهد
امکان انشای خط در عملیات دکلرید در گزیده
و عدد دارد.

* در طراحی که حاس با طول شعری برای عددگیری از ابهام دکلرید (در نگردن سازی) در گزیده، ملات
که با به صورتی انتخاب می کنند که هیچ طبعی که پیچیده طبعی که در دکلرید نباشد.