ENTROPIJA

- Neku situaciju ćemo označiti kao neodredjenost ukoliko postoji više mogućnosti, pri čemu nije poznato koja će od njih biti realizovana ili izabrana.
- Zapravo realne situacije potpune odredjenosti u kojima se ishodi mogu predvidjati sa sigurnošću su više izuzetak nego pravilo.

- Uvedimo na formalniji način pojam neodredjenih situacija, koji je početni gradivni blok savremene teorije informacija.
- Pretpostavimo neku situaciju *S* u kojoj je moguće *m* različitih ishoda.
- Označimo ove mogućnosti sa $e_1, e_2, \dots, e_m$.
- Ovaj skup mogućnosti ćemo označiti sa $S(e_1, e_2, \dots, e_m)$, i nazvati konačnom šemom izbora.

- Po nekom mehanizmu ili na osnovu nekog procesa ili na osnovu nečije volje vrši se izbor jedne od ovih mogućnosti.
- Neodredjenost asocirana zadatoj šemi izbora S, nastaje usled našeg neznanja ili nemogućnosti da znamo koja će od m alternativa biti izabrana.
- Na koji način možemo izmeriti količinu neodredjenosti u datoj šemi izbora S?

- Intuitivno predosećamo da što je veća kardinalnost $|S|$ skupa S, odnosno broj elementa ovog skupa, veća je i neodredjenost.
- Zašto onda ne bi smo jednostavno uzeli $|S|$ kao meru neodredjenosti?
- Iako je to u principu moguće, jedan drugi pristup će se pokazati delotvornijim. Predstavimo sebi sledeću igru: zamisli se jedna od mogućnusti iz skupa $S(e_1, e_2, \dots, e_m)$, a zatim se na osnovu pitanja koja su tako formulisana da se na njih odgovara samo sa da/ne, pokušava pogoditi zamišljena vrednost.

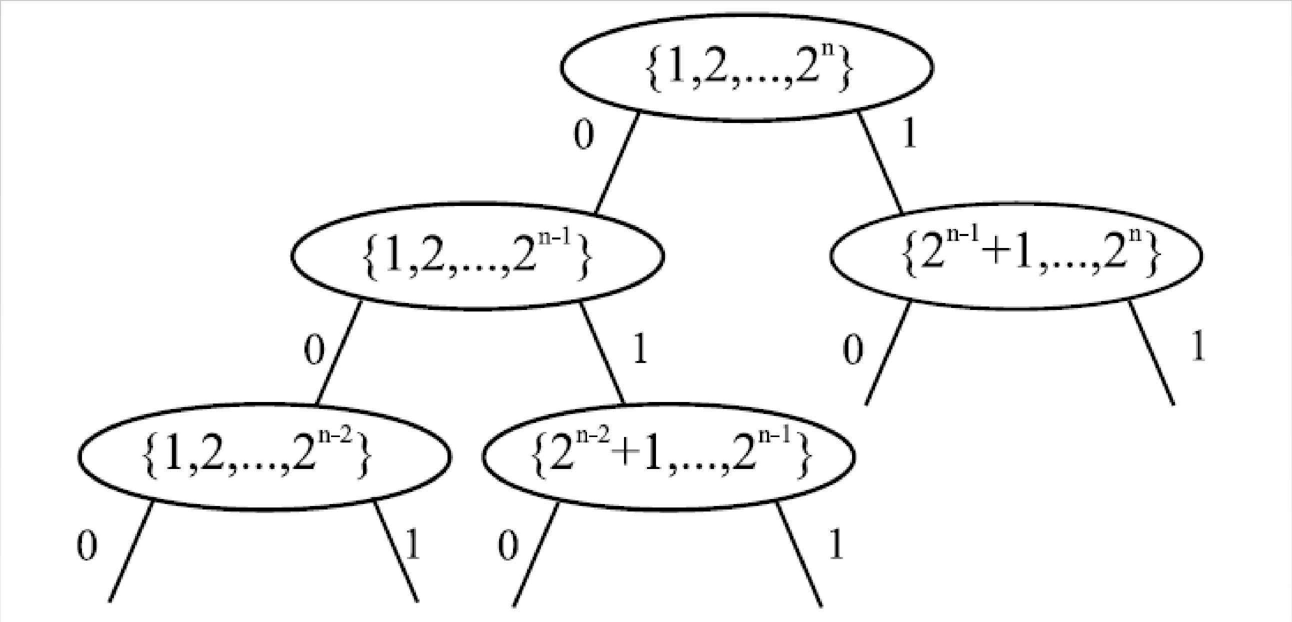
- Mogućnosti uvek možemo označiti redom brojevima $1,2,\ldots,m = |S|$, tako da pitanja mogu biti tipično sledećeg oblika: Da li je zamišljeni broj paran? Da li je zamišljeni broj manji od 20? Da li je zamišljeni broj veći od 12? i td.
- Odmah zapažamo da ukoliko nam je potrebno više pitanja, neodredjenost šeme izbora S je veća.
- Stoga je osnovano smatrati da je broj pitanja tipa da/ne neophodnih za odredjivanje nepoznate zamišljene mogućnosti iz skupa S, dobra mera neodredjenosti šeme izbora S.

Univerzitet Sinergija

- Od važnosti je i kakva pitanja postavljamo. Ako su pitanja tipa da li je zamišljeni broj jednak 1, pa ako nije da li je jednak 2, pa ako nije da li je jednak 3, jasno je da nas to vodI ka postavljanju svih m mogućih pitanja.
- Ovo sigurno nije optimalan način dobijanja tačnog odgovora u najmanjem broju koraka.
- Medjutim ako je prvo pitanje da li je zamišljeni broj manji od m/2, u narednom pitanju sužavamo pretragu na polovinu početnih mogućnosti.

- Ako i u narednim koracima nastavimo da postavljamo pitanja na ovaj način, pokazaće se da je izabrana strategija vrlo efikasna.
- U cilju detaljnijeg upoznavanja sa ovim postupkom, pretpostavimo prvo da je ukupan broj početnih mogućnosti stepen broja 2, odnosno da važi $m = 2^n$.
- Tada možemo podeliti S postavljanjem prvog pitanja (da li je zamišljeni broj veći od 2^{n-1} ?) na dva skupa iste veličine: $\{1,2, \dots , 2^{n-1}\}$ i $\{2^{n-1} + 1, \dots , 2^n\}$.

- Svaka od ovih polovina će bit podeljena na dve polovine narednim pitanjem.
- Ako je odgovor na prvo pitanje - ne, tada sledeće pitanje generiše skupove $\{1,2,\ldots,2^{n-2}\}$ i $\{2^{n-2}+1,\ldots,2^{n-1}\}$.
- Ako je odgovor na prvo pitanje – da, tada naredno pitanje pravi razliku izmedju skupova $\{2^{n-1}+1,\ldots,2^{n-1}+2^{n-2}\}$ i $\{2^{n-1}+2^{n-2}+1,\ldots,2^{n}\}$.
- Prikazani proces postavljanja pitanja i dobijanja odgovora prikazan je odgovarajućom grafskom strukturom na sl.1.1.



Sl.1.1 Stablo pitanje odgovor nastalo tokom binarne pretrage za nepoznatom mogućnošću od ukupno $m = 2^n$ mogućnosti.

- Svakom pitanju odgovara jedan čvor, označen skupom mogućnosti identifikovanih do tog momenta.
- Prvi čvor je stoga označen skupom svih mogućnosti S.
- Čvorovi u prvom sloju su označeni sa dva skupa upola manje dimenzije, drugi sloj čvorova je označen sa četiri skupa dimenzija četiri puta manjih od početnog skupa S i td.
- Odgovori se označavaju granama. Odgovor – ne označavamo sa 0, dok odgovor – da označavamo sa 1.

- Proces sukcesivnog delenja početnog skupa mogućnosti na polovine, završava se nakon n koraka sa ukazivanjem na tačan odgovor.
- Broj n nije ništa drugo do logaritam od $m = 2^n$ za osnovu 2, odnosno $n = \log_2 m$. Ovaj broj je značajno manji od m.
- Stoga izgleda razumno predstaviti količinu neodredjenosti u šemi izbora S pomoću logaritma za osnovu dva broja svih mogućnosti $|S| = 2^n$ tog sistema.
- Ako meru neodredjenosti šeme izbora S označimo sa $h(|S|)$, tada je izbor $h(|S|) = \log|S|$, dobar izbor barem kada je $|S|$ stepena 2.

- Stoga, konačno usvajamo sledeću definiciju:
- **Definicija 1.1** *Količina neodredjenosti šeme izbora.* Za šemu izbora S sa $|S|$ mogućnosti, količina neodredjenosti $h(|S|)$ je definisana sa
- $h(|S|) = \log|S|$.

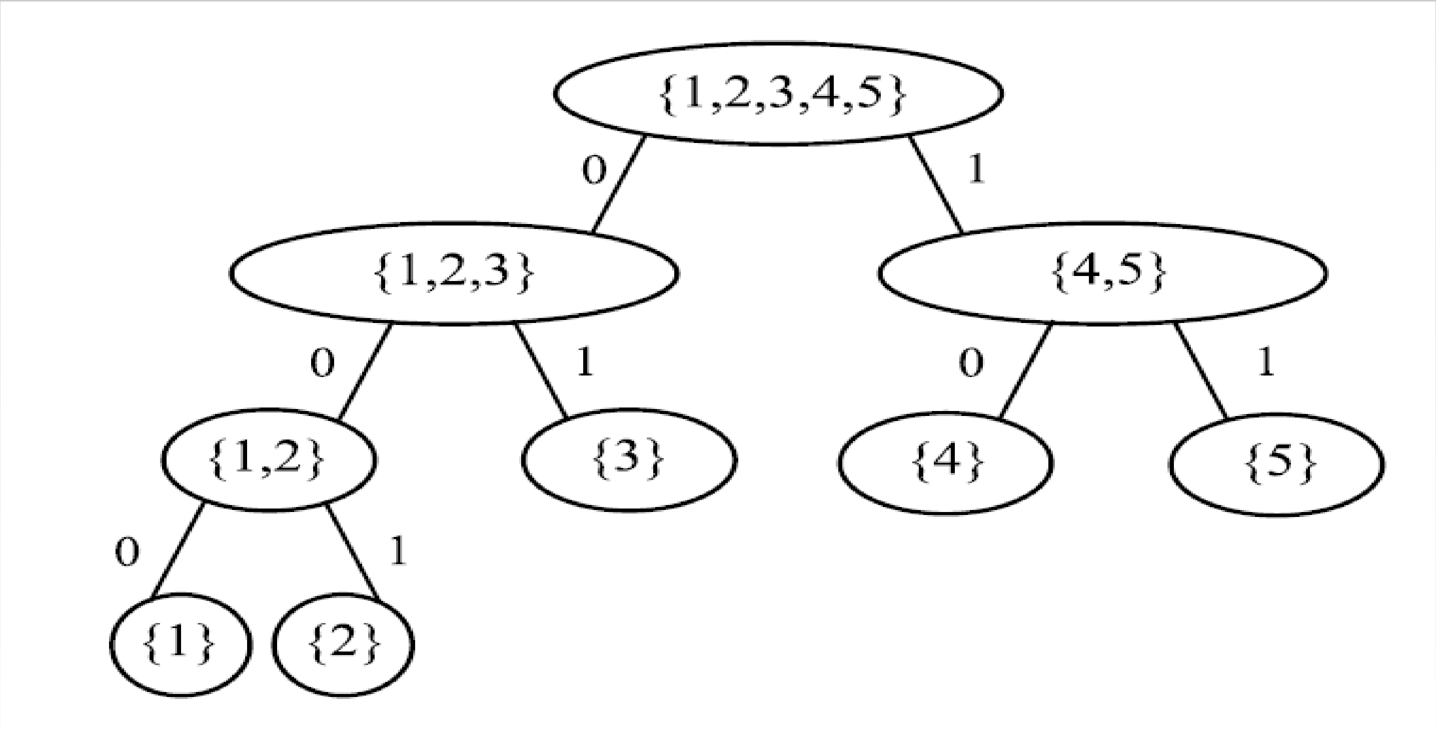
- **Primer 1.1** (Šahovska tabla) Neka je data prazna šahovska table. Postoji ukupno $m = 64 = 2^6$ mogućnosti za postavljanje prve figure. Stoga se odgovarajuća šema izbora može prikazati sa $S = \{1,2, \dots, 64\}$, gde je značenje svakog broja jedno polje na šahovskoj tabli. Količina neodredjenosti u postavljanju jedne figure na praznu šahovsku tablu je $h(|S|) = \log|64| = 6$.

- Kao i svaka druga definicija, i naša definicija mere neodredjenosti je do nekle proizvoljna.
- Opravdanja dobrih definicija se stiču njihovim svojstvima korisnosti i elegancije u praktičnim primenama i olakšavanju teorijskih izvodjenja.
- Budući da je ova definicija mere neodredjenosti preživela prvo stoleće teorije informacija, očigledno da je zadovoljila oba ova kriterijuma.

- Napomenimo da pitanja možemo postavljati i tako da postoji više od dva moguća odgovora. Pretpostavimo da pitanja imaju $k > 2$ mogućih odgovora. U tom slučaju svako pitanje vrši podelu skupa od m mogućnosti u k disjunktnih podskupova (presek prazan skup) sa približno $\frac{m}{k}$ elemenata. Ako je $|S| = k^n$, tada nam je potrebno tačno $n = \log_k|S|$ pitanja. Ako ovog puta usvojimo logaritam za osnovu k, tada je
- $h(|S|) = \log_k|S|$.
- Imajući u vidu da je
- $\log_k|S| = \log_k 2 \cdot \log_2|S|$,
- Vidimo da je promena osnove logaritma po efektima ekvivalentna promeni jedinice merenja i stoga nije od suštinskog značaja.

- Definicijom sistema izbora S i odgovarajućeg stabla pitanja, kao što je ono na sl.1.1, istovremeno je odredjena i sistem kodovanja svih mogućnosti sistema izbora. Kod jedne mogućnosti dobijamo spajanjem simbola 0 i 1 na putu od korena stabla do date mogućnosti. Ako su upućena binarna pitanja, tada dobijamo binarni kod za dati sistem mogućnosti. Primetimo da je dužina koda svake mogućnosti jednaka ili sledećem najmanjem ili sledećem najvećem celom broju od $h(|S|) = \log(|S|)$. Ovim se prvi put susrećemo sa bliskom vezom izmedju mere neodredjenosti i kodovanja.

- Primer 1.2 Binarno stablo pitanja. Neka je sistem mogućnosti data sa $S = \{1,2,3,4,5\}$. Njegova količina neodredjenosti je
- $h(|S|) = \log 5 \approx 2{,}3219$ bit.
- Jedno moguće stablo pitanja je dato na sl.1.3. Vidimo da kod 001 reprezentuje mogućnost $\{2\}$, pri čemu je njegova dužina 3 jednaka narednom najvećem celom broju od $h(|S|)$. Mogućnost $\{3\}$ ima kod dužine 2, $\{4\}$ima kod dužine 2 i td.



Sl.1.3. Stablo pitanje-odgovor nastalo na osnovu binarnih pitanja u sistemu sa 5 mogućnosti.

Nabrojmo neke najjednostavnije osobine mere neodredjenosti $h(|S|)$:

1. Ako su S_1 i S_2 dva sistema mogućnosti i $|S_1| = |S_2|$, tada je i $h(|S_1|) = h(|S_2|)$. Prema tome uticaja na neodredjenost ima samo broj mogućnosti a ne i njihova priroda.
2. Ako su S_1 i S_2 dva sistema mogućnosti i $|S_1| < |S_2|$, tada je i $h(|S_1|) < h(|S_2|)$, budući da je logaritam neopadajuća funkcija. Ova osobina je u skladu sa našim intuitivnim očekivanjima: neodredjenost raste sa porastom broja mogućnosti izbora.

1. Ako je S sistem samo sa dve mogućnosti, tada ako za osnovu logaritma uzmemo osnovu 2, tada je $h(|S|) = \log_2 2 = 1$. Ova jedinica mere se naziva bit (jedinica binarne informacije). Vidimo da je neodredjenost tesno povezana sa informacijom i da se mogu izraziti u istim mernim jedinicama. Budući da je u računarima informacija prirodno zapisana pomoću dva stanja, najpopularniji tip pitanja je binarni.

Entropija-Primer 1.4

- Neka se kocka baca m puta i neka se podrazumeva da su bacanja nezavisna. Ovaj proces se može modelovati pomoću m nezavisnih sistema izbora $S_1, \ldots\ldots, S_m$, pri čemu svaki sistem ima 6 mogućnosti. Prema (1.1), ukupna neodredjenost ishoda prilikom bacanja kocke m uzastopnih puta je

- $h(|S_1 \times S_2 \times \cdots \times S_m|) = h(|S_1|) + h(|S_2|) + \cdots + h(|S_m|) = m \cdot \log 6.$

Entropija-Rezime

- Formalizovali smo situacije neodredjenosti uvodjenjem sistema S, u kome se bira jedna od mogućnosti, ali se ne zna koja.
- Neodredjenost pridružena sistemu mogućnosti se meri najmanjim brojem pitanja koja treba postaviti u cilju nalaženja stvarnog izbora. Na ovaj način smo došli do mere neodredjenosti izražene sa $\log|S|$. Ako su upotrebljena binarna pitanja, jedinica mere neodredjenosti je bit.
- Sistemom postavljanja pitanja u vezi odredjivanja date nepoznate mogućnosti, indukuje se stablo na osnovu koga se može jednoznačno odrediti kod svake od mogućnosti sistema S. Dužine kodova su približno jednake meri neodredjenosti $\log|S|$. Ako su korišćena binarna pitanja, dobija se binarni kod.
- Neodredjenosti sistema nezavisnih mogućnosti se medjusobno sabiraju.

Entropija-izbori mogućnosti sa poznatim verovatnoćama

- Postoje situacije u kojima su nam poznate verovatnoće sa kojima se pojedine mogućnosti pojavljuju u nekom posmatranom sistemu mogućnosti S. Npr. ako neko kuca tekst na srpskom jeziku, pojedini znaci se pojavljuju češće od nekih drugih.
- Uvedimo formalno verovatnoće u sistem mogućnosti $S = \{e_1, e_2, \ldots, e_m\}$ dodelom verovatnoća p_i mogućnostima e_i za $i = 1,2, \ldots, m$.
- Sistem izbora S zajedno sa skupom verovatnoća $P = \{p_1, p_2, \ldots, p_m\}$ obrazuje verovatnosni sistem izbora. Formalno, definišemo ga na sledeći način

- DEFINICIJA 1.2 Ako je S system izbora a P skup verovatnoća nad S, koji zadovoljava uslov (1.2), tada se par (S,P) naziva verovatnosni sistem izbora.

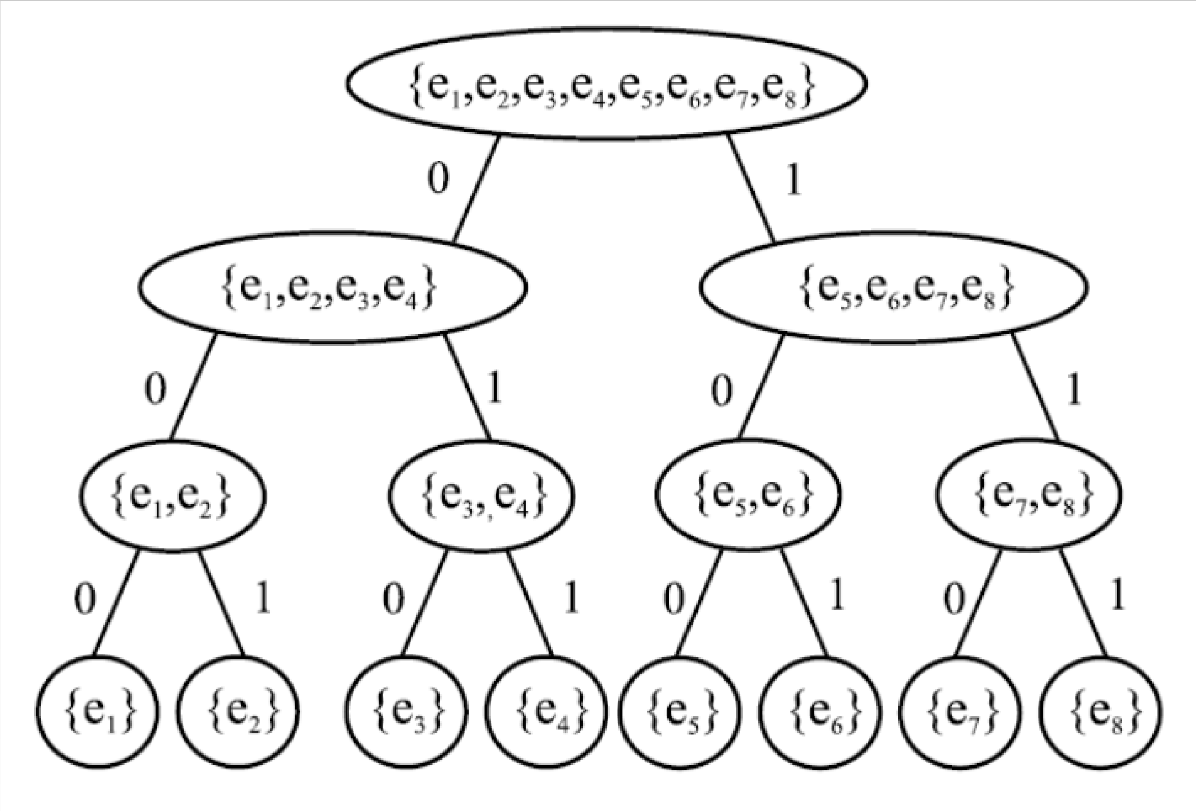
- Kolika je količina neodredjenosti u verovatnosnom sistemu izbora? Možemo pristupiti na isti način postavljanjem niza odgovarajućih pitanja. Medjutim, sada bi podela na skupove jednakih veličina bila neoptimalna strategija, jer njome ne bi bile obuhvaćene verovatnoće kojima raspolažemo.
- Pretpostavimo npr. da je jedna mogućnost, recimo e_1 mnogo verovatnija od svih drugih. Razumno je odmah prvo pitati da li se mogućnost e_1 realizovala, jer sa velikom verovatnoćom postoji šansa da smo pogodili samo na osnovu jednog pitanja. Samo ako je odgovor “ne” nastavlja se sa pitanjima.

Univerzitet Sinergija

Entropija – primer 1.5

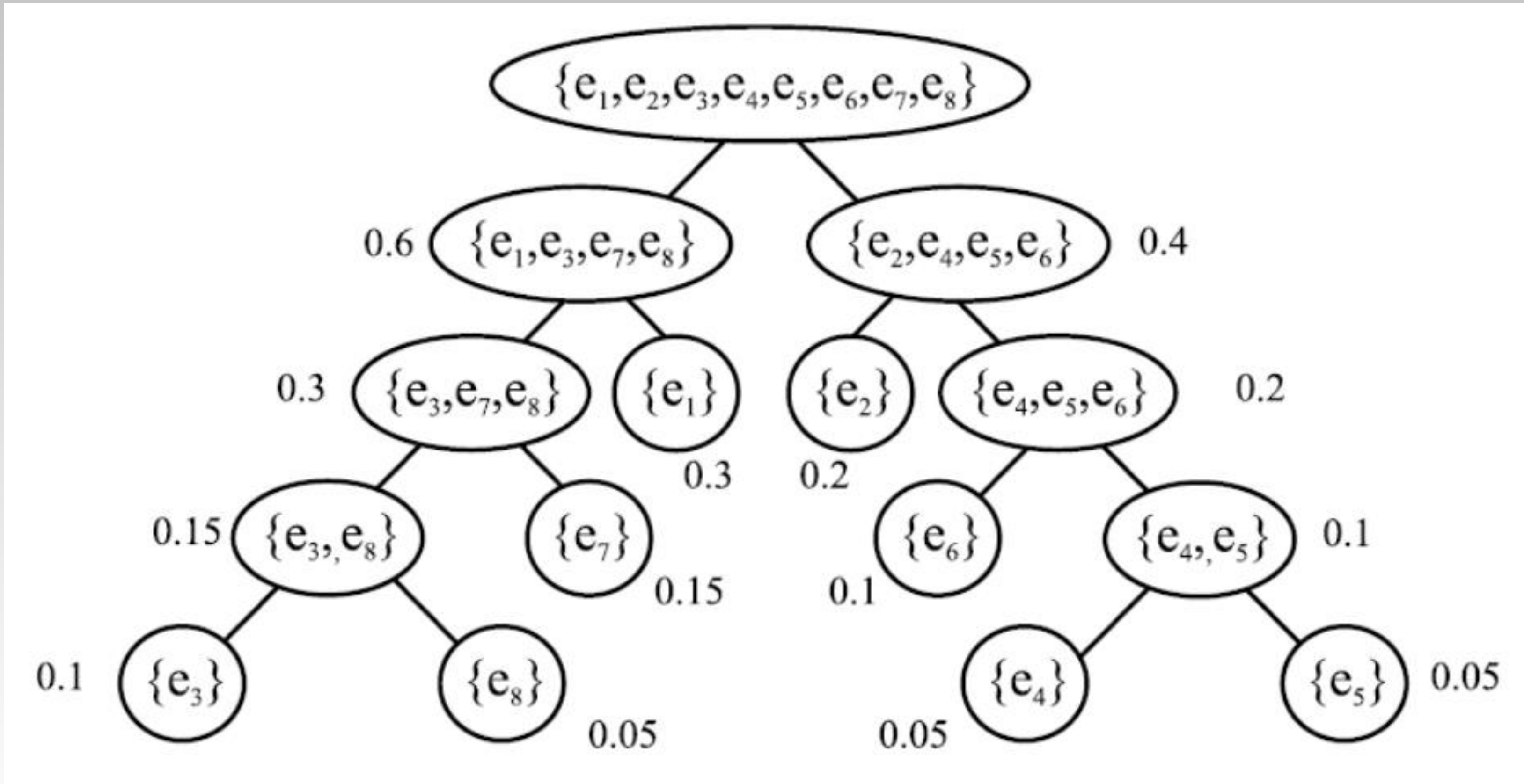
- Primer 1.5 Neka je dat verovatnosni sistem izbora (S,P) sa $S = \{e_1, e_2, \dots, e_8\}$ i $P = \{0.3, 0.2, 0.1, 0.05, 0.05, 0.1, 0.15, 0.05\}$. Odgovarajuće binarno stablo prikazano je na sl.1.4, dok je alternativno stablo koje uzima u obzir verovatnoće, prikazano je na sl.1.5. Jednostavnim računom dolazimo do podatka da je očekivani broj pitanja u slučaju binarnog stabla 3, a u slučaju stabla koje uzima u obzir verovatnoće 2.75.

Primer 1.5



Sl.1.4 Binarno stablo iz primera 1.5 sa očekivanim brojem pitanja 3

Primer 1.5



Sl.1.5. Alternativno stablo iz primera 1.5 kada su u obzir uzete verovatnoće pojavljivanja svih mogućnosti. Očekivani broj pitanja je 2.75.

Definicija entropije

- Minimizacija očekivane vrednosti broja postavljenih pitanja nije trivijalan zadatak. Rešenje ovog problema je poznato i veoma se koristi u teoriji kodovanja. Ključna ideja ovog postupka je da se skup svih mogućnosti ne deli na podskupove jednake kardinalnosti, već jednakih verovatnoća. Iz teorije kodovanja je poznato da je očekivani broj pitanja aproksimativno iznosi
- $-\sum_{i=1}^{m} p_i \log p_i$.

  Ova veličina se naziva entropija, i predstavlja meru neodredjenosti količine neodredjenosti verovatnosnih sistema mogućnosti.

Definicija entropije

- DEFINICIJA 1.3 Neka je (S,P) verovatnosni sistem mogućnosti. Tada je količina neodredjenosti u njemu definisana entropijom

$$H(P) = -\sum_{i=1}^{m} p_i \log p_i \ .$$

U definiciji (1.3) postoji odredjeni stepen proizvoljnosti, budući da nije definisana osnova logaritma. Baza k logaritma odgovara broju mogućih odgovora u eksperimentu postavljanja pitanja i dobijanja odgovora. Promena baze menja samo jedinicu mere. Binarni slučaj je najpopularniji i njemu odgovara jedinica bit.

Primer 1.6

- Primer 1.6 Izračunajmo količinu neodredjenosti za verovatnosni sistem neodredjenosti iz primera 1.5.

  - $H(P) = -\sum_{i=1}^{8} p_i \log p_i =$

$$-0.3 \cdot \log 0.3 - 0.2 \cdot \log 0.2 - 0.2 \cdot \log 0.1 - 0.15 \cdot \log 0.05 - 0.15 \cdot \log 0.15 \approx 2.7087$$

bita.

- H(P) je manje od očekivane vrednosti dužine koda u boljem stablu pitanja sa sl.1.5.

Svojstva entropije

- Budući da je entropija fundamentalan pojam teorije informacija, navešćemo njene osnovne karakteristike. U buduće ćemo za sistem mogućnosti nad verovatnoćama $P = \{p_1,\ p_2, \dots, p_m\}$ označavati entropiju tog sistema sa $H(P) = H(p_1,\ p_2, \dots, p_m)$.
- Prvo ćemo povezati opšti pojam entropije sa pojmom mere neodredjenosti za sisteme mogućnosti u kojima nisu definisane verovatnoće pojedinih izbora. Ako je P uniformno raspodeljeno po svih n mogućnosti, tada je
- $H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = -\sum_{i=1}^{n} \frac{1}{n} \cdot \log \frac{1}{n} = -\log \frac{1}{n} = \log n = h(n)$.

Svojstva entropije

- Osobina 1.1 Entropija uniformne raspodele verovatnoća n mogućnosti jednaka je meri neodredjenosti odgovarajućeg sistema mogućnosti bez definisanih verovatnoća mogućih ishoda.
- Ovaj rezultat je u tesnoj vezi sa poznatim Laplasovim principom nedovoljnog razloga, koji tvrdi da ukoliko ne znamo ništa odredjeno o nekoj pojavi, razumno je pretpostaviti jednake verovatnoćemogućih ishoda.
- U tom pogledu entropija je primenljiva kao mera neodredjenosti i na one seisteme mogućnosti u kojima nam nisu poznate verovatnoće pojedinih ishoda.

Svojstva entropije

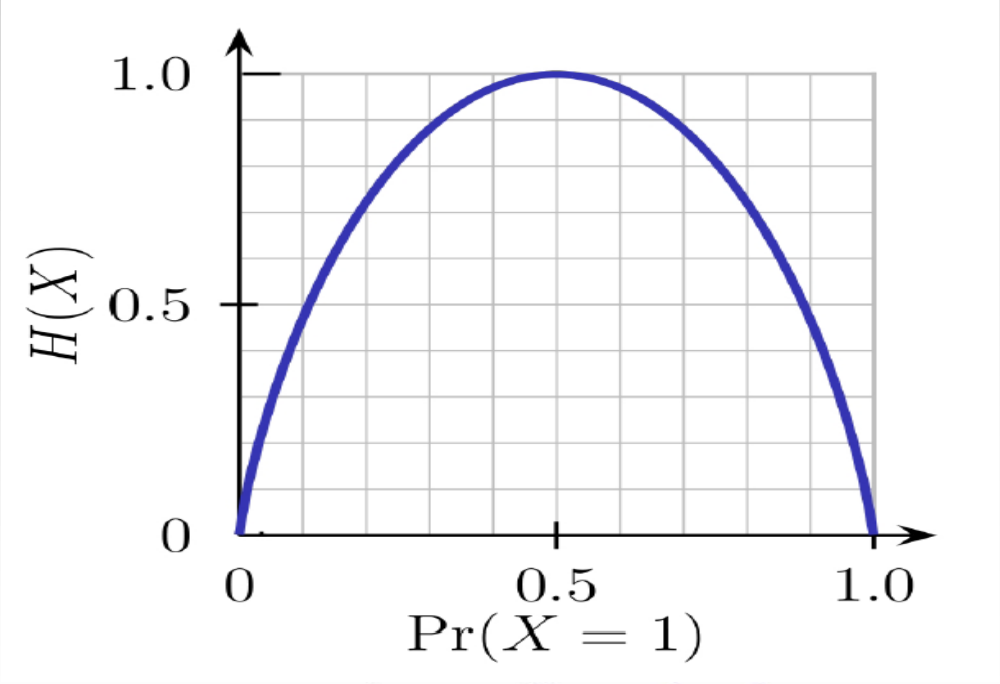
- Za zadati sistem mogućnosti S u kome je $|S| = n$, imamo maksimalnu neodredjenost ukoliko ne poznajemo verovatnoće ishoda ili ako su one medjusobno jednake, odnosno važi
- $H(P) \leq h(S) = \log n$.

Važna svojstva entropije

1. Ako su (S_1, P_1) i (S_2, P_2) dva verovatnosna sistema mogućnosti za koje važi $|S_1| = |S_2|$ i $P_1 = P_2$, tada je i $H(P_1) = H(P_2)$. Ovim se potvrdjuje intuitivan stav da entropija zavisi samo od raspodele verovatnoće, a ne i od prirode mogućnosti e_i u nekom sistemu mogućnosti.
2. U važnosti je relacija
   - $H(p_1, p_2, \ldots, p_m) = H(p_1, p_2, \ldots, p_m, 0)$,

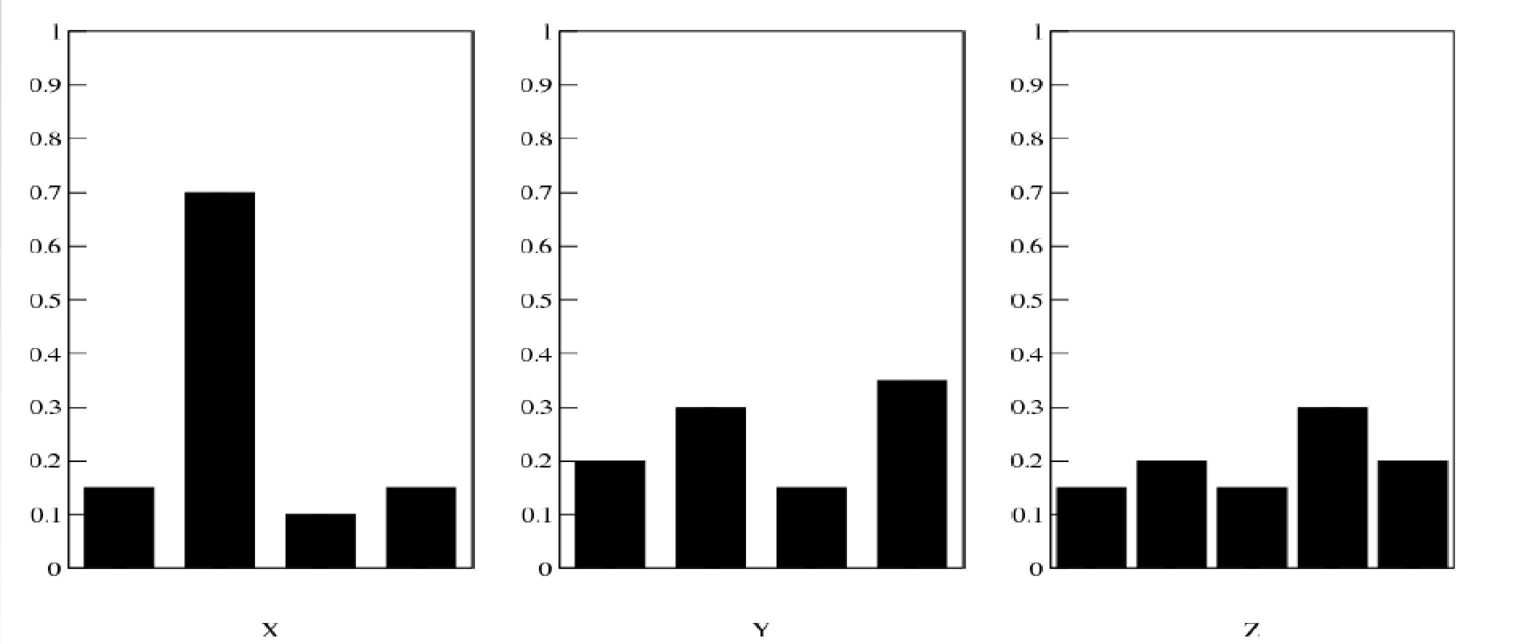
- što potiče od konvencije da je $0 \cdot \log 0 = 0$. Mogućnosti čija je verovatnoća nula ne utiču na ukupnu količinu neodredjenosti. To je sasvim u skladu sa našom intuicijom, pošto smo sigurni da se takve mogućnosti neće nikada realizovati.

Entropija slučajne varijable sa dva ishoda (X=1, X=0)



Univerzitet Sinergija

Primeri entropija različito raspodeljenih slučajnih veličina



$$H(X) < H(Y) < H(Z)$$

Uslovne entropije

- Neka je zadat verovatnosni sistem mogućnosti (S,P) i neka je ovom sistemu asocirana slučajna veličina X. Neka je observirano da se desio dogadjaj $E \subseteq S$. Kako poznavanje ove činjenice utiče na neodredjenost sistema? Ovaj dogadjaj indukuje novi verovatnosni sistem mogućnosti (E, P_E). Ovde je P_E uslovna verovatnoća

$$p_{X|E}(x) = \frac{p_X(x)}{p_X(E)}, \text{za svako } x \in E.$$

Uslovne entropije

- Neodredjenost vezana za ovaj novi sistem je data sa uslovnom entropijom slučajne veličine X kada je poznato E

$$H(X|E) = -\sum_{x \in E} p_{X|E}(x) \log p_{X|E}(x) \ .$$

Primer 1.8

- Neka je X slučajna veličina vezana za verovatnosni sistem mogućnosti (S,P), gde je $S = \{1,2,3,4\}$, $P = \{0.5, 0.25, 0.125, 0.125\}$ i $E = \{1,3\}$ je jedan dogadjaj. Tada je bezuslovna entropija data sa

$$\begin{aligned} H(X) &= -0.5 \log 0.5 - 0.25 \log 0.25 - 0.125 \log 0.125 - 0.125 \log 0.125 \\ &= 1.75 \text{bit}. \end{aligned}$$

- Kako je $p_X(E) = 0.625$, $p_{X|E}(1) = {p_X(1)}/{p_X(E)} = 0.8$, $p_{X|E}(3) = {p_X(3)}/{p_X(E)} = 0.2$, dobijamo

Primer 1.8

$$\begin{aligned} H(X|E) &= -p_{X|E}(1)\log p_{X|E}(1) - p_{X|E}(3)\log p_{X|E}(3) \\ &= -0.8\log 0.8 - 0.2\log 0.2 \\ &\approx 0.7219 \text{ bit.} \end{aligned}$$