Vorlesung 10a

Mehrstufige Zufallsexperimente

und

Markovketten

bisher: von “heute” zu “morgen”

jetzt: von “heute und morgen” zu “übermorgen”, etc.

Für jedes $i = 1, \ldots, n-1$ hat man

Übergangswahrscheinlichkeiten

$$P(a_1 \ldots a_i, a_{i+1}) = \mathbf{P}_{a_1 \ldots a_i}(X_{i+1} = a_{i+1}) \ ,$$

die angeben,

mit welcher Wahrscheinlichkeit in der $(i+1)$-ten Stufe

das Ereignis $\{X_{i+1} = a_{i+1}\}$ eintritt,

gegeben das Eintreten von $\{X_1 = a_1, \ldots, X_i = a_i\}$.

Die gemeinsame Verteilung von $X_1, \ldots, X_n$

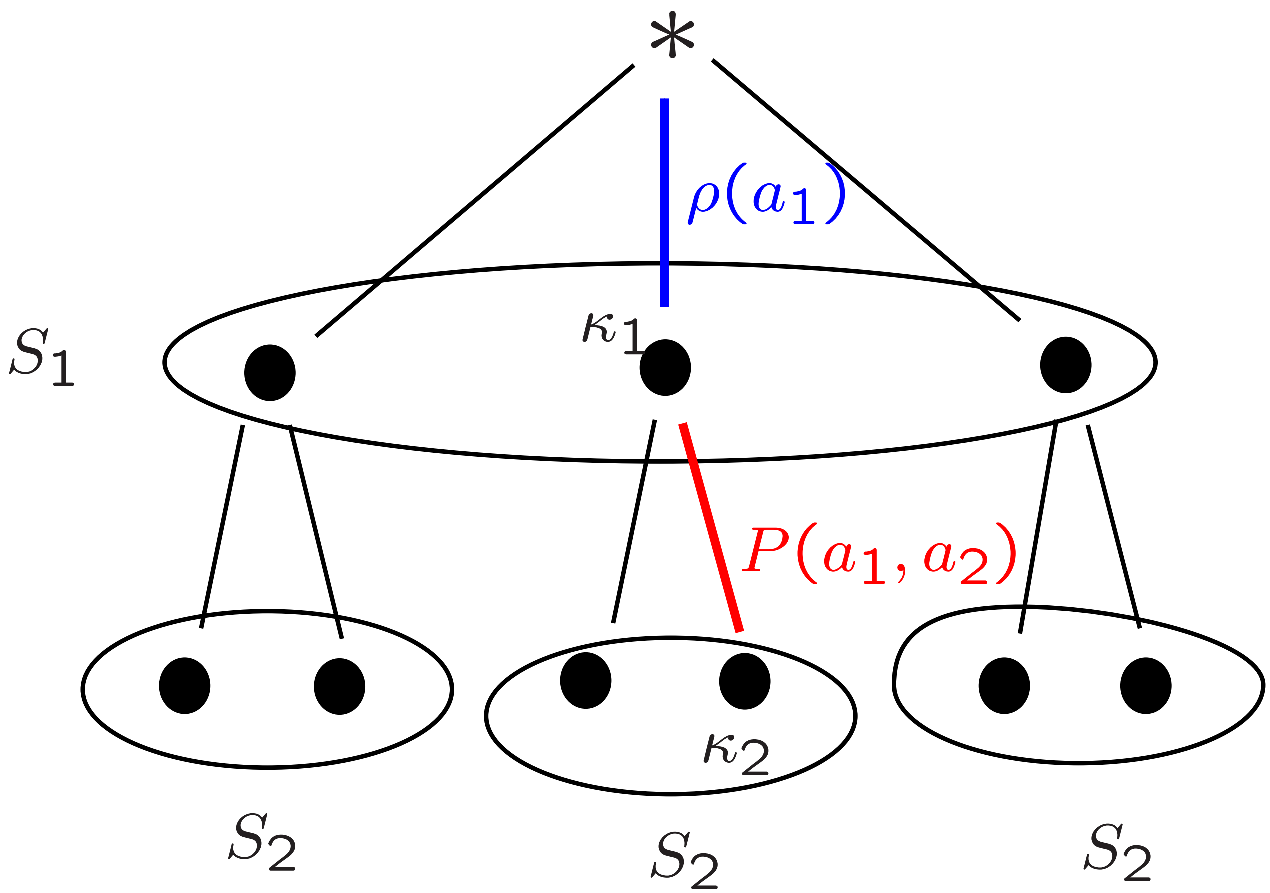
ergibt sich rekursiv als

Multiplikationsregel

$$\mathbf{P}(X_1 = a_1, \ldots, X_n = a_n)$$
$$= \mathrm{P}(X_1 = a_1, \ldots, X_{n-1} = a_{n-1}) P(a_1 \ldots a_{n-1}, a_n)$$
$$= \rho(a_1) P(a_1, a_2) \cdots P(a_1 \ldots a_{n-1}, a_n) \,.$$

Veranschaulichung von zweistufigen Experimenten

durch *Bäume*:



$\kappa_1 = a_1$

$\kappa_2 = a_1 a_2$

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \rho(a_1) P(a_1, a_2)\,.$$

(Produkt der Kantengewichte von $*$ zum Knoten κ_2)

Veranschaulichung von mehrstufigen Experimenten durch ***Bäume***:

$\kappa_i = a_1 \ldots a_i$ ist ein *Knoten der Tiefe* i

Die *Nachfolger* von κ_i sind von der Form $\kappa_i\, a_{i+1}$

Die *Kanten* des Baums erhalten die Gewichte

$$g(*, \kappa_1) := \rho(a_1)\ , \quad g(\kappa_i, \kappa_{i+1}) := P(a_1 \ldots a_i, a_{i+1})$$

mit $\kappa_1 = a_1,\ \ \kappa_i = a_1 \ldots a_i,\ \ \kappa_{i+1} = a_1 \ldots a_{i+1}$.

Die Wahrscheinlichkeit, in einem bestimmten Blatt zu enden, ergibt sich als Produkt der Kantengewichte entlang des Weges von der Wurzel zum Blatt.

Eine wichtige Beispielklasse mehrstufiger Zufallsexperimente

ist die, bei der

die Übergangswahrscheinlichkeiten der nächsten Stufe

nur von der aktuellen Stufe abhängen

(und nicht von den vorhergehenden):

$$P(\ldots a_{i-2}\, a_{i-1}, a_i) = P(a_{i-1}, a_i)$$

In dem Fall spricht man von einer **Markovkette**

mit Übergangsmatrix P.

Die Stufen sind jetzt mit $i = 0, 1, 2, \ldots$ indiziert.

Man denkt sich die Übergangsmatrix P als fest

und notiert die Verteilung ρ von X_0 (die "Startverteilung")

als Subskript bei der Wahrscheinlichkeit $\mathbf{P}$.

Die Multiplikationsregel ergibt:

$$\mathbf{P}_\rho(X_0 = a_0, \ldots, X_n = a_n)$$
$$= \rho(a_0)P(a_0, a_1) \cdots P(a_{n-1}, a_n)$$

Startet die Kette in $a \in S$,

dann ist ρ die auf a konzentrierte Verteilung

(notiert als $\rho = \delta_a$).

Statt $\mathbf{P}_{\delta_a}$ schreibt man auch $\mathbf{P}_a$

und erhält

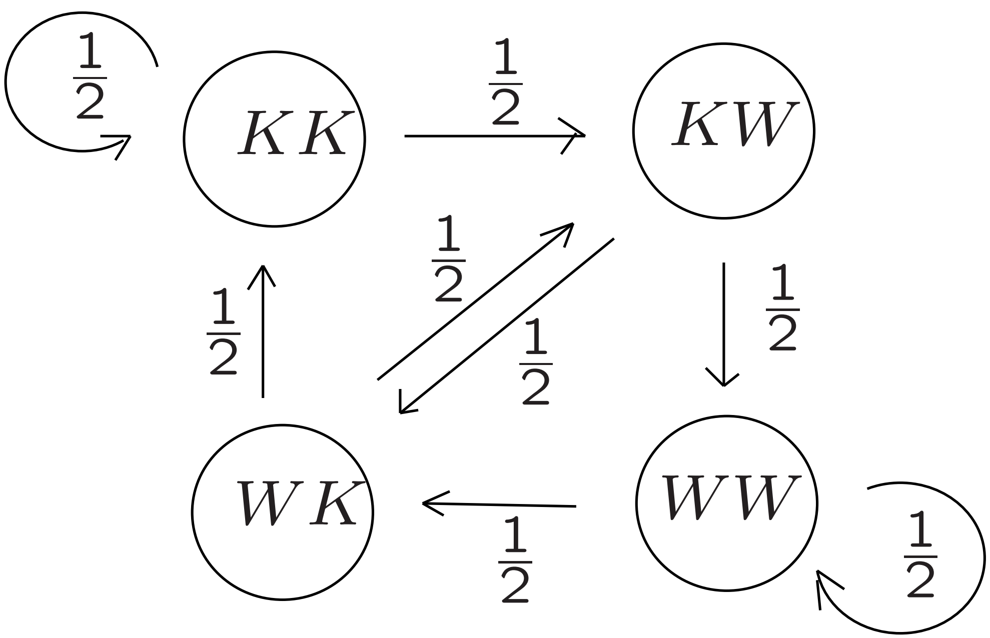
$$\mathbf{P}_a(X_0 = a) = 1,$$

$$\mathbf{P}_a(X_1 = a_1, \ldots, X_n = a_n) = P(a, a_1) \cdots P(a_{n-1}, a_n) \ .$$

Beispiel 1:

Muster der Länge 2 beim fairen Münzwurf

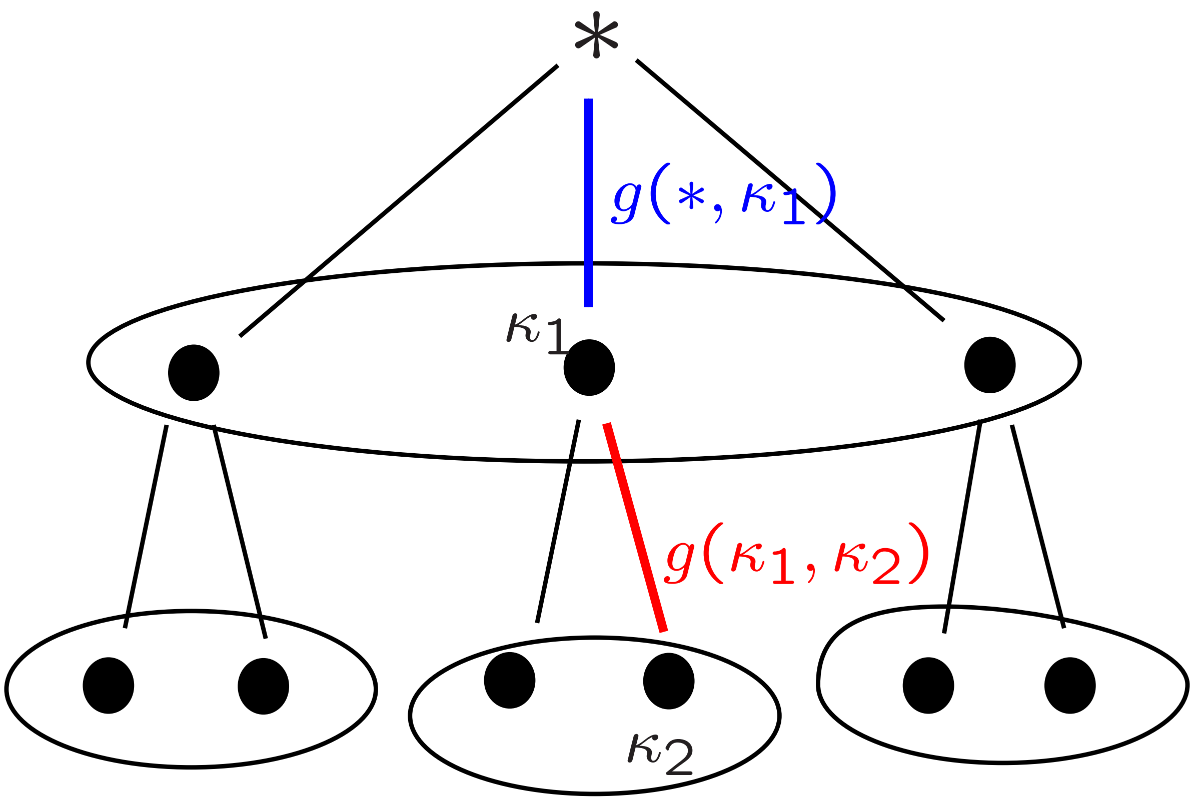
Graph der Übergangswahrscheinlichkeiten:



Beispiel 2:

Zufällige Wanderung durch einen Baum

von der Wurzel zur Krone



$$\mathbf{P}(X_1 = \kappa_1, X_2 = \kappa_2) = g(*,\kappa_1) \cdot g(\kappa_1,\kappa_2)\,.$$

(Produkt der Kantengewichte von $*$ zum Knoten κ_2)

Beispiel 2:

Zufällige Wanderung durch einen Baum

von der Wurzel zur Krone

Die Zustände sind die Knoten;

die Übergangswahrscheinlichkeiten sind die Kantengewichte:

$$P(\kappa,\kappa') := \begin{cases} g(\kappa,\kappa'), & \text{falls } \kappa' \text{ Nachfolger von } \kappa, \\ 0 & \text{sonst.} \end{cases}$$

Die Blätter werden zu *absorbierenden Zuständen*.

Beispiel 3:

(p, q)-Irrfahrt auf $\mathbb{Z}$:

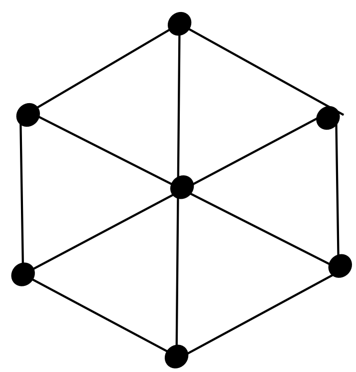
$$S = \mathbb{Z}$$

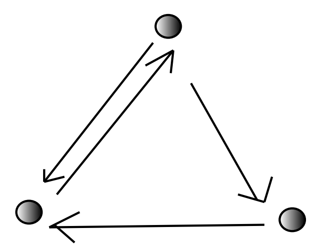
$$P(k, k+1) = p, \quad P(k, k-1) = 1 - p =: q$$

Beispiel 4:

Einfache Irrfahrt

auf einem (ungerichteten oder gerichteten) Graphen





$S :=$ die Menge der Knoten.

Der nächste Schritt erfolgt jeweils

zu einem rein zufällig ausgewählten Nachbarn.

Beispiel 5:

Pólya-Urne

$$S = \{(w, b) : w, b \in \mathbb{N}\} = \mathbb{N}^2$$

$$P((w, b), (w + 1, b)) = \frac{w}{w + b} \,,$$

$$P((w, b), (w, b + 1)) = \frac{b}{w + b} \,.$$

Zur Erinnerung:

Für eine Markovkette $X = (X_0, X_1, \ldots)$

mit Start in a und Übergangsmatrix P

hat man die Wahrscheinlichkeiten

$$\mathbf{P}_a(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n)$$
$$= P(a, a_1)P(a_1, a_2) \cdots P(a_{n-1}, a_n).$$

Zerlegung nach dem ersten Schritt.

$$\mathbf{P}_a(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n)$$

$$= P(a, a_1)P(a_1, a_2) \cdots P(a_{n-1}, a_n)$$

$$= P(a, a_1)\mathbf{P}_{a_1}(X_1 = a_2, \ldots, X_{n-1} = a_n)$$

Summation über $a_1, a_2, \ldots, a_{n-1}$,

mit b statt a_1 und c statt a_n:

$$\mathbf{P}_a(X_n = c) = \sum_{b \in S} P(a, b)\, \mathbf{P}_b(X_{n-1} = c)\ .$$

Treffwahrscheinlichkeiten

Die Frage:

P sei eine Übergangsmatrix auf der Menge S

X sei Markovkette mit Übergangsmatrix P.

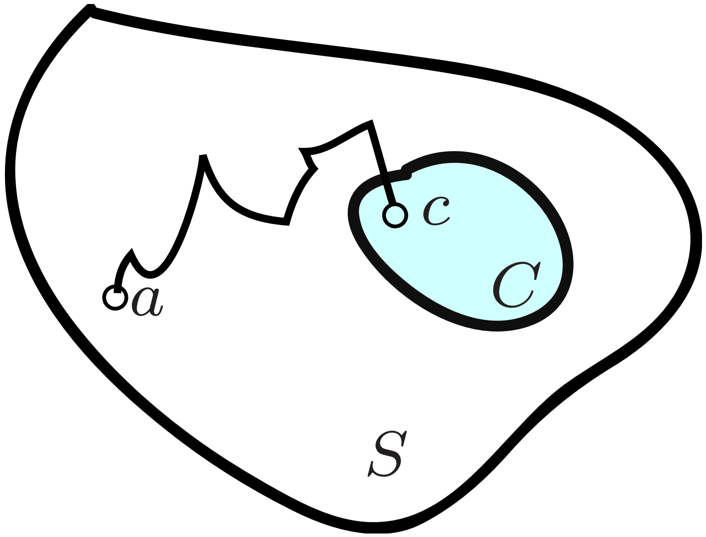
$C \subset S$, $c \in C$ seien fest.

Wie wahrscheinlich ist es,

dass der in $a \in S$ startende Pfad

die Menge C erstmals

im Zustand c trifft?



Treffwahrscheinlichkeiten

Die Antwort:

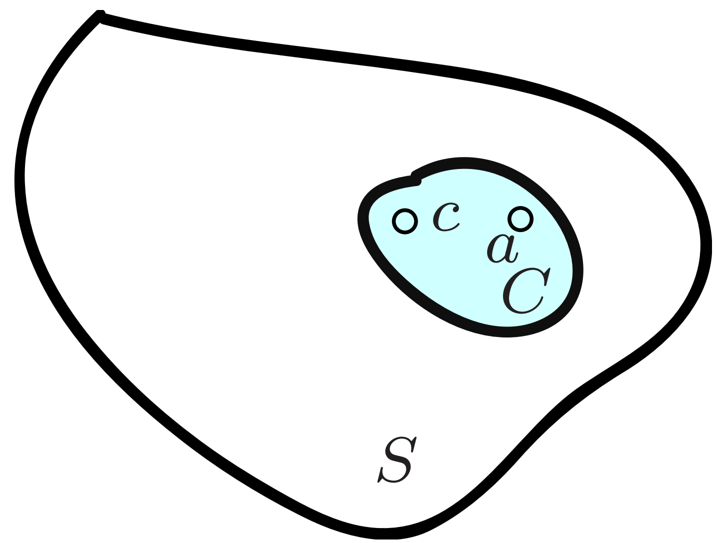
Sei $w(a)$ die gesuchte Wahrscheinlichkeit.

Die $w(a)$, $a \in S$, erfüllen das Gleichungssystem

$$\sum_{b \in S} P(a,b) w(b) = w(a)$$

für $a \in S \setminus C$,

$$w(a) = \delta_{ac} \quad \text{für } a \in C.$$

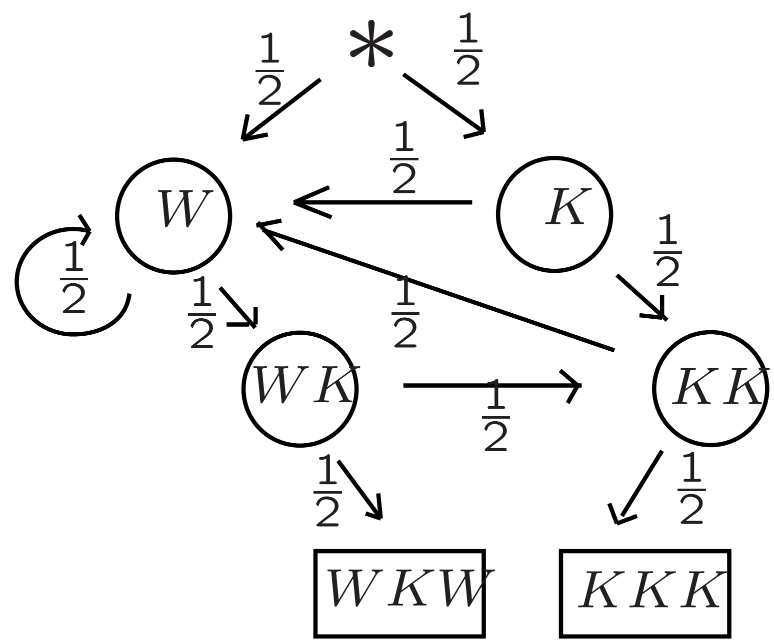


Beispiel A

Welches Muster kommt eher?

Mit welcher W'keit kommt beim fairen Münzwurf das Muster KKK früher als das Muster WKW?

Hier ist ein "reduzierter Graph" der relevanten Zustände und Übergänge:

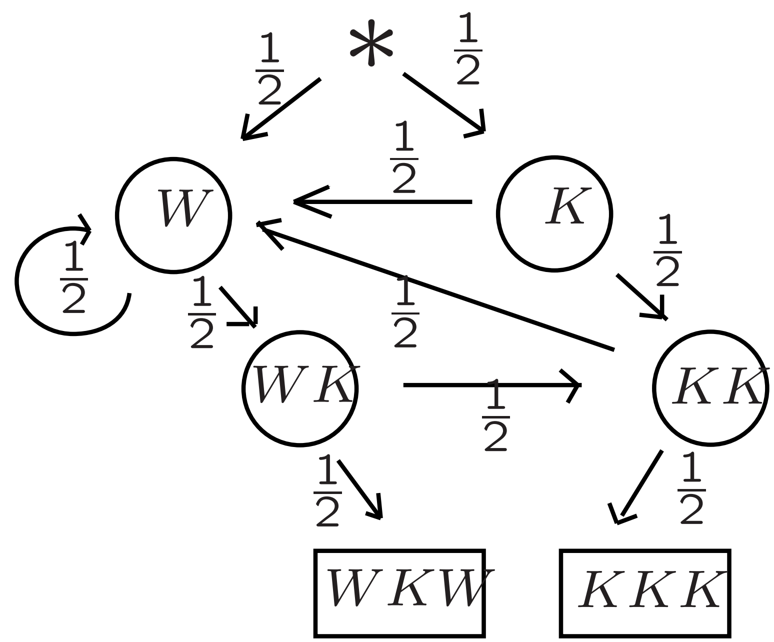


Für $w(a) :=$

$\mathbf{P}_a$(Spiel endet in KKK)

ergibt sich das

Gleichungssystem



$$w(KK) = \tfrac{1}{2} + \tfrac{1}{2}w(W)\,, \quad w(WK) = \tfrac{1}{2}w(KK)$$

$$w(K) = \tfrac{1}{2}w(KK) + \tfrac{1}{2}w(W)\,, \quad w(W) = \tfrac{1}{2}w(WK) + \tfrac{1}{2}w(W)\,.$$

und daraus

$$w(W) = \tfrac{1}{3},\, w(K) = \tfrac{1}{2},\;\; w(*) = \tfrac{1}{2}w(W) + \tfrac{1}{2}w(K) = \tfrac{5}{12}\,.$$

Beispiel B: **Gewinn oder Ruin?**

Eine einfache Irrfahrt auf $\mathbb{Z}$ starte im Punkt 3.

Mit welcher W'keit erreicht sie den Punkt $c = 10$,

bevor sie zum Nullpunkt kommt?

"Zerlegung nach dem ersten Schritt" und Randbedingungen:

$$w(a) = \frac{1}{2}w(a-1) + \frac{1}{2}w(a+1)\,, \quad a = 1, \ldots, c-1\,,$$

$$w(0) = 0, \quad w(c) = 1.$$

Fazit: Die $w(a)$ liegen auf einer Geraden,

$$w(a) = \beta a + \gamma \quad \text{mit } \gamma = 0, \beta = 1/c.$$

Die gesuchte Wahrscheinlichkeit ist also $3/10$.

Erwartete Treffzeiten

Sei X eine Markovkette mit Zustandsraum S
und Übergangsmatrix P.

Für eine Teilmenge $C \subset S$ ist

$$T_C = \min\{n : n \geq 0, X_n \in C\}$$

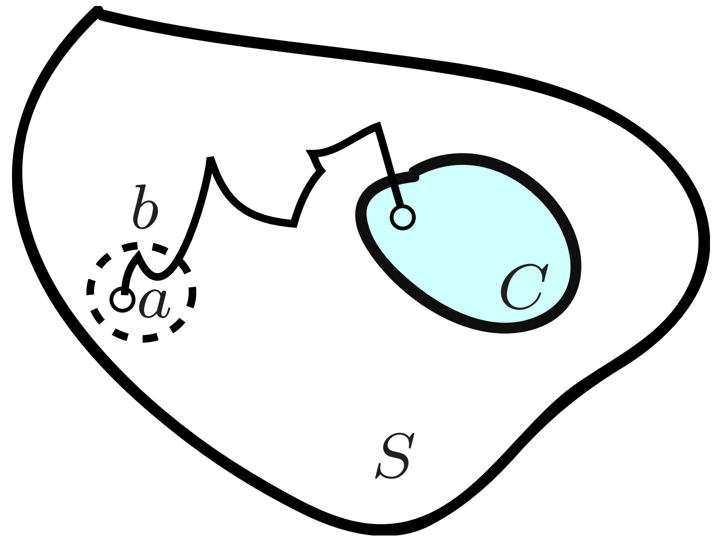
die erste Treffzeit von C.

Es geht um die Berechnung von $\mathbf{E}_a[T_C]$:

Für $a \notin C$

Zerlegung von $\mathbf{E}_a[T_C]$

nach dem ersten Schritt:



Erst ein Schritt

von a nach b gemäß $P(a, b)$,

dann "Neustart" in b:

$$\mathbf{E}_a[T_C] = 1 + \sum_{b \in S} P(a, b)\, \mathbf{E}_b[T_C]$$

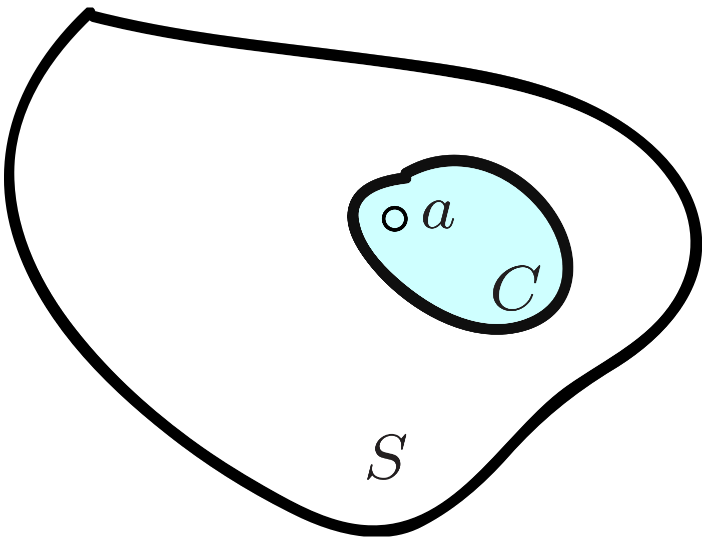
(Formales Argument hierfür: siehe Buch Seite 106.)

Und für $a \in C$

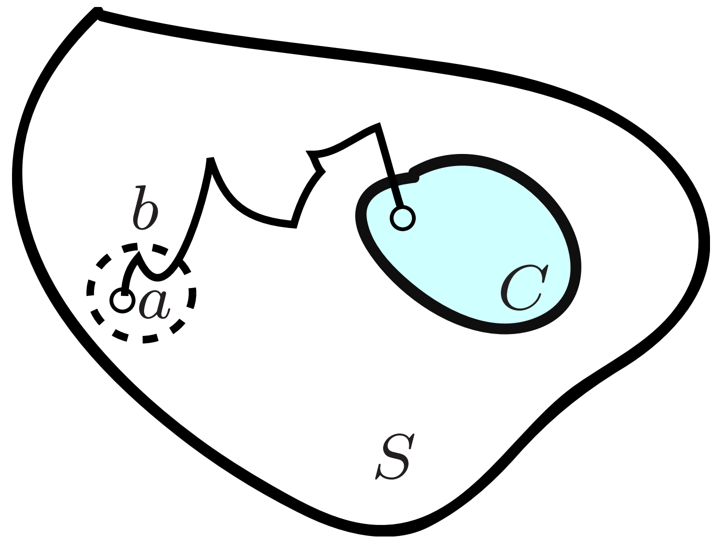
ist $\mathbf{P}_a(T_C = 0) = 1$,

also

$\mathbf{E}_a[T_C] = 0$.



Fazit:



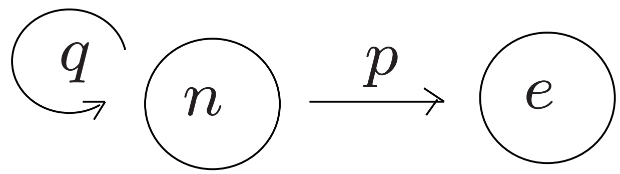
$a \mapsto e(a) := \mathbf{E}_a[T_C]$ erfüllt das Gleichungssystem

$$\mathbf{e}(a) = 1 + \sum_{b \in S} P(a,b)\, e(b) \quad \text{für } a \notin C,$$

$$\mathbf{e}(a) = 0 \quad \text{für } a \in C.$$

Beispiel:

Erwartete Zeit bis zum ersten Erfolg.



$$\mathbf{E}_n[T_e] = 1 + q\mathbf{E}_n[T_e] + p\mathbf{E}_e[T_e] \ .$$

Wegen $\mathbf{E}_e[T_e] = 0$ wird dies gelöst durch

$$\mathbf{E}_n[T_e] = 1/p.$$

Transport von Verteilungen

$\mathbf{P}_\rho(X_n = c)$ zerlegt nach X_{n-1} ergibt die Rekursion

$$\mathbf{P}_\rho(X_n = c) = \sum_{b \in S} \mathbf{P}_\rho(X_{n-1} = b)\, P(b, c)$$

Mit

$$\pi_n(\cdot) := \mathbf{P}_\rho(X_n \in \cdot)\ , \quad n = 0, 1, \ldots$$

lautet diese Rekursion

$$\pi_n(c) = \sum_{b \in S} \pi_{n-1}(b) P(b, c)\ , \quad n \geq 1$$

Zweischritt-Übergangswahrscheinlichkeiten

$$P^2(a,c) := \mathbf{P}_a(X_2 = c)\ , \quad n \geq 0\ ,\ a, c \in S$$

Zerlegungen nach dem Zwischenschritt:

$$\mathbf{P}_a(X_2 = c) = \sum_{b \in S} \mathbf{P}_a(X_1 = b)\mathbf{P}_b(X_1 = c)$$

$$P^2(a,c) = \sum_{b \in S} P(a,b)\, P(b,c)$$

$P^2 = (P^2(a,c))_{a,c \in S}$ lässt sich also auffassen als

Produkt der Matrix P mit sich selbst.

Mehrschritt-Übergangswahrscheinlichkeiten

$$P^n(a,c) := \mathbf{P}_a(X_n = c)\ , \quad n \geq 0\ ,\ a,c \in S$$

Zerlegung nach dem ersten Schritt:

$$P^n(a,c) = \sum_{b \in S} P(a,b)\, P^{n-1}(b,c)$$

Zerlegung nach dem letzten Schritt:

$$P^n(a,c) = \sum_{b \in S} P^{n-1}(a,b)\, P(b,c)$$

$P^n = (P^n(a,c))_{a,c \in S}$ lässt sich also auffassen als n-te Matrixpotenz von P.

Gleichgewichtsverteilungen

Eine Verteilung π auf S heißt

Gleichgewichtsverteilung zur Übergangsmatrix P,

wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(G1) $$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b)\,, \quad b \in S$$

(G2) $$\sum_{a \in S} \pi(a) P(a,b) = \pi(b)\,, \quad b \in S\,.$$

(Dann haben auch $X_2, X_3, \ldots$ die Verteilung π.)

Reversible Gleichgewichtsverteilungen

Hinreichend für

(G1)

$$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b)\ , \quad b \in S$$

ist die Bedingung

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a)\ , \quad a, b \in S$$

(Denn dann ist das Paar (X_0, X_1) so verteilt wie (X_1, X_0), also insbesondere X_0 so wie X_1.)

Gleichbedeutend mit

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a) , \quad a, b \in S$$

ist

$$\pi(a)P(a,b) = \pi(b)P(b,a) , \quad a, b \in S .$$

π heißt dann *reversible Gleichgewichtsverteilung* zu P.

Beispiel

für eine nicht reversible Gleichgewichtsverteilung:

Zyklische Irrfahrt auf $S = \{a, b, c\}$, mit

$$P(a, b) = P(b, c) = P(c, a) := p,$$

$$P(b, a) = P(c, b) = P(a, c) := 1 - p$$

Die uniforme Verteilung auf S

ist eine Gleichgewichtsverteilung zu P.

Nur für $p = 1/2$ ist sie reversibel.

Beispiel:

Die einfache Irrfahrt auf dem Würfel $S = \{0, 1\}^3$:

Von jedem $a \in S$ geht man in einem Schritt
zu einem rein zufällig ausgewählten Nachbarn.
(Zwei Elemente von S heißen *benachbart*,
wenn sie sich in genau einer Komponente unterscheiden.)
Für benachbarte Knoten a und b ist hier $P(a, b) = 1/3$.

Die uniforme Verteilung auf S
ist reversible Gleichgewichtsverteilung.

Beispiel

Die einfache Irrfahrt auf einem ungerichteten Graphen

mit endlicher Knotenmenge S

Von jedem $a \in S$ geht man in einem Schritt

zu einem rein zufällig ausgewählten Nachbarn.

Es stellt sich heraus:

Die Gewichte der Knoten unter der Gleichgewichtsverteilung

sind proportional zur Anzahl der Nachbarn der Knoten

(vgl. Übungen 32 und 38)