BAB III

UKURAN INFORMASI

Dalam sistem komunikasi, informasi ditransmisikan oleh sebuah sumber informasi melalui channel untuk dikirimkan pada tempat tujuannya. Seorang penyiar radio menyampaikan berita melalui gelombang radio yang dipancarkan kepada para pendengarnya. Otak manusia sebagai sumber informasi, menyampaikan sinyal-sinyal perintah kepada sistem gerak tubuh melalui syaraf-syaraf yang tersebar di dalam tubuh.

Setiap sumber informasi menghasilkan sederet pesan -pesan, yang dapat berupa huruf-huruf abjad, not musik, impuls elektrik atau angka-angka. Deretan total dari pesan -pesan yang berbeda, yang dapat ditransmisikan oleh sumber disebut *alphabet* dari sumber tersebut.

Contoh 3.1 :

Seorang menteri mengirim surat dalam bahasa Inggris, maka kandungan alphabetnya terdiri dari huruf a hingga z dan sebuah spasi.

Contoh 3.2 :

Instruktur komputer memasukkan bilangan binary 3 digit ke dalam sebuah komputer, alphabetnya memiliki 2^3 elemen, yaitu

000	010	100	110
001	011	101	111

Beragam karakter dalam suatu alphabet keluar dari sumber dengan membawa probabilitas yang diberikan, sehingga setiap sumber dihubungkan dengan suatu ruang probabilitas.

Contoh 3.3 :

Sebuah dadu sebagai suatu sumber informasi yang mengirim sebuah pesan setiap kali dilemparkan, alphabetnya berupa sederet angka 1, 2, 3, 4, 5, 6. Bila dadu tersebut dianggap seimbang, maka ruang probabilitas yang berhubungan dengan sumber ini adalah probabilitas biasa yang menunjuk $\frac{1}{6}$ untuk masing-masing digit.

Ruang probabilitas $\{\Omega, \mathcal{A}, P\}$ yang berhubungan dengan suatu sumber $\{A_i\}$, i = 1, 2, ..., n menjadi satu himpunan kejadian yang merupakan partisi Ω.

Bila diambil $I(A_i)$ sebuah variabel acak pada ruang $\{\Omega, \mathcal{A}, P\}$ sebagai banyaknya informasi yang diberikan ketika peristiwa A_i terjadi, maka dengan menggunakan Entropy dapat dihitung jumlah rata-rata informasi yang diberitahukan oleh sebuah sumber.

3.1 ENTROPY SEBAGAI UKURAN INFORMASI

Informasi diberikan ketika sumber mengeluarkan sebuah pesan. Dalam ruang probabilitas $\{\Omega, \mathcal{A}, P\}$ yang dihubungkan dengan suatu sumber $\{A_i\}$, i = 1, 2, ... , n menjadi satu himpunan peristiwa yang merupakan partisi dari Ω.

Definisi 3.1.1 :

Notasi $\mathfrak{A} = [A_1, A_2, \ldots, A_n]$ atau secara singkat $\mathfrak{A} = [A_i]$, i = 1, 2, ... , n mempunyai arti bahwa $\mathfrak{A}$ adalah partisi yang terdiri dari peristiwa-peristiwa A_i. Peristiwa-peristiwa tersebut disebut elemen - elemen $\mathfrak{A}$.

Masing-masing peristiwa A_i memiliki probabilitas kemunculan yang tidak selalu sama. Bila $p_i = P(A_i)$, yaitu probabilitas terjadinya peristiwa A_i, akan dibentuk suatu bagan ketidakpastian (uncertainty scheme) dari sumber $\{A_i\}$, i = 1, 2, ... , n

$$\begin{bmatrix} A_1 & A_2 & \ldots & A_n \\ p_1 & p_2 & \ldots & p_n \end{bmatrix}$$

Ketika sumber memancarkan sebuah pesan, yang merupakan salah satu dari peristiwa A_i, informasi diberikan dan ketidakpastian menjadi hilang.

Diambil $I(A_i)$ sebagai besarnya informasi yang diberikan ketika A_i terjadi. I merupakan variabel acak pada $\{\Omega, \mathcal{A}, P\}$, maka

$$E(I) = p_1 I(A_1) + p_2 I(A_2) + \ldots + p_n I(A_n)$$

merupakan jumlah informasi yang diharapkan terkandung dalam bagan ketidakpastian, atau jumlah rata-rata informasi yang diberitahukan oleh sumber ketika A_i benar - benar terjadi.

Contoh 3.1.1 :

Untuk dadu sebagai suatu sumber informasi, dua bagan ketidakpastian yang mungkin dapat diberikan, antara lain

$$\begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} & \{6\} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

dan

$$\begin{bmatrix} \{1,2,3\} & \{4,5\} & \{6\} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

Jumlah rata-rata informasi yang terkandung dalam bagan pertama

$$E(I) = \sum_{i=1}^{6} \frac{1}{6} I(\{n\})$$

dan dalam bagan kedua

$$E(I) = \frac{1}{2} I(\{1,2,3\}) + \frac{1}{3} I(\{4,5\}) + \frac{1}{6} I(\{6\})$$

Di sini tidak ada fungsi I yang didefinisikan secara terperinci. Dengan cara intuisi, semakin tidak pasti suatu peristiwa di dalam bagan akan terjadi, maka semakin besar informasi yang seharusnya diberitahukan. Dalam contoh

3.1.1, bagaimanapun jadinya fungsi I, diharapkan informasi yang lebih besar dari bagan pertama, dari pada bagan kedua. Seseorang akan belajar lebih bila mengetahui bahwa {1} terjadi, dari pada bila ia mengetahui {1,2,3} yang terjadi.

Ide mengenai ketidakpastian yang lebih besar akan menghasilkan informasi yang lebih besar, juga akan diperlihatkan dalam contoh 3.1.2.

Contoh 3.1.2 :

Diberikan dua buah bagan ketidakpastian dengan dua peristiwa (partisi biner). Bagan pertama berhubungan dengan koin seimbang, dan bagan kedua berhubungan dengan koin yang diperberat salah satu sisinya.

$$\begin{bmatrix} H & T \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ dan } \begin{bmatrix} H & T \\ \frac{9}{10} & \frac{1}{10} \end{bmatrix}$$

Di sini diharapkan semakin banyak informasi diberikan untuk kemunculan H dari bagan pertama dari pada jika H muncul dari bagan kedua.

Bagan pertama dalam contoh 3.1.1 maupun contoh 3.1.2 di atas, merupakan gambaran dari pada bagan - bagan ketidakpastian dengan probabilitas sama. Di mana masing - masing peristiwa di dalam bagan terjadi dengan probabilitas yang sama. Hanya sebagai maksud penandaan, peristiwa - peristiwa partisi dalam bagan - bagan dengan probabilitas sama akan ditunjukkan dengan notasi E_i, untuk menggantikan A_i. Sehingga bagan mempunyai bentuk

$$\begin{bmatrix} E_1 & E_2 & \cdots & E_n \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

Meskipun variabel acak I didefinisikan pada ruang sampel di dalam $\{\Omega, \mathscr{A}, P\}$, akan tetapi penggunaannya dalam teori informasi dibatasi untuk kepentingan ekspektasi dari bagan ketidakpastian. Oleh karena itu, akan diberikan istilah baru. Diambil ruang probabilitas $\{\Omega, \mathscr{A}, P\}$ yang dihubungkan dengan sebuah sumber informasi, dan I merupakan sebuah variabel acak pada ruang probabilitas tersebut. Dipandang suatu bagan

$$\begin{bmatrix} A_1 & A_2 & \cdots & A_n \\ p_1 & p_2 & \cdots & p_n \end{bmatrix}.$$

Fungsi

$$H(A_1, A_2, \ldots, A_n) = p_1 I(A_1) + p_2 I(A_2) + \ldots + p_n I(A_n)$$

disebut *Entropy* dari bagan, yang merupakan jumlah rata - rata informasi yang diberitahukan ketika ketidakpastian dihilangkan, yaitu ketika diketahui A_i mana yang terjadi. Dalam hal ini H merupakan suatu fungsi dari bagan ketidakpastian. Karena bagan ketidakpastian dibentuk dari suatu partisi, maka H berlaku jika, dan hanya jika

- $A_i \in \mathscr{A}$ untuk setiap i
- $A_i \cap A_j = \emptyset$, untuk $i \neq j$

- $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$ dan $p_1 + p_2 + \ldots + p_n = 1$

3.2 PENGEMBANGAN ENTROPY

Dalam proses pengiriman pesan dari sebuah sumber, telah diperoleh ide bahwa ketidakpastian yang lebih besar akan menghasilkan informasi yang lebih besar. Dari sini akan diambil beberapa batasan yang beralasan pada variabel acak I dan fungsi H, untuk mendapatkan pernyataan yang jelas bagi keduanya.

Dipandang dua bagan ketidakpastian dengan partisi biner

$$\begin{bmatrix} A & \bar{A} \\ p(A) & 1-p(A) \end{bmatrix} \text{ dan } \begin{bmatrix} B & \bar{B} \\ p(B) & 1-p(B) \end{bmatrix}$$

dimana $0 < p(A) < p(B) \leq \frac{1}{2}$. Bagan kedua mengandung ketidakpastian yang lebih besar dari bagan pertama, karena lebih sulit untuk memperkirakan pesan mana dari B dan $\bar{B}$ yang akan dikirimkan dari pada memperkirakan pesan mana dari bagan pertama yang akan dikirim. Dari sini diharapkan $H(B,\bar{B})$ menjadi lebih besar dari pada $H(A,\bar{A})$. Di dalam bagan pertama dimana $p(A) < \frac{1}{2} < 1-p(A) = p(\bar{A})$, diharapkan dapat diperoleh informasi lebih dengan mengetahui bahwa A terjadi dari pada mengetahui bahwa $\bar{A}$ yang terjadi. Dengan demikian persyaratan pertama bagi I dan H adalah

$$I(A) > I(B) \text{ jika } p(A) < p(B)$$

R_1 :

$$H(A,\bar{A}) < H(B,\bar{B}) \text{ jika } 0 < p(A) < p(B) \le \frac{1}{2}$$

Sebagai kesimpulan singkat pada bagian pertama dari R_1, jika p(A) = p(B) maka akan diperoleh jumlah informasi yang sama dengan mengetahui pesan mana yang terjadi. Sehingga

$$I(A) = I(B) \text{ bilamana } p(A) = p(B).$$

Apa yang telah diperoleh di sini, memperlihatkan bahwa I(A) merupakan suatu fungsi dari p(A), sehingga dalam bagan ketidakpastian

$$\begin{bmatrix} A_1 & A_2 & \ldots & A_n \\ p_1 & p_2 & \ldots & p_n \end{bmatrix}$$

berlaku

$$I(A_i) = I(p_i)$$

dan

$$\begin{aligned} H(A_1,A_2, \ldots ,A_n) &= p_1 I(A_1) + p_2 I(A_2) + \ldots + p_n I(A_n) \\ &= p_1 I(p_1) + p_2 I(p_2) + \ldots + p_n I(p_n) \\ &= H(p_1,p_2, \ldots ,p_n) \end{aligned}$$

Untuk menjelaskan pengertian bahwa informasi dibawa oleh suatu pesan tergantung pada probabilitas pesan yang sedang dikirim, diambil variabel acak X dan I dalam ruang sampel dihubungkan dengan percobaan pelemparan dua buah dadu yang dibedakan. Jika $\mathfrak{D}$ = {1, 2, 3, 4, 5, 6}

$$\Omega = \mathfrak{D} \times \mathfrak{D} = \{(r,g) \mid r,g \in \mathfrak{D}\}$$

Diambil $X(r,g) = r + g$ dan $I(r,g) = I(p(r+g))$. Peristiwa dasar dari Ω beserta probabilitasnya diberikan dalam bagan ketidakpastian :

$$\begin{bmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{bmatrix}$$

$$X(2,2) = 4 < X(3,4) = 7 < X(5,5) = 10$$

$$I(2,2) = I(p(4)) = I\left(\frac{3}{36}\right)$$

$$I(3,4) = I(p(7)) = I\left(\frac{6}{36}\right)$$

$$I(5,5) = I(p(10)) = I\left(\frac{3}{36}\right) = I(2,2)$$

Karena diharapkan informasi meningkat sebagaimana menurunnya probabilitas, maka

$$I(2,2) = I(5,5) = I\left(\frac{3}{36}\right) > I\left(\frac{6}{36}\right) = I(3,4)$$

Karena I dan H lebih bergantung pada probabilitas suatu pesan dari pada bergantung terhadap pesan itu sendiri, maka akan lebih mudah jika menyatakan I dan H sebagai fungsi dari p_i. Sehingga penulisan R_1 :

$$I(p) > I(q) \text{ jika } p < q$$

R_1 :

$$H(p,1-p) < H(q,1-q) \text{ jika } 0 < p < q \leq \frac{1}{2}$$

dan ketika $p = q$ diperoleh $I(p) = I(q)$.

Contoh 3.2.1 :

Diberikan 3 bagan ketidakpastian dengan partisi biner yang berhubungan dengan pelemparan mata uang. Bagan pertama berhubungan dengan mata uang seimbang, sedangkan dua bagan lainnya berhubungan dengan mata uang yang diperberat satu sisinya.

$$\mathfrak{A} = \begin{bmatrix} H & T \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \mathfrak{B} = \begin{bmatrix} H & T \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \quad \mathfrak{C} = \begin{bmatrix} H & T \\ \frac{9}{10} & \frac{1}{10} \end{bmatrix}$$

maka

$$H\left(\frac{1}{2},\frac{1}{2}\right) > H\left(\frac{1}{3},\frac{2}{3}\right) > H\left(\frac{9}{10},\frac{1}{10}\right)$$

Dalam suatu keadaan khusus dari bagan dua-pesan atau partisi biner dimana salah satu pesannya, misalkan A, tidak dapat dikirim, sehingga $\bar{A}$ harus dikirimkan. Karena A tidak dapat terjadi, maka P(A) = p = 0.
Dengan demikian

$$pI(p) = 0 \text{ jika } p = 0$$

Sebaliknya, karena $\bar{A}$ merupakan satu-satunya pesan yang dapat dikirimkan, $P(\bar{A}) = 1$. Maka tidak diperoleh informasi apapun dengan mengetahui pesan yang dikirim, sehingga

$$I(1) = 0$$

Dari sini Entropy dari bagan

$$\mathfrak{A} = \begin{bmatrix} A & \bar{A} \\ 0 & 1 \end{bmatrix}.$$

adalah

$$H(0,1) = 0.I(0) + 1.I(1) = 0$$

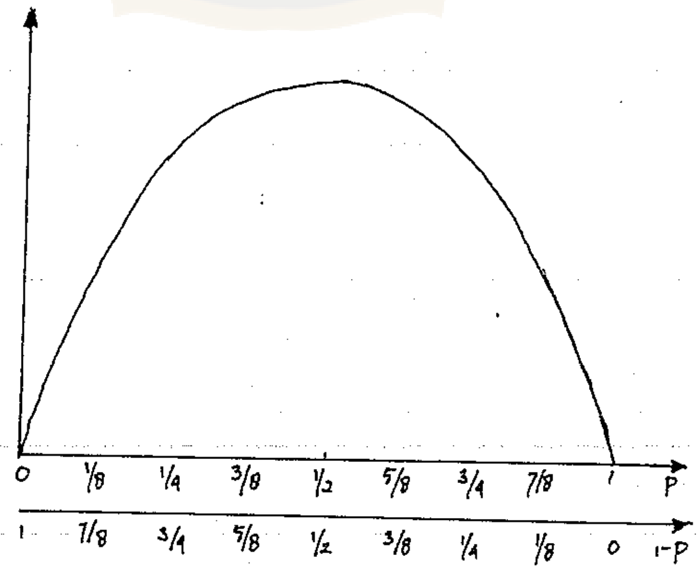
Dengan demikian syarat kedua yang dapat ditetapkan untuk I dan H

$$pI(p) = 0 \text{ jika } p = 0$$

R_2 :

$$H(0,1) = 0$$

Jika R_1 dan R_2 digabungkan dengan simetris dari H(p,1-p), menunjukkan bahwa H(p,1-p) dapat digambarkan sebagai suatu fungsi dari p, dimana diharapkan berbentuk sebuahkurva seperti pada gambar 3.2.1. Dalam hal ini ketepatan sesungguhnya dari kurva tidak dipedulikan. Yang dipentingkan adalah maksimumnya pada $p = \frac{1}{2} = 1-p$. Simetrisnya kira-kira terletak pada titik $p = \frac{1}{2}$. Dan ini meningkat dari p = 0 hingga $p = \frac{1}{2}$, dimana meliputi syarat yang dikenakan oleh R_1 dan R_2.



Gb. 3.2.1 Grafik H(p,1-p) sebagai fungsi dari p

Untuk mengetahui lebih jauh mengenai H, dipandang bagan ketidakpastian dengan tiga peristiwa

$$\begin{bmatrix} A_1 & A_2 & A_3 \\ p_1 & p_2 & p_3 \end{bmatrix}$$

$$H(p_1,p_2,p_3) = p_1I(p_1) + p_2I(p_3) + p_3I(p_3)$$

merupakan jumlah rata-rata informasi yang diberitahukan ketika diketahui mana dari A_1, A_2 dan A_3 yang dikirim. Hal ini akan dilihat dari sudut pandang lain. Dalam menyelidiki apakah A_1, A_2 atau A_3 yang dikirim,pertama ditanyakan apakah A_1 telah dikirimkan, dan jika tidak, manakah dari A_2 dan A_3 yang dikirim. Dengan demikian H akan memenuhi

$$H(A_1,A_2,A_3) = p_1I(p_1)+(1-p_1)\Big\{I(\bar{A}_1)$$

$$+ p(A_2|\bar{A}_1)I(A_2|\bar{A}_1) + p(A_3|\bar{A}_1)I(A_3|\bar{A}_1)\Big\}$$

Dengan menggunakan definisi 2.2.1 dan sebagaimana yang telah diketahui bahwa $A_j \subset \bar{A}_1$, j = 2, 3, maka

$$p(A_j|\bar{A}_1) = \frac{p(A_j \cap A_1)}{p(\bar{A}_1)} = \frac{p(A_j)}{p(\bar{A}_1)} = \frac{p_j}{1-p_1}$$

Mensubstitusikan ini ke dalam persamaan $H(A_1,A_2,A_3)$ dan dituliskan dalam sudut pandang p_1, p_2 dan p_3, diperoleh :

$$H(p_1,p_2,p_3) = p_1I(p_1) + (1-p_1)\Big\{I(1-p_1)$$

$$+ \frac{p_2}{1-p_1} I\left(\frac{p_2}{1-p_1}\right) + \frac{p_3}{1-p_1} I\left(\frac{p_3}{1-p_1}\right)\Bigg\}$$

$$= p_1 I(p_1) + (1-p_1) I(1-p_1)$$

$$+ p_2 I\left(\frac{p_2}{1-p_1}\right) + p_3 I\left(\frac{p_3}{1-p_1}\right)$$

Persamaan ini dapat dikelompokkan menjadi

$$H(p_1,p_2,p_3) = p_1 I(p_1) + (1-p_1) I(1-p_1)$$

$$+ (1-p_1)\left\{ \frac{p_2}{1-p_1} I\left(\frac{p_2}{1-p_1}\right) + \frac{p_3}{1-p_1} I\left(\frac{p_3}{1-p_1}\right)\right\}$$

Ditulis dalam fungsi H

$$H(p_1,p_2,p_3) = H(p_1,1-p_1) + (1-p_1) H\left(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}\right)$$

Dengan menggunakan persamaan terakhir akan dicari suatu syarat bagi I

$$p_1 I(p_1) + p_2 I(p_2) + p_3 I(p_3)$$

$$= p_1 I(p_1) + (1-p_1) I(1-p_1) + p_2 I\left(\frac{p_2}{1-p_1}\right) + p_3 I\left(\frac{p_3}{1-p_1}\right)$$

Karena $p_2 + p_3 = 1 - p_1$, persamaan ini dapat menjadi

$$p_2 I(p_3) + p_3 I(p_3)$$

$$= (p_2 + p_3) I(1-p_1) + p_2 I\left(\frac{p_2}{1-p_1}\right) + p_3 I\left(\frac{p_3}{1-p_1}\right)$$

$$p_2\left[I(p_2) - I(1-p_1) - I\left(\frac{p_2}{1-p_1}\right)\right] + p_3\left[I(p_3) - I(1-p_1) - I\left(\frac{p_3}{1-p_1}\right)\right] = 0$$

Persamaan ini pasti akan terpenuhi jika untuk dua peristiwa A dan B dengan $A \subset B$, dimana $p = p(A) < p(B) = q$

$$I(p) - I(q) = I\left(\frac{p}{q}\right)$$

Satu fungsi umum dimana fungsi perbandingan dua bilangan sama dengan selisih dari fungsi dua bilangan tersebut adalah logaritma. Oleh karena itu, logaritma merupakan pilihan yang mungkin untuk I.

Untuk memperoleh satu syarat akhir bagi H, persamaan

$$H(p_1,p_2,p_3) = H(p_1,1-p_1) + (1-p_1)H\left(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}\right)$$

akan diperluas menjadi persamaan dengan lebih dari tiga peristiwa. Dalam bagan

$$\begin{pmatrix} A_1 & A_2 & , \dots & A_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

himpunan $A_1 \cup A_2 \cup \dots \cup A_{n-2} = A$ dan $p_1+p_2+\dots+p_{n-2}=1$, dan dibentuk bagan

$$\begin{pmatrix} A & A_{n-1} & A_n \\ p & p_{n-1} & p_n \end{pmatrix}$$

Yang dikehendaki adalah entropy-entropy dari kedua bagan di atas memiliki besar yang sama, karena keduanya mengandung himpunan pesan yang sama. Dari sini diharapkan

$$H(p_1,p_2,\dots,p_n) = H(p,p_{n-1},p_n)$$

Dengan menggunakan persamaan $H(p_1,p_2,p_3)$ untuk menyatakan $H(p,p_{n-1},p_n)$,

$$H(p_1,p_2,\ldots,p_n) = H(p,1-p) + (1-p)H\left(\frac{p_{n-1}}{1-p}, \frac{p_n}{1-p}\right).$$

Dinyatakan dalam fungsi I

$$\sum_{i=1}^{n} p_i I(p_i) = pI(p)+(1-p)I(1-p)+p_{n-1}I\left(\frac{p_{n-1}}{1-p}\right)+p_n I\left(\frac{p_n}{1-p}\right)$$

Sebagaimana dalam syarat bagi I, bahwa $I(p)-I(q) = I\left(\frac{p}{q}\right)$

bilamana $p < q$, dan karena $p_{n-1} + p_n = 1 - p$, maka

$$p_{n-1}\Big[I(p_{n-1})-I(1-p)\Big] + p_n\Big[I(p_n)-I(1-p)\Big]$$

$$= p_{n-1}I\left(\frac{p_{n-1}}{1-p}\right) + p_n I\left(\frac{p_n}{1-p}\right)$$

$$p_{n-1}I(p_{n-1}) + p_n I(p_n) = (p_{n-1}+p_n)I(1-p) + p_{n-1}I\left(\frac{p_{n-1}}{1-p}\right) + p_n I\left(\frac{p_n}{1-p}\right)$$

$$p_{n-1}I(p_{n-1}) + p_n I(p_n) = (1-p)I(1-p) + p_{n-1}I\left(\frac{p_{n-1}}{1-p}\right) + p_n I\left(\frac{p_n}{1-p}\right)$$

Persamaan di atas dimasukkan ke dalam fungsi I, diperoleh

$$\sum_{i=1}^{n} p_i I(p_i) = \sum_{i=1}^{n-2} p_i I(p_i) + p_{n-1}I(p_{n-1}) + p_n I(p_n)$$

$$= \sum_{i=1}^{n-2} p_i I(p_i) + (1-p)I(1-p) + p_{n-1}I\left(\frac{p_{n-1}}{1-p}\right) + p_n I\left(\frac{p_n}{1-p}\right)$$

dan ditulis dalam notasi H

$H(p_1, p_2, \ldots, p_n)$

$$= H(p_1, \ldots, p_{n-2}, 1-p) + (1-p)H\left(\frac{p_{n-1}}{1-p}, \frac{p_n}{1-p}\right)$$

Dengan demikian syarat ketiga yang dikenakan pada I dan H adalah

$$I(p) - I(q) = I\left(\frac{p}{q}\right) \text{ bilamana } p < q$$

R_3 :

$H(p_1, p_2, \ldots, p_n)$

$$= H(p_1, \ldots, p_{n-2}, 1-p) + (1-p)H\left(\frac{p_{n-1}}{1-p}, \frac{p_n}{1-p}\right)$$

dimana $p_1 + p_2 + \ldots + p_{n-2} = p$

Dari perluasan di atas, secara ringkas akan diberikan definisi mengenai I dan H, beserta syarat-syarat yang diharapkan dari keduanya. Diberikan suatu sumber informasi dan ruang probabilitas yang berhubungan dengannya, $\{\Omega, \mathcal{A}, P\}$. Dibentuk suatu bagan ketidakpastian

$$\begin{pmatrix} A_1 & A_2 & \ldots & A_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

dimana himpunan peristiwa $\{A_i\}$, $i = 1, 2, \ldots, n$ merupakan suatu partisi dari Ω dan $\sum_{i=1}^{n} p_i = 1$.

Dimasukkan suatu variabel acak I, yang disebut unit informasi, dan mendefinisikan Entropy H sebagai suatu fungsi dalam bagan ketidakpastian. Dengan definisi H adalah ekspektasi dari variabel acak I, besarnya informasi dalam bagan ketidakpastian, yaitu

$$H(p_1,p_2,\ldots,p_n) = p_1 I(p_1) + p_2 I(p_2) + \ldots + p_n I(p_n) \qquad \ldots\ldots 3.2.1$$

Untuk memperoleh ungkapan yang jelas bagi I dan H, diberikan tiga syarat yang sesuai dengan pengertian informasi secara umum. Syarat-syarat tersebut adalah

R_1 :

$$I(p) > I(q) \text{ jika } p < q$$

$$H(p,1-p) < H(q,1-q) \text{ jika } 0 < p < q \le \tfrac{1}{2}$$

R_2 :

$$pI(p) = 0 \text{ jika } p = 0$$

$$H(0,1) = 0$$

R_3 :

$$I(p) - I(q) = I\left(\frac{p}{q}\right) \text{ bilamana } p < q$$

$$H(p_1,p_2,\ldots,p_n)$$

$$= H(p_1,\ldots,p_{n-2},1-p) + (1-p)H\left(\frac{p_{n-1}}{1-p}, \frac{p_n}{1-p}\right)$$

dimana $\sum_{i=1}^{n-2} p_i = p$

3.3 SIFAT-SIFAT ENTROPY

Teorema 3.3.1 :

Entropy dari bagan

$$\begin{bmatrix} A_1 & A_2 & \dots & A_n & B_1 & B_2 & \dots & B_m \\ p_1 & p_2 & \dots & p_n & q_1 & q_2 & \dots & q_m \end{bmatrix}$$

dimana

$$\sum_{i=1}^{n} p_i = p \quad \text{dan} \quad \sum_{j=1}^{m} q_j = q = 1 - p$$

memenuhi persamaan

$$H(p_1,p_2,\dots,p_n,q_1,q_2,\dots,q_m)$$

$$= H(p_1,p_2,\dots,p_n,q) + qH\left(\frac{q_1}{q}, \frac{q_2}{q}, \dots, \frac{q_m}{q}\right)$$

Bukti :

Jika m = 2, maka teorema ini memenuhi syarat H dalam R_3. Pembuktian teorema ini akan dijalankan dengan menggunakan Prinsip Induksi Matematika.

Jika teorema benar untuk m = k

$$H(p_1,p_2,\dots,p_n,q_1,q_2,\dots,q_k)$$

$$= H(p_1,p_2,\dots,p_n,q) + qH\left(\frac{q_1}{q}, \frac{q_2}{q}, \dots, \frac{q_k}{q}\right) \qquad \dots\dots 3.3.1$$

dimana $q = \sum_{j=1}^{k} q_j$.

Akan dibuktikan bahwa teorema benar untuk m = k + 1.

Diambil m = k + 1 dan diberikan $\bar{q} = \sum_{j=1}^{k} q_{j+1} = q - q_1$.

Karena teorema memenuhi untuk m = k, maka

$$H(p_1,p_2,\ldots,p_n,q_1,q_2,\ldots,q_{k+1})$$

$$= H(p_1,p_2,\ldots,p_n,q_1,\bar{q}) + \bar{q}H\left(\frac{q_2}{\bar{q}},\ldots,\frac{q_{k+1}}{\bar{q}}\right)$$

$$= H(p_1,p_2,\ldots,p_n,q) + qH\left(\frac{q_1}{q},\frac{\bar{q}}{q}\right)$$

$$+ \bar{q}H\left(\frac{q_2}{\bar{q}},\ldots,\frac{q_{k+1}}{\bar{q}}\right) \quad \ldots\ldots 3.3.2$$

dimana persamaan (*3.3.2*) diperoleh dari syarat H dalam R_3 dengan $q = q_1 + \bar{q}$.

Di lain pihak, karena teorema benar untuk m = k, dari persamaan (*3.3.1*)

$$H\left(\frac{q_1}{q},\frac{q_2}{q},\ldots,\frac{q_{k+1}}{q}\right)$$

$$= H\left(\frac{q_1}{q},\frac{\bar{q}}{q}\right) + \frac{\bar{q}}{q}H\left(\frac{q_2/q}{\bar{q}/q},\ldots,\frac{q_{k+1}/q}{\bar{q}/q}\right)$$

$$H\left(\frac{q_1}{q},\frac{q_2}{q},\ldots,\frac{q_{k+1}}{q}\right) = H\left(\frac{q_1}{q},\frac{\bar{q}}{q}\right) + \frac{\bar{q}}{q}H\left(\frac{q_2}{\bar{q}},\ldots,\frac{q_{k+1}}{\bar{q}}\right)$$

dimana $q = q_1 + \bar{q}$ dan $\bar{q} = \sum_{j=1}^{k} q_{j+1}$.

Persamaan di atas dikalikan dengan q diperoleh

$$qH\left(\frac{q_1}{q},\frac{q_2}{q},\ldots,\frac{q_{k+1}}{q}\right) = qH\left(\frac{q_1}{q},\frac{\bar{q}}{q}\right) + \bar{q}H\left(\frac{q_2}{\bar{q}},\ldots,\frac{q_{k+1}}{\bar{q}}\right)$$

Dari sini

$$H(p_1,p_2,\ldots,p_n,q) + qH\left(\frac{q_1}{q}, \frac{q_2}{q}, \ldots, \frac{q_{k+1}}{q}\right)$$

$$= H(p_1,p_2,\ldots,p_n,q) + qH\left(\frac{q_1}{q}, \frac{\bar{q}}{q}\right) + \bar{q}H\left(\frac{q_2}{\bar{q}}, \ldots, \frac{q_{k+1}}{\bar{q}}\right)$$

$$H(p_1,\ldots,p_n,q) + qH\left(\frac{q_1}{q}, \ldots, \frac{q_{k+1}}{q}\right)$$

$$= H(p_1,\ldots,p_n,q_1,\ldots,q_{k+1})$$

Karena teorema benar untuk $m = k + 1$, maka teorema benar untuk seluruh $m \in I^+$. ■

Contoh 3.3.1 :

Diberikan suatu bagan ketidakpastian

$$\begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} & \{6\} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Entropy dari bagan di atas dapat dihitung dengan menggunakan teorema 3.3.1

$$H\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$$

$$= H\left(\frac{1}{6}, \frac{1}{6}, \frac{2}{3}\right) + \frac{2}{3}H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right).$$

Teorema 3.3.2 :

Dalam bagan

$$\begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1m_1} & A_{21} & A_{22} & \ldots & A_{2m_2} & \ldots & A_{n1} & A_{n2} & \ldots & A_{nm_n} \\ p_{11} & p_{12} & \ldots & p_{1m_1} & p_{21} & p_{22} & \ldots & p_{2m_2} & \ldots & p_{n1} & p_{n2} & \ldots & p_{nm_n} \end{bmatrix}$$

dimana

$$p_i = p_{i1} + p_{i2} + \dots + p_{im_i} \quad \text{sehingga} \quad \sum_{i=1}^{n} p_i = 1$$

maka

$$H\left(p_{11}, p_{12}, \dots, p_{1m_1}, \dots, p_{n1}, p_{n2}, \dots, p_{nm_n} \right)$$

$$= H(p_1, p_2, \dots, p_n) + \sum_{i=1}^{n} p_i H\left(\frac{p_{i1}}{p_i}, \frac{p_{i2}}{p_i}, \dots, \frac{p_{im_i}}{p_i} \right).$$

Bukti :

Dari teorema 3.3.1

$$H\left(p_{11}, \dots, p_{n1}, p_{n2}, \dots, p_{nm_n}\right)$$

$$= H\left(p_{11}, \dots, p_{n-1m_{n-1}}, p_n\right) + p_n H\left(\frac{p_{n1}}{p_n}, \frac{p_{n2}}{p_n}, \dots, \frac{p_{nm_n}}{p_n} \right).$$

Karena H simetrik dalam seluruh pernyataannya, maka p_n dapat diputar ke posisi yang paling kiri.

$$H\left(p_{11}, p_{12}, \dots, p_{nm_n}\right)$$

$$= H\left(p_n, p_{11}, \dots, p_{n-2\, m_{n-2}}, p_{n-1}\right)$$

$$+ p_{n-1} H\left(\frac{p_{n-1\,1}}{p_{n-1}}, \dots, \frac{p_{n-1\, m_{n-1}}}{p_{n-1}} \right) + p_n H\left(\frac{p_{n1}}{p_n}, \dots, \frac{p_{nm_n}}{p_n} \right).$$

$$= H\left(p_{n-1}, p_n, p_{11}, \dots, p_{n-3\, m_{n-3}}, p_{n-2}\right)$$

$$+ p_{n-2} H\left(\frac{p_{n-2\,1}}{p_{n-2}}, \dots, \frac{p_{n-2\, m_{n-2}}}{p_{n-2}} \right) + p_{n-1} H\left(\frac{p_{n-1\,1}}{p_{n-1}}, \dots, \frac{p_{n-1\, m_{n-1}}}{p_{n-1}} \right)$$

$$+ p_n H\left(\frac{p_{n1}}{p_n}, \ldots, \frac{p_{nm_n}}{p_n}\right).$$

$$= H\left(p_{n-1}, p_n, p_{11}, \ldots, p_{n-3\ m_{n-3}}, p_{n-2}\right)$$

$$+ \sum_{j=0}^{2} p_{n-j} H\left(\frac{p_{n-j\ 1}}{p_{n-j}}, \ldots, \frac{p_{n-j\ m_{n-j}}}{p_{n-j}}\right).$$

Jika proses ini dijalankan untuk seluruh i

$$H\left(p_{11}, p_{12}, \ldots, p_{1m_1}, \ldots, p_{n1}, p_{n2}, \ldots, p_{nm_n}\right)$$

$$= H(p_1, p_2, \ldots, p_n) + \sum_{i=1}^{n} p_i H\left(\frac{p_{i1}}{p_i}, \frac{p_{i2}}{p_i}, \ldots, \frac{p_{im_i}}{p_i}\right).$$

dan teorema terbukti. ∎

Contoh 3.3.2 :

Dengan menggunakan teorema 3.3.2 untuk bagan

$$\begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} & \{6\} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}.$$

Entropynya adalah

$$H\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$+ \left[\frac{1}{3} H\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{3} H\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{3} H\left(\frac{1}{2}, \frac{1}{2}\right)\right].$$

$$= H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) + 3\left[\frac{1}{3} H\left(\frac{1}{2}, \frac{1}{2}\right)\right].$$

$$H\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) + H\left(\frac{1}{2}, \frac{1}{2}\right).$$

Dengan menggunakan teorema tersebut di atas, akan dicari suatu fungsi numerik bagi I. Diberikan suatu bagan ketidakpastian dengan probabilitas sama

$$\begin{bmatrix} E_1 & E_2 & \dots & E_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix}.$$

Ditentukan Entropy dari bagan di atas

$$H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = F(n) \qquad \dots\dots 3.3.3$$

Jika masing-masing E_i dinyatakan sebagai gabungan dari m peristiwa dengan probabilitas yang sama, yaitu jika

$E_i = E_{i1} \cup E_{i2} \cup \dots \cup E_{im}$ dan $P(E_{ij}) = \frac{1}{nm}$ untuk seluruh j, maka menggunakan teorema 3.3.2

$$H\left(\frac{1}{nm}, \frac{1}{nm}, \dots, \frac{1}{nm}\right)$$

$$= H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) + \frac{1}{n}\sum_{i=1}^{n} H\left[\frac{\frac{1}{nm}}{\frac{1}{n}}, \frac{\frac{1}{nm}}{\frac{1}{n}}, \dots, \frac{\frac{1}{nm}}{\frac{1}{n}}\right].$$

$$= H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) + \frac{1}{n}\left[nH\left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}\right)\right].$$

$$= H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) + H\left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}\right) \qquad \dots\dots 3.3.4$$

Bila persamaan (*3.3.4*) di atas ditulis dalam bentuk fungsi F,

$$F(nm) = F(n) + F(m).$$

Satu fungsi utama yang memenuhi persamaan fungsional ini adalah logaritma. Diambil F(n) = log n

$$\log nm = \log n + \log m.$$

Sebagaimana dalam persamaan (*3.2.1*) bahwa

$$H(p_1, p_2, \dots, p_n) = p_1 I(p_1) + p_2 I(p_2) + \dots + p_n I(p_n)$$

maka

$$H\left(\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n}\right) = \frac{1}{n} I\left(\frac{1}{n}\right) + \ldots + \frac{1}{n} I\left(\frac{1}{n}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} I\left(\frac{1}{n}\right)$$

$$= \frac{1}{n}\left\{nI\left(\frac{1}{n}\right)\right\} = I\left(\frac{1}{n}\right)$$

Dengan memandang persamaan (3.3.3)

$$F(n) = H\left(\frac{1}{n},\frac{1}{n},\ldots,\frac{1}{n}\right) = I\left(\frac{1}{n}\right)$$

maka penetapan $F(n) = \log n$ menghasilkan

$$I\left(\frac{1}{n}\right) = \log n.$$

Dengan demikian, pilihan $\log n$ bagi $F(n)$ memberikan suatu ukuran yang berguna dari informasi untuk suatu bagan ketidakpastian dengan probabilitas sama. Hal ini memberikan suatu arahan dalam menemukan jumlahan yang tepat untuk informasi, jika pesan yang akan disampaikan tidak seluruhnya muncul dengan probabilitas sama.

Karena $\log n = -\log \frac{1}{n}$, maka besarnya informasi yang diberitahukan dengan mengetahui mana dari pesan berprobabilitas sama E_i yang dikirim, sama dengan negatif dari probabilitas pesan tersebut. Yaitu

$$I(E_i) = I\left(\frac{1}{n}\right) = \log n = -\log\frac{1}{n} = -\log p(E_i).$$

Hal ini memperlihatkan bahwa suatu definisi yang dapat dikerjakan dari informasi berupa

$$I(p) = -\log p$$

dan dengan ini dapat diperoleh

$$H(p_1,p_2,\ldots,p_n) = -\sum_{i=1}^{n} p_i \log p_i.$$

Demikian pula untuk $p_i = \frac{1}{n}$

$$H(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n}$$

$$= -n(\frac{1}{n} \log \frac{1}{n}) = \log n.$$

Karena $p < 1$ maka $\log p < 0$ dan $I(p) = -\log p > 0$.

Sekarang akan dibuktikan apakah fungsi pilihan untuk I ini memenuhi syarat-syarat R_1, R_2 dan R_3. Pertama akan dilihat pada syarat bagi I.

Jika $p < q$

$$I(p) = -\log p > -\log q = I(q).$$

Demikian pula jika $p < q$

$$I(p) - I(q) = -\log p + \log q = \log \frac{q}{p} = -\log \frac{p}{q} = I\left(\frac{p}{q}\right)$$

Lebih jauh lagi, berdasarkan pada metoda analisis bahwa di dalam persyaratan bagi informasi ini tidak dikerjakan suatu asumsi, syarat bahwa untuk $p = 0$ maka $pI(p) = 0$ merupakan suatu konsekuensi logis dari pemilihan $I(p) = -\log p$. Sedangkan dengan menggunakan d'Hospital diperoleh

$$\lim_{p \to 0} -p \log p = \lim_{p \to 0} \frac{\log p}{-^1/p}$$

$$= \lim_{p \to 0} \frac{^1/p}{^1/p^2}$$

$$= \lim_{p \to 0} p = 0.$$

Dengan demikian pemilihan $I(p) = -\log p$ memenuhi seluruh persyaratan bagi I.

Kemudian akan dibuktikan apakah pemilihan $I(p) = -\log p$ memenuhi persyaratan bagi H.

Karena $pI(p) = 0$ untuk $p = 0$, dan karena $\log 1 = 0$, maka

$$H(0,1) = -1 \log 1 = 0.$$

Dan R_2 dipenuhi. Demikian pula karena pilihan $I(p) = -\log p$ menghasilkan $I\left(\frac{p}{q}\right) = I(p) - I(q)$, dengan menggunakan (3.2.1)

$$H(p_1,p_2,\ldots,p_n) = \sum_{i=1}^{n} p_i I(p_i)$$

$$= -\sum_{i=1}^{n-2} p_i \log p_i - p_{n-1}\log p_{n-1} - p_n \log p_n$$

$$+ \{(1-p)\log(1-p) - (1-p)\log(1-p)\}$$

$$= -\sum_{i=1}^{n-2} p_i \log p_i - (1-p)\log(1-p)$$

$$-p_{n-1}\log p_{n-1} - p_n \log p_n - (1-p)\log(1-p)$$

dimana $p = \sum_{i=1}^{n-2} p_i$. Karena $p_{n-1} + p_n = 1 - p$, maka

$$-p_{n-1}\log p_{n-1} - p_n \log p_n - (1-p)\log(1-p)$$

$$= -p_{n-1}\log p_{n-1} - p_n \log p_n - (p_{n-1} + p_n)\log(1-p)$$

$$= p_{n-1}\left[\log(1-p) - \log p_{n-1}\right] + p_n\left[\log(1-p) - \log p_n\right]$$

$$= p_{n-1}\log\left(\frac{1-p}{p_{n-1}}\right) + p_n \log\left(\frac{1-p}{p_n}\right)$$

$$= -p_{n-1}\log\left(\frac{p_{n-1}}{1-p}\right) - p_n \log\left(\frac{p_n}{1-p}\right)$$

sehingga

$$H(p_1,p_2,\ldots,p_n) = \left(-\sum_{i=1}^{n-2} p_i \log p_i - (1-p)\log(1-p)\right)$$

$$+ \left(-p_{n-1}\log\left(\frac{p_{n-1}}{1-p}\right) - p_n \log\left(\frac{p_n}{1-p}\right)\right)$$

$$= \sum_{i=1}^{n-2} p_i I(p_i) + (1-p)I(1-p) +$$

$$(1-p)\left[\frac{p_{n-1}}{1-p} I\left(\frac{p_{n-1}}{1-p}\right) + \frac{p_n}{1-p} I\left(\frac{p_n}{1-p}\right)\right]$$

$$= H(p_1,\ldots,p_{n-2},1-p) + (1-p)H\left(\frac{p_{n-1}}{1-p}, \frac{p_n}{1-p}\right)$$

Dan syarat R_3 dipenuhi.

Untuk menunjukkan bahwa R_1 dipenuhi, yaitu untuk membuktikan bahwa $I(p) = -\log p$ memenuhi syarat $H(p,1-p) < H(q,1-q)$ bilamana $0 < p < q \le \frac{1}{2}$, dimisalkan suatu fungsi

$$H(x,1-x) = -x \log x -(1-x)\log(1-x) = f(x)$$

dimana $0 < x < 1$.

Untuk memperoleh nilai stasioner dari $f(x)$, akan dicari $f'(x)$.

$$f'(x) = -x\frac{1}{x} - \log x -(1-x)\frac{-1}{(1-x)} -(-1)\log(1-x)$$

$$= -1 - \log x + 1 + \log (1-x)$$

$$= \log\frac{1-x}{x}.$$

Dari sini, fungsi $f(x)$ merupakan fungsi naik dalam interval $(0, \frac{1}{2})$, karena dalam interval tersebut $f'(x) > 0$ dengan titik kritisnya berada pada $x = \frac{1}{2}$.

$$f'(x) = 0 \Leftrightarrow \log\frac{1-x}{x} = 0$$

$$\frac{1-x}{x} = 1$$

$$1-x = x$$

$$x = \frac{1}{2}$$

Jika diambil $0 < x < y \le \frac{1}{2}$ maka $f'(x) < f'(y)$. Ini berarti bahwa untuk $0 < x < y \le \frac{1}{2}$ berlaku

$$f(x) < f(y)$$

$$H(x,1-x) < H(y,1-y)$$

Dengan demikian pilihan $I(p) = -\log p$ memenuhi seluruh syarat R_1, R_2 dan R_3 yang dikenakan pada I dan H. Karena $I(p) = -\log p$ merupakan satu-satunya fungsi yang secara serentak memenuhi ketiga syarat tersebut, maka sebagai suatu ukuran informasi, $-\log p$ adalah tunggal.

Dari keseluruhan yang telah diuraikan di atas, untuk setiap peristiwa A dengan probabilitas p dan untuk setiap bagan ketidakpastian

$$\begin{bmatrix} A_1 & A_2 & \cdots & A_n \\ p_1 & p_2 & \cdots & p_n \end{bmatrix}$$

yang didefinisikan pada ruang probabilitas $\{\Omega, \mathcal{A}, P\}$, berlaku

Definisi 3.3.1 :

$$I(A) = I(p) = -\log p, \text{ dan}$$

$$H(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} p_i I(p_i) = -\sum_{i=1}^{n} p_i \log p_i .$$

Sifat-sifat lain yang penting dari Entropy H :

– Pecahan Peristiwa.

Diberikan partisi $\mathfrak{A} = [A_1, A_2, \ldots, A_n]$. Dan dibentuk partisi $\mathfrak{B} = [B_a, B_b, A_2, \ldots, A_n]$ yang didapat dengan memecah peristiwa A_1 ke dalam dua peristiwa B_a dan B_b. Bagan ketidakpastiannya berbentuk

$$\mathfrak{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_n \\ p_1 & p_2 & \dots & p_n \end{bmatrix}$$

dan

$$\mathfrak{B} = \begin{bmatrix} B_a & B_b & A_2 & \dots & A_n \\ q_a & q_b & p_2 & \dots & p_n \end{bmatrix}$$

dimana $p_1 = q_a + q_b$. Maka berlaku

$$H(p_1,p_2,\dots,p_n) < H(q_a,q_b,p_2,\dots,p_n)$$

Bukti :

$$H(p_1,p_2,\dots,p_n) = -p_1 \log p_1 - \sum_{i=2}^{n} p_i \log p_i$$

dan

$$H(q_a,q_b,p_2,\dots,p_n) = -q_a \log q_a - q_b \log q_b - \sum_{i=2}^{n} p_i \log p_i . \qquad \dots\dots 3.3.5$$

Karena $p_1 = q_a + q_b$, maka

$$H(p_1,p_2,\dots,p_n) = -(q_a+q_b) \log (q_a+q_b) - \sum_{i=2}^{n} p_i \log p_i . \qquad \dots\dots 3.3.6$$

Dari persamaan (3.3.5) dan (3.3.6), serta jika diberikan fungsi $F(p) = -p \log p$ diperoleh

$$H(p_1,p_2,\dots,p_n) - F(q_a+q_b) = H(q_a,q_b,p_2,\dots,p_n) - F(q_a) - F(q_b).$$

Dengan menggunakan teorema 2.4.1 bahwa

$$F(q_a+q_b) < F(q_a) + F(q_b)$$

maka

$$H(p_1,p_2,\dots,p_n) < H(q_a,q_b,p_2,\dots,p_n) \qquad \blacksquare$$

- **Entropy Maksimum.**

Nilai maksimum dari Entropy diperoleh dari bagan dengan ketidakpastian terbesar.

$$H(p_1,p_2,\ldots,p_n) \leq H(\tfrac{1}{n},\tfrac{1}{n},\ldots,\tfrac{1}{n})$$

Bukti :

$$H(p_1,p_2,\ldots,p_n) = -\sum_{i=1}^{n} p_i \log p_i$$

dan

$$H(\tfrac{1}{n},\tfrac{1}{n},\ldots,\tfrac{1}{n}) = -\sum_{i=1}^{n} \tfrac{1}{n} \log \tfrac{1}{n}$$

$$= -n.\tfrac{1}{n} \log \tfrac{1}{n}$$

$$= \log n.$$

Dengan menggunakan pertidaksamaan (*2.4.3*), dan mengambil

$$a_i = p_i \quad \text{dan} \quad b_i = \tfrac{1}{n}$$

maka

$$-\sum_{i=1}^{n} p_i \log p_i \leq -\sum_{i=1}^{n} p_i \log \tfrac{1}{n} = \log n \sum_{i=1}^{n} p_i = \log n.$$

Sehingga

$$H(p_1,p_2,\ldots,p_n) \leq H(\tfrac{1}{n},\tfrac{1}{n},\ldots,\tfrac{1}{n}) \quad \blacksquare$$

Sebenarnya tidak ada ketentuan dalam mengambil bilangan pokok logaritma, baik logaritma umum dengan bilangan pokok 10 atau logaritma dasar dengan bilangan pokok e. Dalam aplikasi logaritma di sini, akan digunakan logaritma dengan bilangan pokok 2. Dan satuan dari entropy disebut *bit* informasi per elemen peristiwa. Bit merupakan kependekan dari binary digit, yang mempunyai arti bahwa informasi yang diberikan dalam bentuk bilangan biner.

Contoh 3.3.3 :

Memandang contoh 3.2.2, contoh 3.3.1 dan contoh 3.3.2. Dengan menggunakan definisi 3.3.1 akan dihitung entropy dari bagan

$$\begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} & \{6\} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}.$$

$$H(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6})$$

$$= H(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{3}) + \tfrac{1}{3} H(\tfrac{1}{2}, \tfrac{1}{2})$$

$$= 4(-\tfrac{1}{6} \log_2 \tfrac{1}{6}) - \tfrac{1}{3} \log_2 \tfrac{1}{3} + \tfrac{1}{3}(2[-\tfrac{1}{2} \log_2 \tfrac{1}{2}])$$

$$= 4(0{,}43120) + 0{,}52827 + \tfrac{2}{3}(0{,}49999)$$

$= 2{,}5864$ bit per elemen peristiwa

$$H(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6})$$

$$= H(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{2}{3}) + \tfrac{2}{3} H(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4})$$

$$= 2(-\tfrac{1}{6} \log_2 \tfrac{1}{6}) - \tfrac{2}{3} \log_2 \tfrac{2}{3} + \tfrac{2}{3}(4[-\tfrac{1}{4} \log_2 \tfrac{1}{4}])$$

$$= 2(0{,}43120) + 0{,}39054 + \tfrac{8}{3}(0{,}49999)$$

$= 2{,}5863$ bit per elemen peristiwa.

$$H(\tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}) =$$

$$= H(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}) + H(\tfrac{1}{2}, \tfrac{1}{2})$$

$$= 3(-\tfrac{1}{3} \log_2 \tfrac{1}{3}) + 2(-\tfrac{1}{2} \log_2 \tfrac{1}{2})$$

$$= 3(0{,}52827) + 2(0{,}49999)$$

$= 2{,}5848$ bit per elemen peristiwa.

Dan dari definisi 3.3.1,

$$H\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = 6\left(-\frac{1}{6}\log_2 \frac{1}{6}\right)$$

$$= 6(0{,}43120)$$

$= 2{,}5872$ bit per elemen peristiwa.

3.4 ENTROPY BERSYARAT

Ruang probabilitas $\{\Omega, \mathscr{A}, P\}$ dihubungkan dengan suatu sumber informasi. Dan diberikan dua bagan ketidakpastian dari dua partisi $\mathfrak{A} = [A_1, A_2, \ldots, A_n]$ dan $\mathfrak{B} = [B_1, B_2, \ldots, B_m]$

$$\mathfrak{A} = \begin{bmatrix} A_1 & A_2 & \ldots & A_n \\ p_1 & p_2 & \ldots & p_n \end{bmatrix} \qquad \mathfrak{B} = \begin{bmatrix} B_1 & B_2 & \ldots & B_m \\ q_1 & q_2 & \ldots & q_m \end{bmatrix}$$

Kemudian dibentuk sebuah bagan ketiga $\mathfrak{AB}$ dimana elemennya merupakan pasangan berurutan (A_i, B_j), $A_i \in \mathfrak{A}$, $B_j \in \mathfrak{B}$ terjadi dengan probabilitas $p(A_i, B_j) = p(B_j | A_i)p(A_i)$.

Untuk ringkasnya, ditulis sebagai berikut

$$H(p_1, p_2, \ldots, p_n) = H(\mathfrak{A})$$

$$H(q_1, q_2, \ldots, q_m) = H(\mathfrak{B})$$

$$H(p(A_1, B_1), p(A_1, B_2), \ldots, p(A_n, B_m)) = H(\mathfrak{A}, \mathfrak{B})$$

Maka dengan menggunakan definisi 3.3.1

$$H(\mathfrak{A}, \mathfrak{B}) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(A_i, B_j)\log p(A_i, B_j)$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} p(B_j | A_i)p(A_i)\log p(B_j | A_i)p(A_i)$$

$$= -\sum_{i=1}^{n} \sum_{j=1}^{m} p(B_j|A_i)p(A_i)\left[\log p(B_j|A_i) + \log p(A_i)\right]$$

$$= -\sum_{i=1}^{n} p(A_i) \sum_{j=1}^{m} p(B_j|A_i)\log p(B_j|A_i)$$

$$-\sum_{i=1}^{n} p(A_i)\log p(A_i)\cdot \sum_{j=1}^{m} p(B_j|A_i)$$

Untuk A_i tetap, peristiwa (A_i, B_j) merupakan suatu bentuk partisi dari A_i. Karena $\bigcup_{j=1}^{m} B_j = \Omega$

$$\sum_{j=1}^{m} p(B_j|A_i) = \sum_{j=1}^{m} \frac{p(A_i, B_j)}{p(A_i)} = \frac{\sum_{j=1}^{m} p(A_i, B_j)}{p(A_i)}$$

$$= \frac{p\left[A_i, \bigcup_{j=1}^{m} B_j\right]}{p(A_i)} = \frac{p(A_i, \Omega)}{p(A_i)} = \frac{p(A_i)}{p(A_i)} = 1$$

sehingga

$$H(\mathfrak{A},\mathfrak{B}) = -\sum_{i=1}^{n} p(A_i) \sum_{j=1}^{m} p(B_j|A_i) \log p(B_j|A_i)$$

$$-\sum_{i=1}^{n} p(A_i) \log p(A_i) \cdot 1$$

$$= -\sum_{i=1}^{n} p(A_i) \sum_{j=1}^{m} p(B_j|A_i) \log p(B_j|A_i) + H(\mathfrak{A}).$$

..... *3.4.1*

Jumlahan

$$-\sum_{j=1}^{m} p(B_j|A_i) \log p(B_j|A_i)$$

merupakan bentuk dari suatu entropy, dan diberikan sebagai ketidakpastian pada $\mathfrak{B}$ jika diketahui A_i yang muncul.

Dengan demikian Entropy Bersyarat dari $\mathfrak{B}$ jika diketahui peristiwa A_i terjadi

$$H(\mathfrak{B}|A_i) = -\sum_{j=1}^{m} p(B_j|A_i) \log p(B_j|A_i) \qquad \text{..... } 3.4.2$$

Secara langsung dapat dilihat bahwa $H(\mathfrak{B}|A_i)$ merupakan variabel acak dalam bagan ketidakpastian $\mathfrak{A}$. Dan ekspektasi dari variabel acak $H(\mathfrak{B}|A_i)$ dengan probabilitas $p(A_i)$

$$p(A_1)H(\mathfrak{B}|A_1) + p(A_2)H(\mathfrak{B}|A_2) + \ldots + p(A_n)H(\mathfrak{B}|A_n)$$

$$= \sum_{i=1}^{n} p(A_i)H(\mathfrak{B}|A_i)$$

$$= H(\mathfrak{B}|\mathfrak{A}) \qquad \text{..... } 3.4.3$$

adalah Entropy Bersyarat $\mathfrak{B}$ setelah $\mathfrak{A}$ dirata - rata, atau jumlah rata-rata dari ketidakpastian dalam $\mathfrak{B}$ ketika diketahui mana dari peristiwa A_i partisi $\mathfrak{A}$ yang telah terjadi. Jika hasil akhir di atas (*3.4.3*) dimasukkan dalam pernyataan bagi $H(\mathfrak{A},\mathfrak{B})$ pada (*3.4.1*) diperoleh

$$H(\mathfrak{A},\mathfrak{B}) = H(\mathfrak{B}|\mathfrak{A}) + H(\mathfrak{A}) \qquad \text{..... } 3.4.4$$

Karena pernyataan $H(\mathfrak{A},\mathfrak{B})$ simetris terhadap $\mathfrak{A}$ dab $\mathfrak{B}$, dengan cara yang sama dapat diperoleh

$$H(\mathfrak{A},\mathfrak{B}) = H(\mathfrak{A}|\mathfrak{B}) + H(\mathfrak{B}).$$

Contoh 3.4.1 :

Diambil sebuah dadu sebagai sumber informasi, dan diberikan dua partisi dari ruang sampel Ω

$$\mathfrak{A} = \begin{bmatrix} \{1,2\} & \{3\} & \{4,5,6\} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

dan

$$\mathfrak{B} = \begin{bmatrix} \{1\} & \{2\} & \{3\} & \{4\} & \{5\} & \{6\} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}.$$

Dari dua bagan di atas diperoleh tabel $p(B_j|A_i)$

$\mathfrak{B} \backslash \mathfrak{A}$	{1,2}	{3}	{4,5,6}
{1}	$\frac{1}{2}$	0	0
{2}	$\frac{1}{2}$	0	0
{3}	0	1	0
{4}	0	0	$\frac{1}{3}$
{5}	0	0	$\frac{1}{3}$
{6}	0	0	$\frac{1}{3}$

Entropy bersyarat dari $\mathfrak{B}$ jika diketahui $A_1 = \{1,2\}$ yang terjadi

$$H(\mathfrak{B}|A_1) = -\sum_{j=1}^{6} p(B_j|A_1) \log_2 p(B_j|A_1)$$

$$= -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} + 0 + 0 + 0 + 0$$

$= 0{,}99998$ bit per elemen peristiwa.

$$H(\mathfrak{B}|A_2) = -\sum_{j=1}^{6} p(B_j|A_2) \log_2 p(B_j|A_2)$$

$$= 0 + 0 - 1 \log_2 1 + 0 + 0 + 0$$

$= 0$ bit per elemen peristiwa.

$$H(\mathfrak{B}|A_3) = -\sum_{j=1}^{6} p(B_j|A_3) \log_2 p(B_j|A_3)$$

$$= 0 + 0 + 0 - \frac{1}{3}\log_2 \frac{1}{3} - \frac{1}{3}\log_2 \frac{1}{3} - \frac{1}{3}\log_2 \frac{1}{3}$$

$= 1{,}58481$ bit per elemen peristiwa.

Maka entropy bersyarat $\mathfrak{B}$ jika diketahui $\mathfrak{A}$ merupakan ekspektasi dari bagan ketidakpastian $\mathfrak{A}$ dengan variabel acak $H(\mathfrak{B}|A_i)$

$$H(\mathfrak{B}|\mathfrak{A}) = \sum_{i=1}^{3} p(A_i)H(\mathfrak{B}|A_i)$$

$$= \frac{1}{3}(0{,}99998) + \frac{1}{6}.0 + \frac{1}{2}(1{,}58481)$$

$= 1{,}12573$ bit per elemen peristiwa.

Teorema 3.4.1 :

Diberikan dua partisi sebarang. Entropy dalam suatu partisi dimana padanya diketahui peristiwa dari partisi lain terjadi, tidak akan pernah melebihi besarnya entropy dari partisi tersebut. Untuk sebarang partisi sebarang $\mathfrak{A}$ dan $\mathfrak{B}$,

$$H(\mathfrak{A}|\mathfrak{B}) \leq H(\mathfrak{A})$$

Bukti :

Karena $H(\mathfrak{A}|\mathfrak{B}) = H(\mathfrak{A},\mathfrak{B}) - H(\mathfrak{B})$ maka teorema di atas akan dibuktikan bahwa

$$H(\mathfrak{A},\mathfrak{B}) - H(\mathfrak{B}) \leq H(\mathfrak{A}).$$

Sebagaimana diketahui

$$H(\mathfrak{A}) = -\sum_i p(A_i)\log p(A_i).$$

Karena

$$p(A_i) = p(A_i,\Omega) = p(A_i,\bigcup_j B_j) = \sum_j p(A_i,B_j)$$

maka

$$H(\mathfrak{A}) = -\sum_i \sum_j p(A_i,B_j)\log p(A_i).$$

Dengan cara yang sama

$$H(\mathfrak{B}) = -\sum_i \sum_j p(A_i,B_j)\log q(B_j).$$

Jika $H(\mathfrak{A})$ dan $H(\mathfrak{B})$ dijumlahkan

$$H(\mathfrak{A}) + H(\mathfrak{B}) = -\sum_i \sum_j p(A_i,B_j)\left\{\log p(A_i) + \log q(B_j)\right\}$$

$$= -\sum_i \sum_j p(A_i,B_j)\log p(A_i)q(B_j).$$

Sementara itu

$$H(\mathfrak{A},\mathfrak{B}) = -\sum_i \sum_j p(A_i,B_j)\log p(A_i,B_j).$$

Karena

$$-\sum_i \sum_j p(A_i,B_j) = 1 \quad \text{dan} \quad -\sum_i \sum_j p(A_i)q(B_j) = 1,$$

maka menggunakan pertidaksamaan berguna *(2.4.3)* dengan mengambil $a_i = p(A_i,B_j)$ dan $b_i = p(A_i)q(B_j)$

$$-\sum_i p(A_i,B_j)\log p(A_i,B_j) \le -\sum_i p(A_i,B_j)\log p(A_i)q(B_j).$$

Bila pertidaksamaan di atas dijumlahkan untuk seluruh j

$$-\sum_i \sum_j p(A_i,B_j)\log p(A_i,B_j) \le -\sum_i \sum_j p(A_i,B_j)\log p(A_i)q(B_j)$$

$$H(\mathfrak{A},\mathfrak{B}) \le H(\mathfrak{A}) + H(\mathfrak{B})$$

$$H(\mathfrak{A},\mathfrak{B}) - H(\mathfrak{B}) \le H(\mathfrak{A}).$$

Dengan demikian terbukti bahwa

$$H(\mathfrak{A}|\mathfrak{B}) \le H(\mathfrak{A}).$$

■

Teorema 3.4.2 :

Jika dua partisi $\mathfrak{A}$ dan $\mathfrak{B}$ saling independen, maka

$$H(\mathfrak{A},\mathfrak{B}) = H(\mathfrak{A}) + H(\mathfrak{B})$$

Bukti :

Dua partisi $\mathfrak{A} = [A_i]$ dan $\mathfrak{B} = [B_j]$ disebut saling independen bila peristiwa-peristiwa A_i dan B_j saling independen untuk setiap i dan j. Dengan demikian berlaku

$$p(A_i,B_j) = p(A_i)q(B_j).$$

Sehingga

$$\begin{aligned} H(\mathfrak{A},\mathfrak{B}) &= -\sum_i \sum_j p(A_i,B_j)\log p(A_i,B_j) \\ &= -\sum_i \sum_j p(A_i)q(B_j)\log p(A_i)q(B_j) \\ &= -\sum_i \sum_j p(A_i)q(B_j)\left(\log p(A_i) + \log q(B_j)\right) \\ &= -\sum_i p(A_i)\log p(A_i)\cdot\sum_j q(B_j) \\ &\qquad - \sum_j q(B_j)\log q(B_j)\cdot\sum_i p(A_i) \\ &= -\sum_i p(A_i)\log p(A_i)\cdot 1 - \sum_j q(B_j)\log q(B_j)\cdot 1 \end{aligned}$$

$$H(\mathfrak{A},\mathfrak{B}) = H(\mathfrak{A}) + H(\mathfrak{B}).$$

■