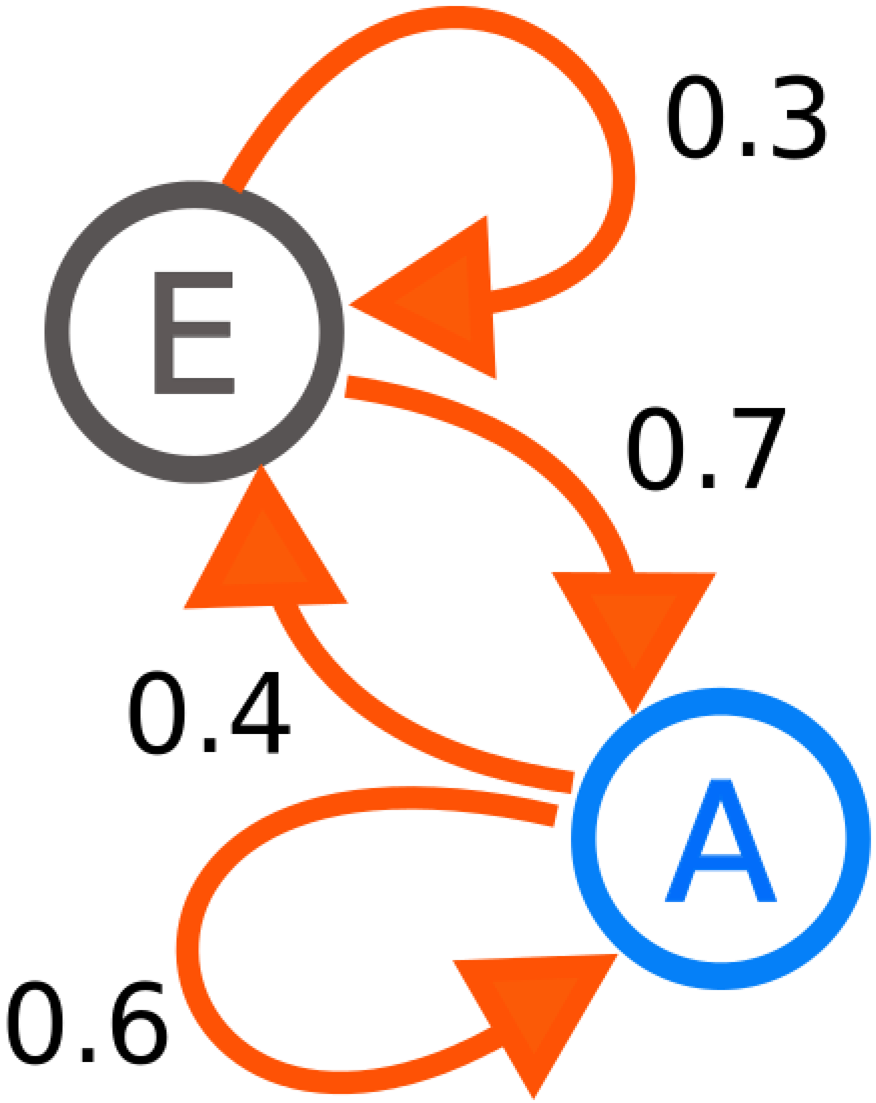
Sistemi probabilistici e stocastici

«Alcune scelte sono più probabili di altre»



Valeria Tateo 635002

Alcune definizioni:

Definizione 1.1.1. *Una **σ-algebra** è un sottoinsieme $\mathscr{F}$ delle parti di un insieme Ω tale che*

1. $\Omega \in \mathscr{F}$;
2. *se* $A \in \mathscr{F}$, *allora* $A^c \in \mathscr{F}$;
3. *se* $(A_n)_{n \geq 1} \subseteq \mathscr{F}$, *allora* $\bigcup_n A_n \in \mathscr{F}$.

*Un'**algebra** è definita analogamente, ma è chiusa per unioni finite (invece che unioni numerabili).*

Esempio: $\Omega = \{a,b,c,d\}$

$\mathscr{F} = \{ \varnothing, \{a,b\}, \{c,d\}, \{a,b,c,d\}\}$

Alcune definizioni:

Definizione 1.1.2. *Dato Ω un insieme e $\mathscr{F}$ una σ-algebra, $\mu : \mathscr{F} \longrightarrow [0, +\infty]$ è una* ***misura*** *se*

- $\mu(A) \geq 0 \ \forall A \in \mathscr{F}$;
- *è σ-additiva: $(A_n)_{n\geq 1} \subseteq \mathscr{F}$ a due a due disgiunti, allora $\mu\Big(\bigsqcup_n A_n\Big) = \sum_n \mu(A_n)$;*

In generale per una misura μ vale che:

- *se $\mu(\Omega) < \infty$, allora μ è finita;*
- *se Ω è unione di una famiglia numerabile di insiemi (di $\mathscr{F}$) di misura finita, allora μ è σ-finita;*
- *se $\mu(\Omega) = 1$, allora μ è una* ***probabilità****.*

$(\Omega, \mathscr{F}, \mathbb{P})$ SPAZIO DI PROBABILITÀ

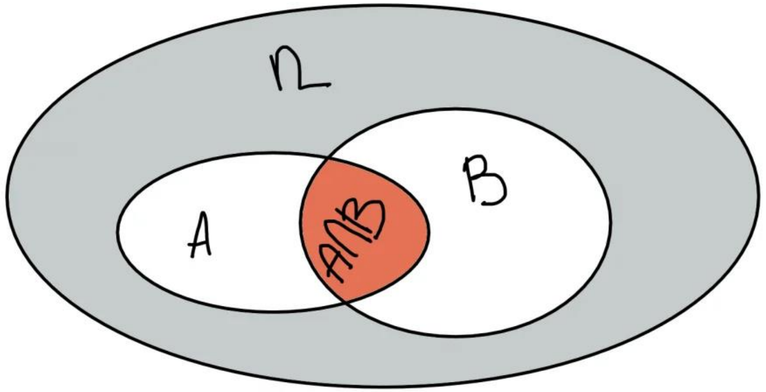
PROPRIETÀ:

- Principio di inclusione e esclusione:

$$P(\Omega \setminus A) = 1 - P(A)$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$$P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2)$$

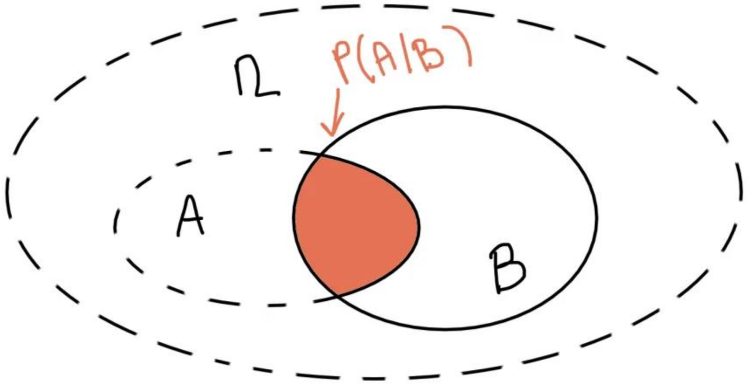


- Probabilità condizionata:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$



Variabili aleatorie continue:

Dato $(\Omega, \mathcal{A}, P)$ spazio di probabilità, una funzione misurabile $X : \Omega \to \mathbb{R}$ è una **variabile aleatoria** reale se

$$\forall x \in \mathbb{R}.\ \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A}$$

E si definisce la **funzione di ripartizione**

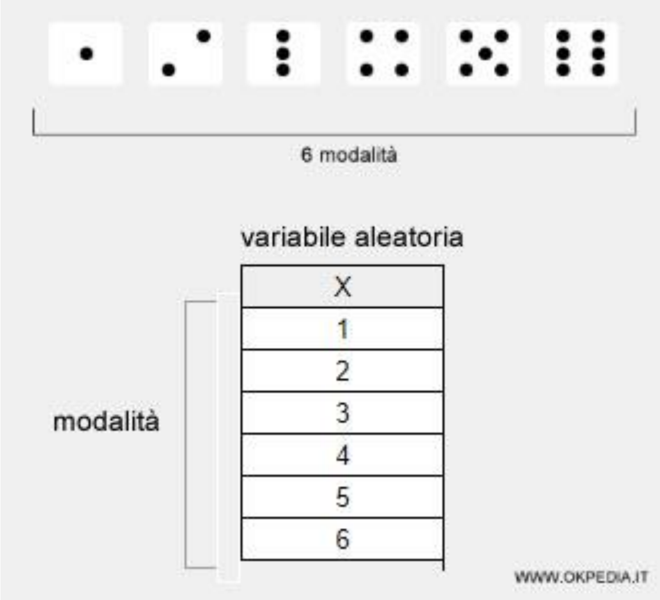
$$F_X(x) \triangleq P(X \leq x) = P(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

Proprietà:

$$P(X \leq a) = F_X(a)$$

$$P(X > a) = 1 - F_X(a)$$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$



Variabili aleatorie continue:

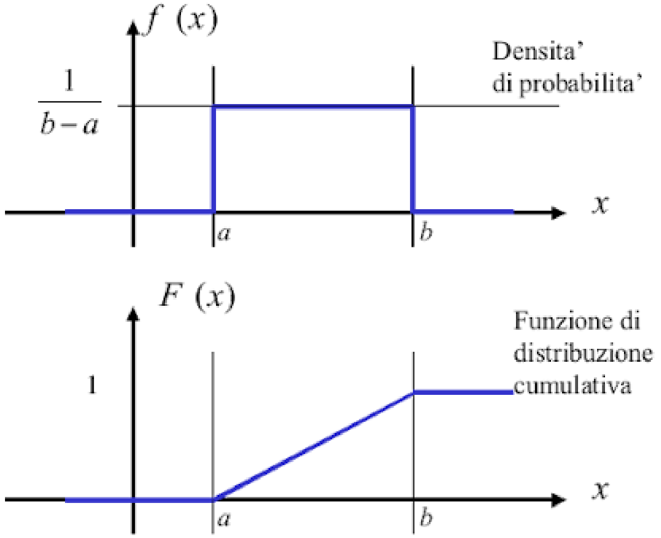
Dato $(\Omega, \mathcal{A}, P)$ spazio di probabilità e data una variabile aleatoria reale $X : \Omega \to \mathbb{R}$,

una funzione integrabile $f_X : \mathbb{R} \to [0, +\infty)$ si dice **densità di probabilità** se è tale che

$$F_X(a) = \int_{-\infty}^{a} f_X(x)\ dx$$

Vale anche che

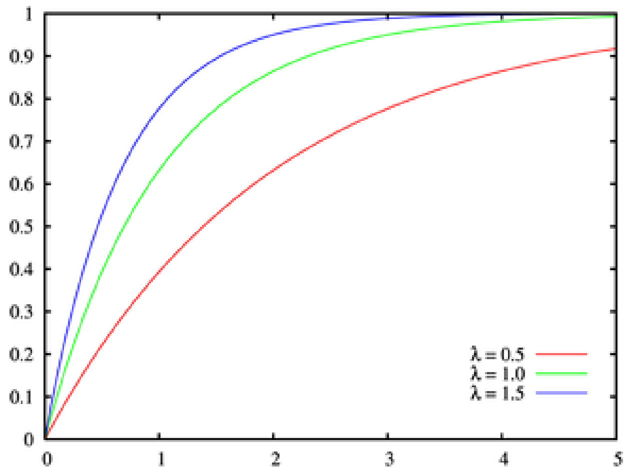
$$P(a < X \leq b) = \int_{a}^{b} f_X(x)\ dx$$

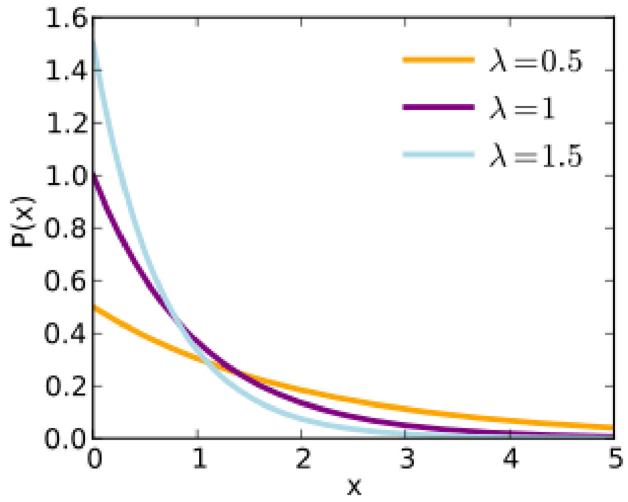


Distribuzione esponenziale e proprietà

La distribuzione esponenziale è così definita

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geqslant 0 \\ 0 & x < 0 \end{cases} \qquad f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geqslant 0 \\ 0 & x < 0 \end{cases}$$





Proprietà distribuzione esponenziale:

Memoryless: Sia X una variabile aleatoria esponenziale di parametro **λ**, allora

$$\boxed{P(X \leq t_0 + t \mid X > t_0) = P(X \leq t)}$$

Proof. Since X is exponentially distributed, its probability law is

$$F_X(t) = \int_0^t \lambda e^{-\lambda x} dx$$

so we need to prove

$$\frac{P(t_0 < X \leq t_0 + t)}{P(X > t_0)} = \frac{\int_{t_0}^{t_0+t} \lambda e^{-\lambda x} dx}{\int_{t_0}^{\infty} \lambda e^{-\lambda x} dx} \stackrel{?}{=} \int_0^t \lambda e^{-\lambda x} dx = P(X \leq t)$$

Since $\int_a^b \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_a^b = \left[e^{-\lambda x}\right]_b^a$ it follows that

$$\frac{\int_{t_0}^{t_0+t} \lambda e^{-\lambda x} dx}{\int_{t_0}^{\infty} \lambda e^{-\lambda x} dx} = \frac{\left[e^{-\lambda x}\right]_{t_0+t}^{t_0}}{\left[e^{-\lambda x}\right]_{\infty}^{t_0}} = \frac{e^{-\lambda t_0} - e^{-\lambda t} \cdot e^{-\lambda t_0}}{e^{-\lambda t_0}} = \frac{\cancel{e^{-\lambda t_0}}(1 - e^{-\lambda t})}{\cancel{e^{-\lambda t_0}}} = 1 - e^{-\lambda t}$$

We conclude by

$$\int_0^t \lambda e^{-\lambda x} dx = \left[e^{-\lambda x}\right]_t^0 = 1 - e^{-\lambda t}$$

□

Proprietà distribuzione esponenziale:

Siano X_1 e X_2 due variabili aleatorie esponenziali di parametri λ_1 e λ_2, allora

$$\boxed{P(\min\{X_1, X_2\} \leq t) = 1 - e^{-(\lambda_1+\lambda_2)t}}$$

Proof. We recall that for any two events (not necessarily disjoint) we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

and that for two independent events we have

$$P(A \cap B) = P(A) \times P(B)$$

Then

$$\begin{aligned}
P(\min\{X_1, X_2\} \leq t) &= P(X_1 \leq t \vee X_2 \leq t) \\
&= P(X_1 \leq t) + P(X_2 \leq t) - P(X_1 \leq t \wedge X_2 \leq t) \\
&= P(X_1 \leq t) + P(X_2 \leq t) - P(X_1 \leq t) \times P(X_2 \leq t) \\
&= (1 - e^{-\lambda_1 t}) + (1 - e^{-\lambda_2 t}) - (1 - e^{-\lambda_1 t})(1 - e^{-\lambda_2 t}) \\
&= 1 - e^{-\lambda_1 t} e^{-\lambda_2 t} \\
&= 1 - e^{-(\lambda_1+\lambda_2)t}
\end{aligned}$$

□

Proprietà distribuzione esponenziale:

Siano X_1 e X_2 due variabili aleatorie esponenziali di parametri λ_1 e λ_2, allora

$$\boxed{P(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}}$$

Proof. Imagine you are at some time t and neither of the two variables has fired. The probability that X_1 fires in the infinitesimal interval dt while X_2 fires in some later instant is

$$\lambda_1 e^{-\lambda_1 t} \left(\int_t^{\infty} \lambda_2 e^{-\lambda_2 t_2} dt_2 \right) dt$$

from which we derive

$$\begin{aligned} P(X_1 < X_2) &= \int_0^{\infty} \lambda_1 e^{-\lambda_1 t_1} \left(\int_{t_1}^{\infty} \lambda_2 e^{-\lambda_2 t_2} dt_2 \right) dt_1 \\ &= \int_0^{\infty} \lambda_1 e^{-\lambda_1 t_1} \left[e^{-\lambda_2 t_2} \right]_{\infty}^{t_1} dt_1 \end{aligned}$$

$$\begin{aligned} &= \int_0^{\infty} \lambda_1 e^{-\lambda_1 t_1} \cdot e^{-\lambda_2 t_1} dt_1 \\ &= \int_0^{\infty} \lambda_1 e^{-(\lambda_1 + \lambda_2) t_1} dt_1 \\ &= \left[\frac{\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2) t} \right]_{\infty}^{0} \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2} \end{aligned}$$

Processi stocastici

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, e sia T un insieme. Allora un processo stocastico è una funzione $X : \Omega \times T \to \mathbb{R}$ tale che

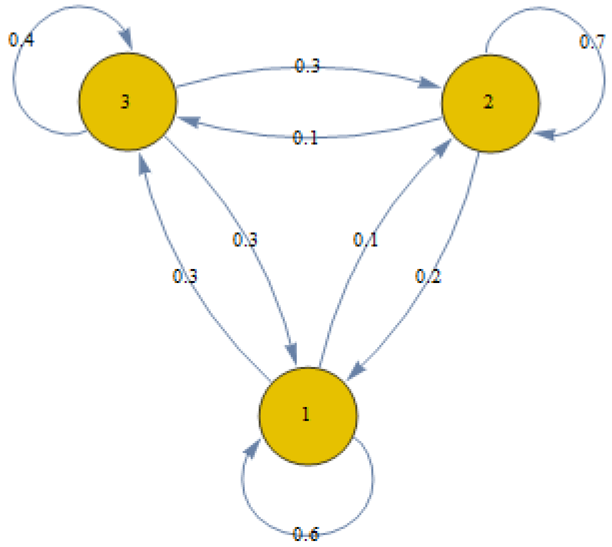
$$\forall t \in T.\, X(\cdot, t) : \Omega \to \mathbb{R} \text{ è una variabile aleatoria}$$

Se T ⊆ ℕ, allora il processo stocastico si dice **discreto.**

Se T ⊆ ℝ, allora il processo stocastico si dice **continuo.**

L'insieme degli stati è

$$S = \{x \mid \exists \omega \in \Omega, t \in T.X(\omega, t) = x\}$$



Catene di Markov

Se T ⊆ ℕ -> DTCM
Se T ⊆ ℝ -> CTCM

Sia $(\Omega, \mathcal{A}, P)$ spazio di probabilità, $\{X_t\}_{t \in T}$ processo stocastico è una Catena di Markov se $\forall t_0 < t_1 < \cdots < t_n < t$ possibili tempi e $\forall x, x_0, x_1, \cdots x_n$ possibili stati si ha

$$\boxed{P(X_{t_{n+1}} = x \mid X_{t_n} = x_n, \ldots, X_{t_0} = x_0) = P(X_{t_{n+1}} = x \mid X_{t_n} = x_n)}$$

Consideriamo solo il caso di Catene di Markov omogenee, ovvero che non dipendono dal tempo. Quindi possiamo riformulare

$$\boxed{P(X_t = x | X_{t_n} = x_n) = P(X_{t-t_n} = x | X_0 = x_n)}$$

Catene di Markov Discrete (DTCM)

Possiamo rappresentare la DTCM attraverso matrici e LTS così definiti:

$$P = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix}$$

$$i \xrightarrow{a_{i,j}} j$$

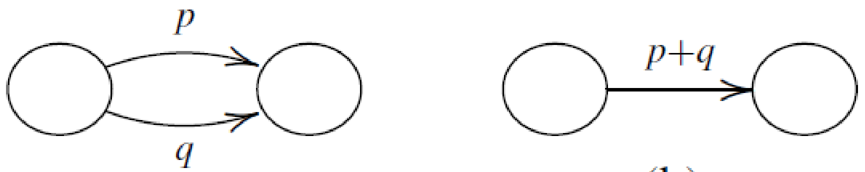
$$a_{i,j} = P(X_1 = j | X_0 = i)$$

Probabilità di transizione

$$\forall i, j \in \{1, ..., N\}.\ 0 \leq a_{i,j} \leq 1$$

$$\forall i \in \{1, ..., N\}.\ \sum_{j=1}^{N} a_{i,j} = 1$$

PTS



(a) (b)

Un **PTS** è un LTS le cui etichette sono le probabilità di transizione.

La **funzione di transizione** è così definita:

$$\alpha_D(s) = \begin{cases} \lambda s'.\, P(X_1 = s' \mid X_0 = s) & \text{if } s \text{ is not a deadlock state} \\ * & \text{otherwise} \end{cases}$$

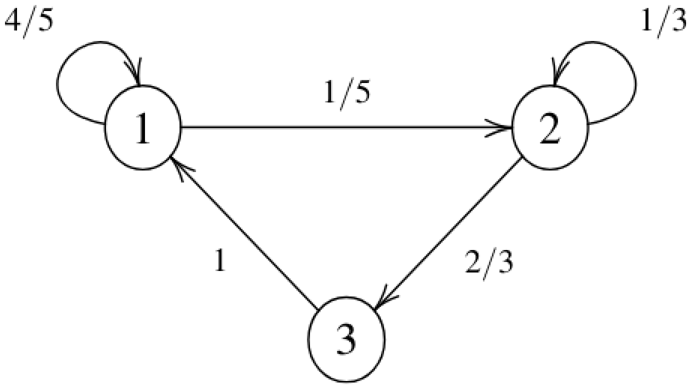
Cioè

$$a_{i,j} = P(X_1 = j | X_0 = i) \; = \alpha_D(s_i)(s_j)$$

Possiamo adesso sfruttare le proprietà di algebra lineare per capire come evolve il sistema, infatti

$$\pi^{(t+1)} = \pi^{(t)} P = \begin{vmatrix} \pi_1^{(t)} & \pi_2^{(t)} & \pi_3^{(t)} \end{vmatrix} \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}$$

$$
\begin{array}{cc}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
P = & \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 1/3 & 2/3 \\ 1 & 0 & 0 \end{bmatrix} \begin{array}{c} 1 \\ 2 \\ 3 \end{array}
\end{array}
$$



$$\pi^{(t)} = |1/3 \ 1/3 \ 1/3|$$

La somma fa 1

$$\pi^{(t+1)} = |1/3 \ 1/3 \ 1/3| \begin{vmatrix} 4/5 & 1/5 & 0 \\ 0 & 1/3 & 2/3 \\ 1 & 0 & 0 \end{vmatrix} = |3/5 \ 8/45 \ 2/9|$$

Esempio:

Come possiamo calcolare la probabilità di andare dallo stato 1 -> 2 -> 3 ->1 ?

FINITE PATH PROBABILITY:

$$P(1\ 2\ 3\ 1) = a_{1,2} \cdot a_{2,3} \cdot a_{3,1} = \frac{1}{5} \cdot \frac{2}{3} \cdot 1 = \frac{2}{15}$$

Distribuzione stazionaria

Si definisce la distribuzione stazionaria $\pi = \left| \pi_1 \dots \pi_n \right|$ di una DTCM come

$$\forall i \in [1, n].\ \pi_i = \lim_{t \to \infty} \pi_i^{(t)}$$

quando tale limite esiste.

Catene di Markov Ergodiche

Una catena di Markov si dice **Ergodica** se:

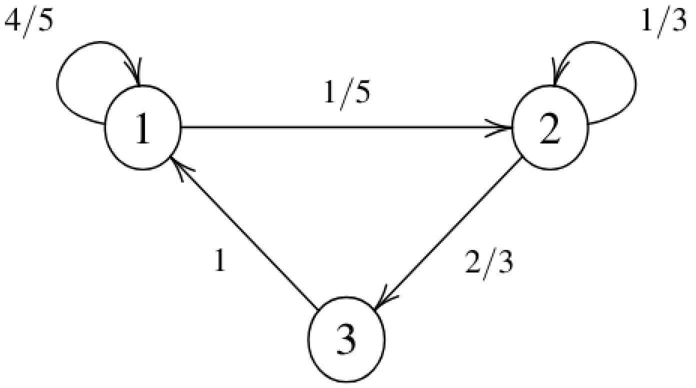
- è irriducibile: ovvero ogni stato è raggiungibile da ogni altro stato;
- è aperiodica: ovvero il massimo comune divisore delle lunghezze di tutti i cammini è 1.

In questo caso, la distribuzione stazionaria si calcola risolvendo il sistema lineare

$$\begin{cases} \pi = \pi \cdot P \\ \sum_{i=1}^{N} \pi_i = 1 \end{cases}$$

Esempio:

$$P = \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 1/3 & 2/3 \\ 1 & 0 & 0 \end{bmatrix}$$



$$\begin{cases} \frac{4}{5}\pi_1 + \pi_3 & = & \pi_1 \\ \frac{1}{5}\pi_1 + \frac{1}{3}\pi_2 & = & \pi_2 \\ \frac{2}{3}\pi_2 & = & \pi_3 \\ \pi_1 + \pi_2 + \pi_3 & = & 1 \end{cases}$$

$\downarrow$

$$\pi = [\ 2/3\ ,\ 1/5\ ,\ 2/15\ ]$$

Catene di Markov Continue (CTCM)

Sia $(\Omega, \mathcal{A}, P)$ spazio di probabilità, $\{X_t\}_{t \in \mathbb{R}}$ CTCM tale che

$$\boxed{P(X_{t_n+\Delta t} = x | X_{t_n} = x_n, \cdots, X_{t_0} = x_0) = P(X_{\Delta t} = x | X_0 = x_n)}$$

Sia $T_{i,j}$ il tempo speso nello stato i prima di andare nello stato j, allora

$$\boxed{P(T_{i,j} > t + \Delta t | T_{i,j} > t) = P(T_{i,j} > \Delta t)}$$

Cioè è una variabile memoryless, ovvero una variabile esponenziale di parametro $\lambda_{i,j}$

Catene di Markov Continue (CTCM)

Con le condizioni $$T_i = \min_{j \neq i} T_{i,j} \qquad \lambda_i = \sum_{j \neq i} \lambda_{i,j}$$

si definisce **l'embedded DTCM** di CTCM la catena di Markov discreta rappresentata dalla matrice

$$P = \begin{bmatrix} 0 & p_{1,2} & \cdots & p_{1,N} \\ p_{2,1} & 0 & \cdots & p_{2,N} \\ \vdots & \vdots & & \vdots \\ p_{N,1} & p_{N,2} & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & \lambda_{1,2}/\lambda_1 & \cdots & \lambda_{1,N}/\lambda_1 \\ \lambda_{2,1}/\lambda_2 & 0 & \cdots & \lambda_{2,N}/\lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_{N,1}/\lambda_N & \lambda_{N,2}/\lambda_N & \cdots & 0 \end{bmatrix}$$

dove

$$p_{i,j} = P\left(T_{i,j} < \min_{k \neq i,j} T_{i,k}\right) = \frac{\lambda_{i,j}}{\lambda_i}$$

CTCM Bisimilarità

Dato un LTS definiamo la funzione γ che preso uno stato *p,* un'azione l e un insieme di stati I restituisce **true** se esiste uno stato *q* in I tale che *q* è raggiungibile da *p* tramite l; restituisce **false** altrimenti.

$$\gamma(p,\ell,I) \stackrel{\text{def}}{=} \exists q \in I \,.\, p \xrightarrow{\ell} q$$

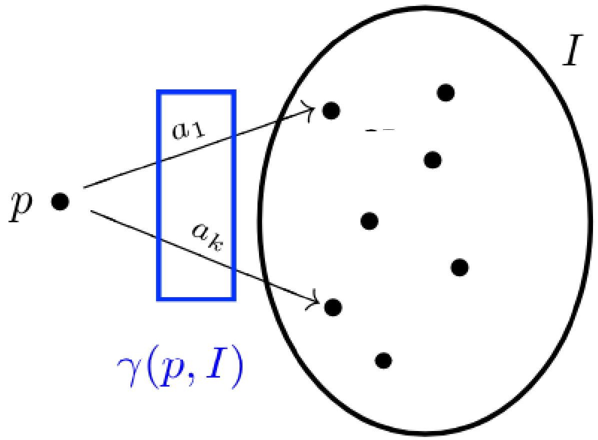
Ora prendiamo l'equivalenza $\equiv_R$ indotta da una relazione R.

Sia $S_{\equiv_R}$ l'insieme delle classi di equivalenza.

Se J ∈ $S_{\equiv_R}$, p,q ∈ J (cioè appartengono alla stessa classe di equivalenza)

if $p \xrightarrow{\mu} p'$ for some μ, p' then $q \xrightarrow{\mu} q'$ for some q' with $p' \equiv_{\mathbf{R}} q'$
(and vice versa) (i.e. $\exists I \in S_{|\equiv_{\mathbf{R}}} .\ p', q' \in I$)

CTCM Bisimilarità



Se consideriamo la funzione $\Phi : \wp(S \times S) \rightarrow \wp(S \times S)$

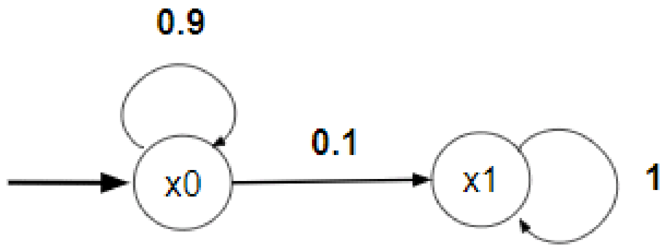
$$p \; \Phi(\mathbf{R}) \; q \triangleq \forall I \in S_{|\equiv_{\mathbf{R}}} . \; \gamma(p, I) = \gamma(q, I)$$

Allora una bisimulazione è tale che $\mathbf{R} \subseteq \Phi(\mathbf{R})$ ← CTCM BISIMULAZIONE

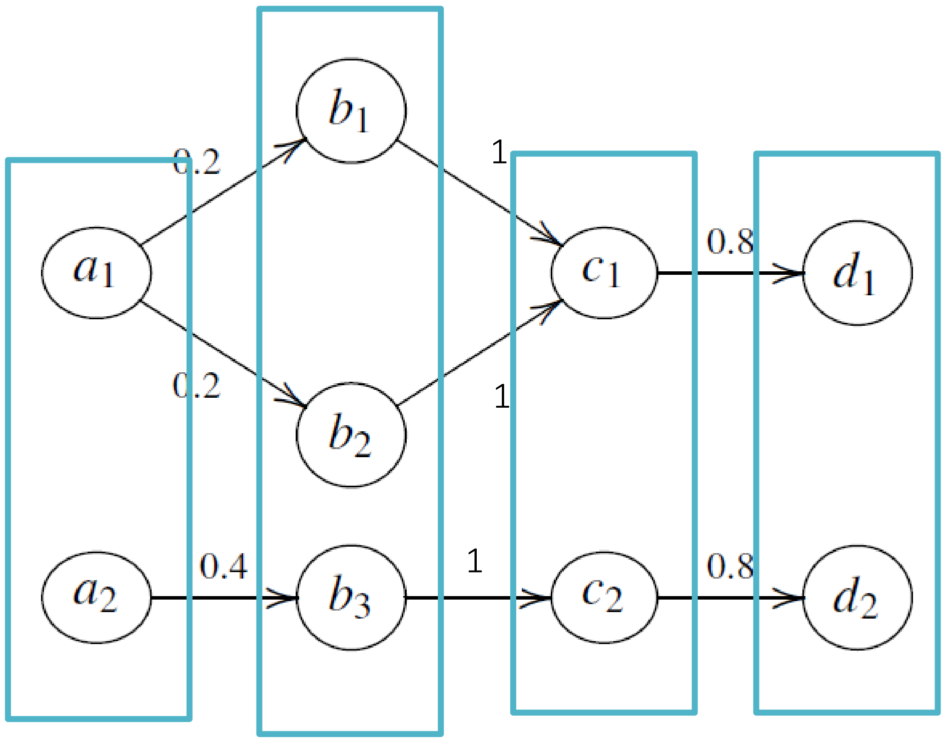
$$\simeq \triangleq \bigcup_{\mathbf{R} \subseteq \Phi(\mathbf{R})} \mathbf{R}$$ è la più grande bisimulazione ← CTCM BISIMILARITA’

DTCM Bisimilarità

- Senza stallo, banale;
- Con stallo, analoga con la condizione che:
  1. Due qualsiasi stati di stallo sono bisimili;
  2. Ogni stato di stallo non è bisimile ad uno stato di non stallo.



Esempio:



$$\equiv_{\mathbf{R}} = \{\ \{a_1, a_2\}\ ,\ \{b_1, b_2, b_3\}\ ,\ \{c_1, c_2\}\ ,\ \{d_1, d_2\}\ \}$$

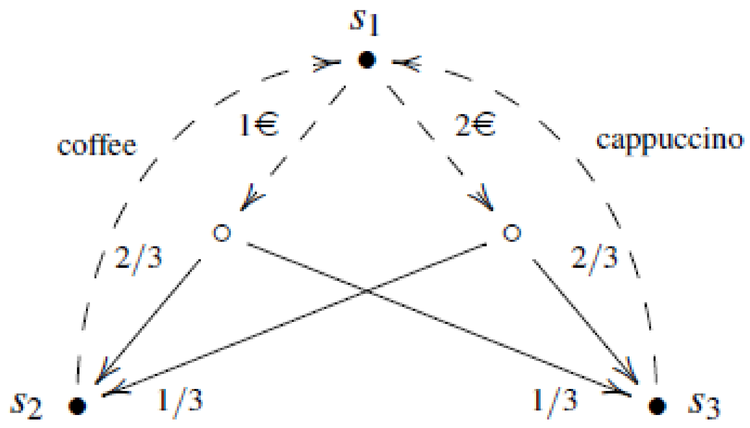
Reactive PTS

$$\alpha_R : S \to L \to \mathbb{D}(S) \cup \{\star\}$$

Sono Catene di Markov che eseguono azioni.

La bisimilarità è analoga a quella già introdotta.

.



Larsen-Skou Logic

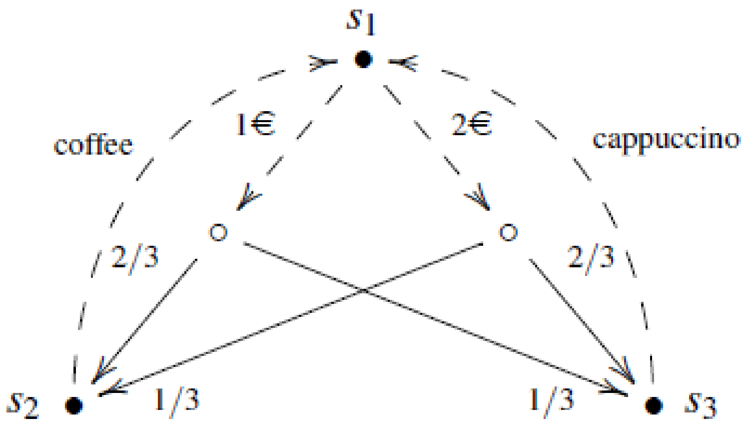
$$\varphi \quad ::= \quad true \quad | \quad \varphi_1 \wedge \varphi_2 \quad | \quad \neg\varphi \quad | \quad \langle \ell \rangle_q \varphi.$$

Teorema

Due stati di un reactive probabilist system sono bisimili se e solo se soddisfano le stesse formule della logica di Larsen-Skou senza la negazione.

$$\begin{array}{ll} s \models true & \\ s \models \varphi_1 \wedge \varphi_2 & \text{if } s \models \varphi_1 \text{ and } s \models \varphi_2 \\ s \models \neg\varphi & \text{if } \neg s \models \varphi \\ s \models \langle \ell \rangle_q \varphi & \text{if } \gamma_r(s)(\ell)\,[\![\varphi]\!] \geq q \text{ where } [\![\varphi]\!] = \{s' \in S \mid s' \models \varphi\} \end{array}$$

Consiste nell'abilità di raggiungere un nuovo stato attraverso la formula φ con un'azione l con probabilità almeno q



$$s_1 \overset{?}{\models} \langle 1€\rangle_{\frac{1}{2}} \langle \text{coffee}\rangle_1 \mathbf{tt} \qquad \gamma_R(s_1, 1€, [\![\langle \text{coffee}\rangle_1 \mathbf{tt}]\!]) \overset{?}{\geq} \frac{1}{2}$$

$$\begin{aligned} [\![\langle \text{coffee}\rangle_1 \mathbf{tt}]\!] &= \{u \mid u \models \langle \text{coffee}\rangle_1 \mathbf{tt}\} \\ &= \{u \mid \gamma_R(u, \text{coffee}, [\![\mathbf{tt}]\!]) \geq 1\} \\ &= \{u \mid \gamma_R(u, \text{coffee}, S) \geq 1\} \\ &= \{s_2\} \end{aligned}$$

$$\gamma_R(s_1, 1€, \{s_2\}) = \frac{2}{3} \geq \frac{1}{2}$$

Esempio:

GRAZIE PER L' ATTENZIONE!

Bibliografia: Roberto Bruni, Ugo Montanari, "Models of Computation", Springer Texts in Computer Science, 2017.