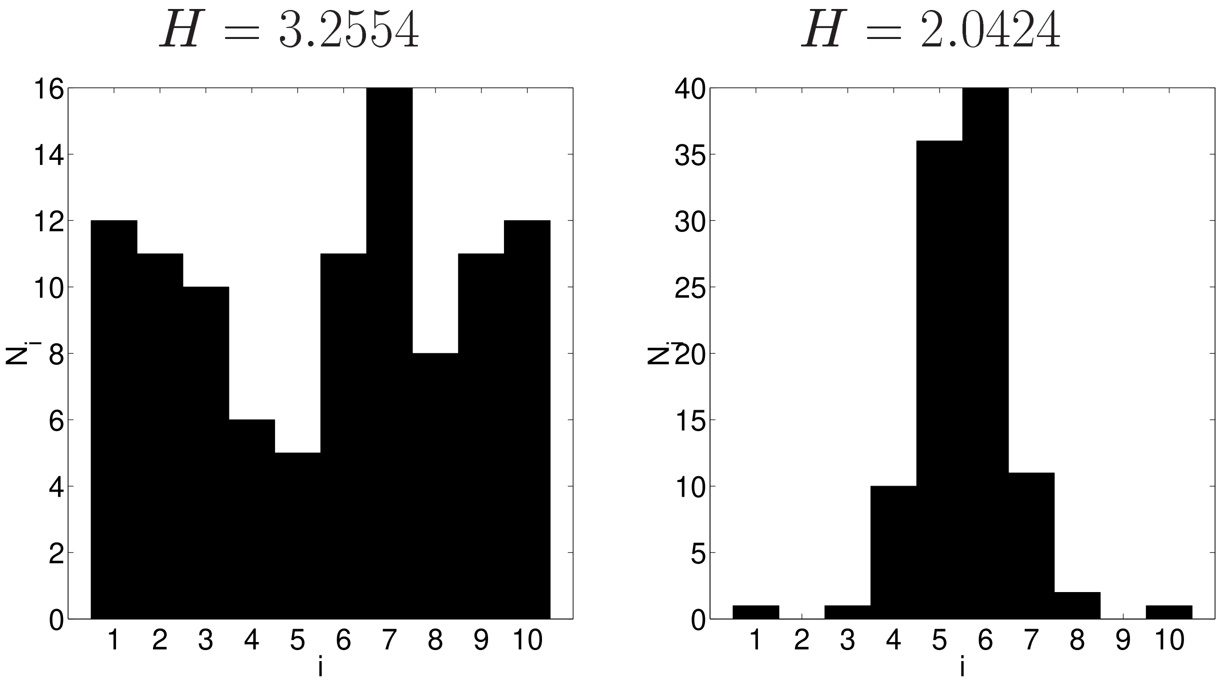
Entropie

1. Entropie jako míra neuspořádanosti.
2. Entropie jako míra informace.
3. Pokrytí množiny elementárních jevů.
4. Vlastnosti entropie.
5. Podmíněná entropie.
6. Vlastnosti podmíněné entropie.
7. Vzájemná informace.

Entropie jako míra neuspořádanosti

Jev x_i: kulička padla do přihrádky i



$N = 100$ stejných objektů: $\{b_1, b_2, \ldots b_N\}$

$N!$ všech permutací

$N_i!$ permutací v každé přihrádce

Celkový počet přerovnání objektů v histogramu je

$$W = \frac{N!}{\prod_{i=1}^{N} N_i!}$$

Entropie

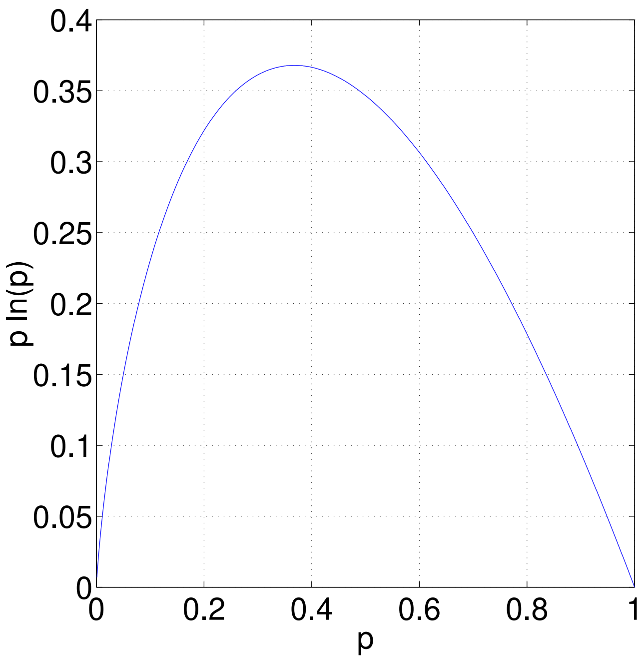
$$\lim_{N\to\infty} -\frac{1}{N}\ln W =$$

$$= \lim_{N\to\infty} -\sum_{i=1}^{m} \frac{N_i}{N} \ln \frac{N_i}{N} = -\sum_{i=1}^{m} p_i \ln p_i$$

Počet miktostavů (rozdělení do m přihrádek) které dávají za vznik stejnému makrostavu (histogram).

- Entopie je malá pro úzká rozdělení
- Entropie je velká pro široká rozdělení

$p \ln p$



Entropie jako míra informace

Hledáme funkci $s(p)$:

1. je spojitá
2. je monotonně klesající s p
3. $s(1) = 0$ (nulová nejistota)

$$s(p) = -s(\tfrac{1}{e}) \ln p$$

$$s(\tfrac{1}{e}) = 1$$ definujeme

Střední míra informace (v Natech) na množině jevů $\mathbf{x} = \{x_k\}$

$$H(\mathbf{x}) = -\sum_k p(x_k) \ln p(x_k)$$

Vlastnosti rozkladu množiny elementárních jevů na třídy ekvivalence

Definice

1. Elementární rozklad $\mathcal{E} = \{A_1, A_2, \dots, A_n\}$
2. Zjemnění rozkladu

$$\mathcal{B} \prec \mathcal{A} \quad \text{právě když} \quad B_j \subset A_i$$

3. Součin rozkladů $\mathcal{C} = \mathcal{A} \cdot \mathcal{B}$ je největší společné zjemnění
$$\mathcal{A} \cdot \mathcal{B} \prec \mathcal{A}, \quad \mathcal{A} \cdot \mathcal{B} \prec \mathcal{B}$$

Vlastnosti

1. $\mathcal{E} \prec \mathcal{A}$ pro každé $\mathcal{A}$
2. $\mathcal{A} \cdot \mathcal{B} = \mathcal{B} \cdot \mathcal{A}$
3. $\mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C}) = (\mathcal{A} \cdot \mathcal{B}) \cdot \mathcal{C}$
4. Jestliže $\mathcal{A}_1 \prec \mathcal{A}_2 \prec \mathcal{A}_3$ potom $\mathcal{A}_1 \prec \mathcal{A}_2$
5. Jestliže $\mathcal{B} \prec \mathcal{A}$ potom $\mathcal{A} \cdot \mathcal{B} = \mathcal{B}$

Entropie na rozkladu množiny elementárních jevů

Entropie

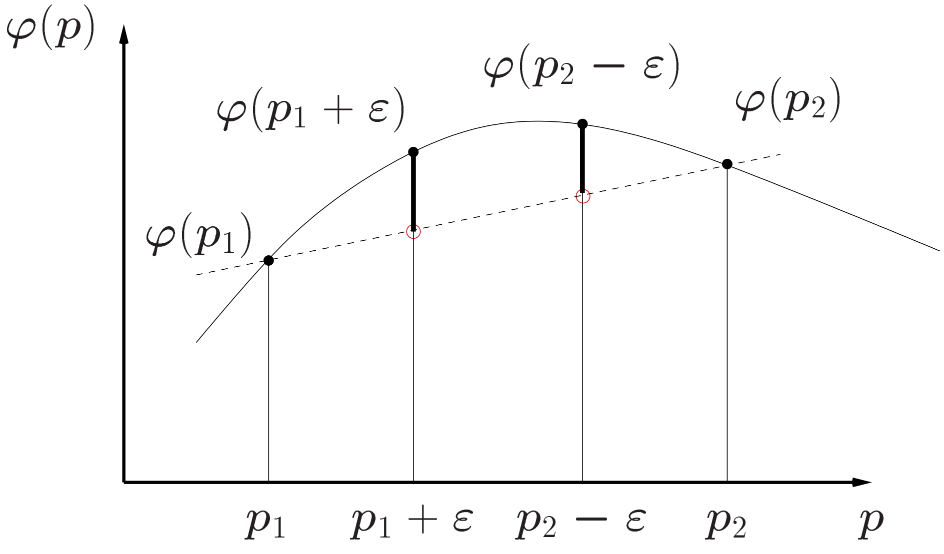
$$H(\mathcal{A}) = -\sum_{i=1}^{N} p_i \ln p_i = \sum_{i=1}^{N} \varphi(p_i)$$

$$p_i = P(A_i) \qquad \varphi(p_i) = -p_i \ln p_i$$

Pomocná věta:

$$\varphi(p_1 + p_2) \leq \varphi(p1) + \varphi(p_2) \leq \varphi(p_1 + \varepsilon) + \varphi(p_2 - \varepsilon)$$

pokud $\quad p_1 < p_1 + \varepsilon \leq p_2 - \varepsilon < p_2$



Vlastnosti

V1. Jestliže $\mathcal{B} \prec \mathcal{A}$ potom $H(\mathcal{B}) \geq H(\mathcal{A})$

Zjemněním rozkladu se entropie zvýší.

Pozn: zjemněním histogramu se entropie zvýší.

V2. Pro každé $\mathcal{A}$ platí $H(\mathcal{A}) \leq H(\mathcal{E})$

Entropie každého rozkladu je menší nebo rovna entropii elementárního rozkladu.

V3. Pro každé $\mathcal{A}$ a $\mathcal{B}$ platí

$$H(\mathcal{A} \cdot \mathcal{B}) \geq H(\mathcal{A}), \quad H(\mathcal{A} \cdot \mathcal{B}) \geq H(\mathcal{B}),$$

Pozn: $\mathcal{A} \cdot \mathcal{B} \prec \mathcal{A}, \quad \mathcal{A} \cdot \mathcal{B} \prec \mathcal{B}$

V4. Entropie $\mathcal{A}$ je maximální pokud všechny jeho prvky mají stejnou pravděpodobnost: $p_i = P(A_i) = p$

Podmíněná entropie

Def. Nechť rozklady $\mathcal{A}$ a $\mathcal{B}$ jsou

$$\mathcal{A} = [A_1, A_2, \dots, A_{N_A}]$$
$$\mathcal{B} = [B_1, B_2, \dots, B_{N_B}]$$

Potom podmíněná entropie rozkladu $\mathcal{A}$ za předpokladu, že nastal jev B_j je

$$H(\mathcal{A} \mid B_j) = -\sum_{i=1}^{N_A} P(A_i \mid B_j) \ln P(A_i \mid B_j)$$

Podmíněná entropie $H(\mathcal{A} \mid \mathcal{B})$ je pak střední hodnota

$$\begin{aligned} H(\mathcal{A} \mid \mathcal{B}) &= \sum_{j=1}^{N_B} P(B_j)\, H(\mathcal{A} \mid B_j) \\ &= -\sum_{j=1}^{N_B} P(B_j) \sum_{i=1}^{N_A} P(A_i \mid B_j) \ln P(A_i \mid B_j) \\ &= -\sum_i \sum_j P(A_i\, B_j) \ln P(A_i \mid B_j) \end{aligned}$$

Pozn: Střední nejistota o $\mathcal{A}$ je-li pozorováno $\mathcal{B}$ je $H(\mathcal{A} \mid \mathcal{B})$.

Vlastnosti

Pro každé $\mathcal{A}$ a $\mathcal{B}$ platí

V5. Jestliže $\mathcal{B} \prec \mathcal{A}$ potom $H(\mathcal{A} \mid \mathcal{B}) = 0$.

Intuice: víme, které jevy v $\mathcal{A}$ nastaly.

V6. Jsou-li $\mathcal{A}$ a $\mathcal{B}$ nezávislé, potom

$$H(\mathcal{A} \mid \mathcal{B}) = H(\mathcal{A}), \qquad H(\mathcal{B} \mid \mathcal{A}) = H(\mathcal{B})$$

Intuice: Jsou-li jevy nezávislé, potom pozorováním $\mathcal{B}$ nezískáme žádnou informaci o $\mathcal{A}$.

V7. $H(\mathcal{A} \cdot \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})$

Intuice: entropie složeného jevu není větší než entropie dílčích jevů.

V8. Jsou-li $\mathcal{A}$ a $\mathcal{B}$ nezávislé, potom

$$H(\mathcal{A} \cdot \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B})$$

V9. $H(\mathcal{A} \cdot \mathcal{B}) = H(\mathcal{B}) + H(\mathcal{A} \mid \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B} \mid \mathcal{A})$

Intuice: Pozorujeme-li $\mathcal{B}$, získáme o $\mathcal{A}$ informaci $H(\mathcal{B})$, takže zbytková nejistota je $H(\mathcal{A} \mid \mathcal{B}) = H(\mathcal{A} \cdot \mathcal{B}) - H(\mathcal{B})$.

V10. $H(\mathcal{A}) - H(\mathcal{A} \mid \mathcal{B}) = H(\mathcal{B}) - H(\mathcal{B} \mid \mathcal{A})$

V11. $H(\mathcal{B}) \leq H(\mathcal{A} \cdot \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})$

V12. Pro každé $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ platí:

$$\begin{aligned} H(\mathcal{B} \cdot \mathcal{C} \mid \mathcal{A}) &= H(\mathcal{B} \mid \mathcal{A}) + H(\mathcal{C} \mid \mathcal{A} \cdot \mathcal{B}) \\ &= H(\mathcal{C} \mid \mathcal{A}) + H(\mathcal{B} \mid \mathcal{C} \cdot \mathcal{A}) \end{aligned}$$

Pozn: důsledek vlastnosti V9.

V13. $H(\mathcal{A} \mid \mathcal{B}) \leq H(\mathcal{A})$

Intuice:

1. Pozorováním $\mathcal{B}$ se nejistota o $\mathcal{A}$ nemůže zvětšit.
2. Entropie podmnožiny rozkladu je menší.

V14. Jestliže $\mathcal{B} \prec \mathcal{C}$ potom $H(\mathcal{A} \mid \mathcal{B}) \leq H(\mathcal{A} \mid \mathcal{C})$.

Intuice: Jemnějším rozkladem se dozvíme více o $\mathcal{A}$.

Vzájemná informace

Pozorování $\mathcal{B}$ redukuje nejistotu o $\mathcal{A}$ z $H(\mathcal{A})$ na $H(\mathcal{A} \mid \mathcal{B})$ a získá tedy o $\mathcal{A}$ informaci $I(\mathcal{A}, \mathcal{B}) = H(\mathcal{A}) - H(\mathcal{A} \mid \mathcal{B})$.

Def.

$$I(\mathcal{A}, \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B}) - H(\mathcal{A} \cdot \mathcal{B})$$

Vlastnosti

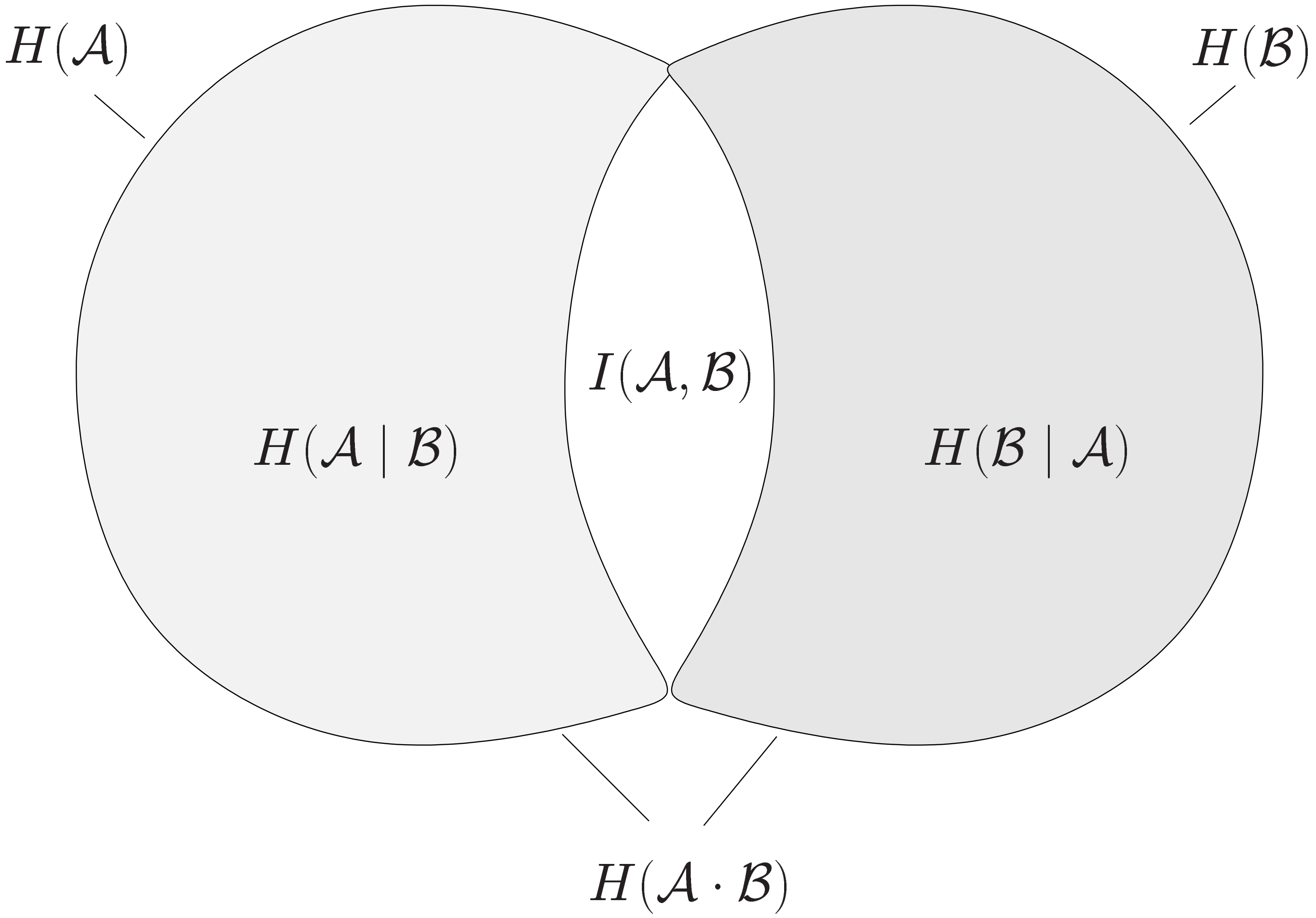
$$\begin{aligned} I(\mathcal{A}, \mathcal{B}) &= H(\mathcal{A}) - H(\mathcal{A} \mid \mathcal{B}) \\ I(\mathcal{A}, \mathcal{B}) &\geq 0 \\ I(\mathcal{A}, \mathcal{B}) &= I(\mathcal{B}, \mathcal{A}) \end{aligned}$$

$I(\mathcal{A}, \mathcal{B})$

1. Informace o $\mathcal{A}$ obsažená v $\mathcal{B}$
2. Informace o $\mathcal{B}$ obsažená v $\mathcal{A}$

$$I(\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_k) = \sum_{i=1}^{k} H(\mathcal{A}_i) - H(\mathcal{A}_1 \cdot \mathcal{A}_2 \cdots \mathcal{A}_k)$$

Vzájemný vztah mezi podmíněnými druhy entropie



Entropie náhodné proměnné

Nechť X je diskrétní náhodná proměnná nabývající hodnot x_i s pravděpodobnostmi

$$P(X = x_i) = p_i$$

Sjednocení událostí $\{X = x_i\}$ tvoří rozklad $\mathcal{A}_X$.

Def. 1

Entropie $H(X)$ diskrétní náhodné proměnné X je rovna

$$H(X) = H(\mathcal{A}_X) = -\sum_i p_i \ln p_i$$

Def. 2

Entropie $H(X)$ spojité náhodné proměnné X je rovna

$$H(X) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx$$

Pozn: události $\{X = x_i\}$ spojité X netvoří rozklad.

Sdružená entropie náhodné proměnné

$$H(X,Y) = H(\mathcal{A}_X \cdot \mathcal{A}_Y) = -\sum_i \sum_j p_{ij} \ln p_{ij}$$

Podmíněná entropie

$$H(X \mid Y) = H(\mathcal{A}_X \mid \mathcal{A}_Y)$$

Jestliže $y = f(x)$ je vzájemně jednoznačné zobrazení, potom

V15. $H(Y \mid X) = H(X \mid Y) = 0$

V16. $I(X,Y) = H(Y), \quad I(X,Y) = H(X).$

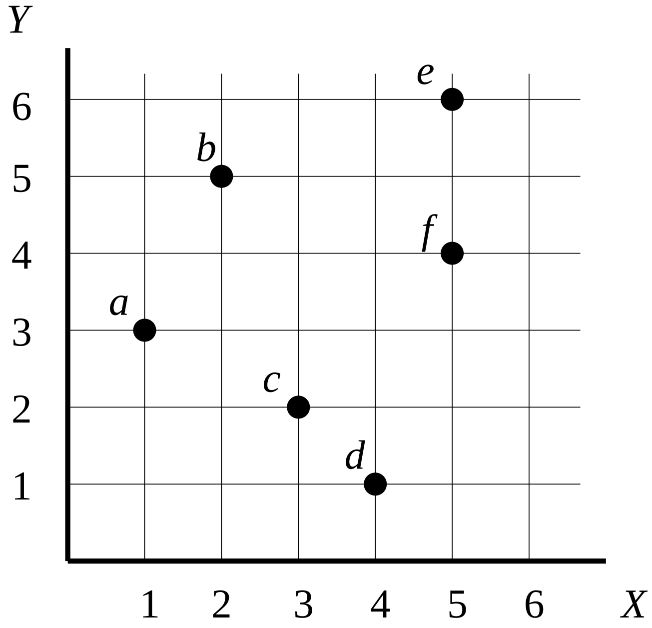
Normovaná střední vzájemná informace

$$i(X,Y) = \frac{I(X,Y)}{H(X,Y)}, \qquad 0 \le i(X,Y) \le 1$$

$i(X,Y) = 0$ pokud X, Y jsou nezávislé

$i(X,Y) = 1$ pokud mezi X, Y existuje jednoznačná deterministická funkční závislost

Příklad



$$
\begin{aligned}
X &= \{1, 2, 3, 4, 5, 5\} \\
Y &= \{3, 5, 2, 1, 6, 4\} \\
X \times Y &= \{(1,3), (2,5), (3,2), (4,1), (5,6), (5,4)\} \\
H(X) &= 1.5607 \\
H(Y) &= 1.7918 \\
H(X,Y) &= 1.7918 \\
I(X,Y) &= 1.5607
\end{aligned}
$$

normovaná stř. vzájemná inf.	$i(X,Y) = 0.8710$
normalizovaný korelační koeficient	$\rho(X,Y) = 0.1964$
Spearmanův rankový koeficient	$r(X,Y) = 0.2609$

Entropie systému s chováním

Stavový prostor vzorkovacích proměnných

$$\mathbf{S} = \mathbf{S}_1 \times \mathbf{S}_2 \times \cdots \times \mathbf{S}_n$$

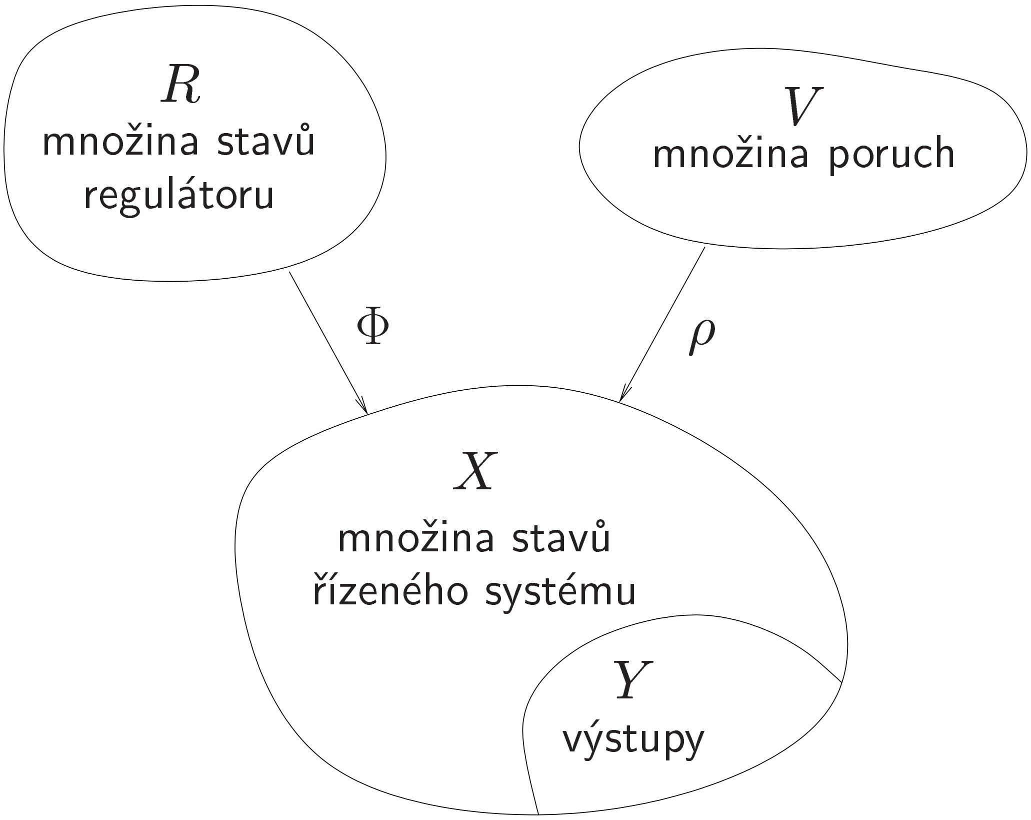
- Množina všech stavů tvoří elementární rozklad $\mathbf{S}$
- Každá proměnná s_i definuje rozklad $\mathbf{S}_i = \mathcal{A}_{s_i}$
- Entropie systému $\mathbf{S} = \{s_1, s_2, \ldots, s_k\}$ je

$$H(\mathbf{S}) = H(\mathbf{S}_1 \cdot \mathbf{S}_2 \cdots \mathbf{S}_k)$$

- Generativní nejistota je podmíněná entropie

$$H(\mathbf{G} \mid \bar{\mathbf{G}})$$

Optimální řízení



$$\Phi\colon R \to X, \qquad \rho\colon V \to X$$

Pozorování

Kapacita kanálu řízení $I(X,R)$ ovlivňuje kvalitu regulace:

$$I(Y,R) \leq I(X,R)$$

Nejlepší řízení nastane pro $I(X,R) \mapsto \max$, tj. bude-li Φ vzájemně jednoznačné zobrazení. Regulátor musí tedy být modelem systému.