

52. Markovské řetězce a základní techniky pro jejich analýzu.

Markovské řetězce

Uspořádanou trojici $M = (S, s_0, P)$ nazveme **Markovským řetězcem** (pracujícím s diskretním časem) DTMC, pokud

- S je konečná a neprázdná množina stavů
- $s_0 \in S$ je počáteční stav
- $P : S \times S \rightarrow \langle 0, 1 \rangle$ je přechodová matice pravděpodobností, kde
 - $\forall s \in S : \sum_{t \in S} P(s, t) = 1$, tedy suma ohodnocení všech hran vystupujících z libovolného stavu s sumuje do jedničky
 - zobrazení definujeme jako totální

Pozn.: Typicky se Markovský řetězec definuje jako čtveřice $M = (S, s_0, P, L)$, kde význam složek S, s_0, P odpovídá definici uvedené výše a $L : S \rightarrow 2^{AP}$ je zobrazení anotující stavy atomickými pozorováními z množiny AP . Anotování stavů má význam pro rozhodování vlastností o těchto modelech, přičemž dané vlastnosti jsou popsány pomocí pravděpodobnostních logik. V rámci této otázky se ovšem anotování stavů neprojeví v žádném z prezentovaných témat, je tedy pro jednoduchost vypuštěno z definice.

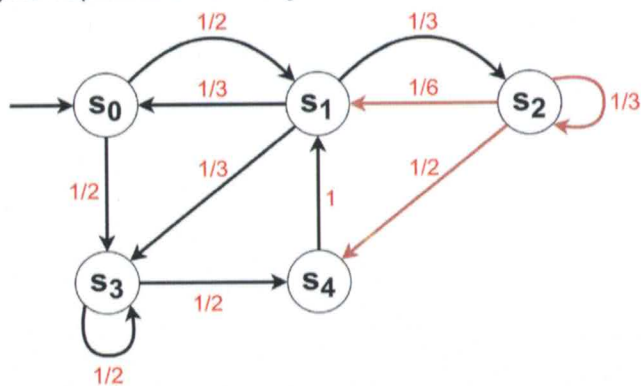
Markovský řetězec je možné znázornit

- **maticově**
 - využijeme matici P , která obsahuje všechny potřebné informace
 - matice má rozměry $|S| \times |S|$
 - ze záhlaví matice lze poznat, který stav je počáteční (označen s_0)
 - buňka $p_{a,b}$ matice P (buňka na řádku a , ve sloupci b) obsahuje hodnotu $P(a, b)$
 - všimněme si kupříkladu řádku s_2 , který sumuje do jedničky, neboť platí, že $1/6 + 1/3 + 1/2 = 1$
 - zde musí každý řádek sumovat do jedničky

P	s_0	s_1	s_2	s_3	s_4
s_0	0	1/2	0	1/2	0
s_1	1/3	0	1/3	1/3	0
s_2	0	1/6	1/3	0	1/2
s_3	0	0	0	1/2	1/2
s_4	0	1	0	0	0

- **graficky**

- o využijeme náčrt podobný konečným automatům
- o tento model ovšem nebude mít žádné koncové stavy a přechody budou ohodnoceny reálnými čísly z intervalu $\langle 0, 1 \rangle$, nikoliv libovolnými symboly
- o pokud $P(s, t) = 0$ pro nějaké dva stavy $s, t \in S$, pak příslušnou hranu vynecháme
- o uvažujme například o následujícím Markovském řetězci



- všimněme si kupříkladu stavu s_2 , kde

$$\sum_{t \in S} P(s_2, t) = P(s_2, s_1) + P(s_2, s_2) + P(s_2, s_4) = \frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1$$

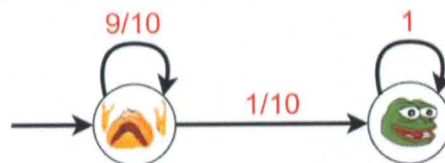
Markovský řetězec je tedy stavový stochastický model, který popisuje diskrétní přechody mezi stavy pomocí pravděpodobnostní distribuce.

Pozn.: Místo počátečního stavu lze také případně definovat iniciální distribuci, tedy nějaké iniciální zobrazení $l: S \rightarrow \langle 0, 1 \rangle$, kde $\sum_{s \in S} l(s) = 1$. V takovém případě se všechna pravděpodobnost nemusí nutně na začátku nacházet ve stavu s_0 , ale může být jinak rozprostřená skrze Markovský řetězec. Zde si ale vystačíme s iniciálním stavem.

Definujeme následující pojmy:

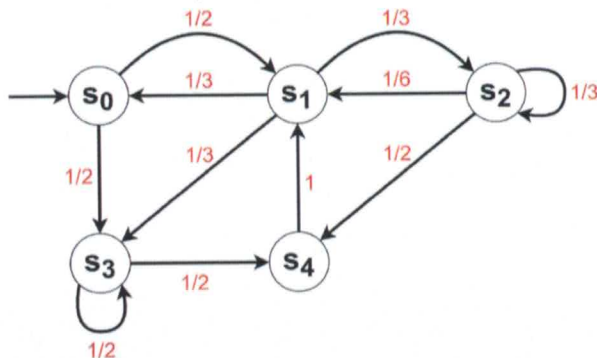
- **vlastnost bezpaměťovosti** (*memorylessness, Markovská vlastnost*)
 - o ať X je náhodná veličina nad nějakým pravděpodobnostním prostorem (Ω, \mathcal{S}, P)
 - o řekneme, že náhodná veličina X má **vlastnost bezpaměťovosti**, pokud platí:
 - $P(X > s + t | X \geq s) = P(X > t)$
 - musí platit $P(X \geq s) > 0$, aby tato podmíněná pravděpodobnost byla dobře definovaná
 - to lze ekvivalentně zapsat jako $P(X > s + t) = P(X \geq s) \cdot P(X > t)$
 - o pro lepší ilustraci uvažujme o následujícím příkladu:
 - snažím se v tělocvičně opakovaně házet míč na koš, trefím se s pravděpodobností $1/10$, s pravděpodobností $9/10$ tragicky selžu
 - během tohoto výkonu se žádným způsobem nemění mé schopnosti, házím pořád stejně kvalitně
 - počítám, kolik se objevilo neúspěšných pokusů před prvním úspěchem
 - první tři hody byly neúspěšné

- do tělocvičny náhle vstoupí kamarád, který rovněž začne počítat moje neúspěšné pokusy před prvním úspěchem
- kamarád nemá ponětí o tom, že už jsem třikrát hodil míčem, proto začíná počítat od nuly
- následující pokus je z mého pohledu čtvrtý, z jeho pohledu první, **avšak náhodná veličina popisující toto snažení má vlastnost bezpaměťovosti, tedy minulé pokusy neovlivní úspěch v budoucnu**
- bude tedy například platit $P(X > 3 + 2 | X \geq 3) = P(X > 2)$
 - tedy pravděpodobnost, že neúspěšných bude více než pět pokusů **za předpokladu, že již byly provedeny tři neúspěšné pokusy**, je stejná jako pravděpodobnost, že neúspěšné budou více než dva pokusy
 - $P(X > 5 | X \geq 3) = P(X > 2)$
 - tedy tři již provedené neúspěšné pokusy nemají naprosto žádný vliv na budoucí úspěch
 - po úpravě $P(X > 5) = P(X \geq 3) \cdot P(X > 2)$
 - *použili jsme vzorec pro podmíněnou pravděpodobnost*
 - $1 - F(5) = (1 - F(2)) \cdot (1 - F(2))$
 - *upravili jsme do podoby $P(X > 5) = P(X > 2) \cdot P(X > 2)$, abychom mohli použít distribuční funkci*
 - $\left(1 - \frac{1}{10}\right)^{5+1} = \left(1 - \frac{1}{10}\right)^{2+1} \cdot \left(1 - \frac{1}{10}\right)^{2+1}$
 - *jde o distribuční funkci geometrického rozdělení pravděpodobnosti*
- **Markovské řetězce mají vlastnost bezpaměťovosti**, tedy tuto situaci by bylo možné namodelovat pomocí níže uvedeného Markovského řetězce
 - lze snadno nahlédnout, že po neúspěšném hodu zůstaneme ve stejném stavu, aniž bychom si tento konkrétní neúspěch jakkoliv pamatovali



- **geometrické rozdělení pravděpodobnosti** je jediné diskrétní rozdělení pravděpodobnosti s vlastností bezpaměťovosti
- **exponenciální rozdělení pravděpodobnosti** je jediné spojité rozdělení pravděpodobnosti s vlastností bezpaměťovosti
- **cesta v Markovském řetězci**
 - ať $M = (S, s_0, P)$ je Markovský řetězec
 - posloupnost stavů $s_0 s_1 s_2 \dots$ nazveme **cestou** v Markovském řetězci M , pokud
 - $\forall k \in \mathbb{N} : P(s_k, s_{k+1}) > 0$
 - tedy cesta začíná v počátečním stavu a pokračuje způsobem, který povolují existující hrany

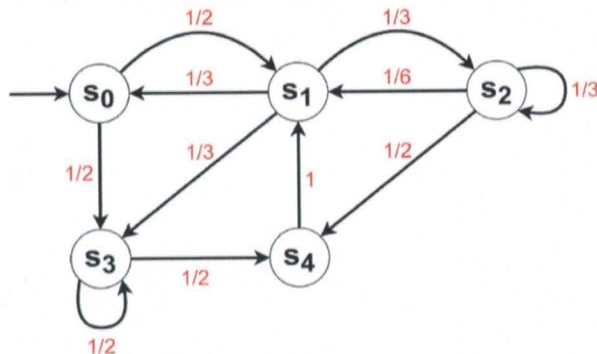
- o cesta může být konečná, nebo nekonečná
- o mějme Markovský řetězec vyobrazený níže
 - posloupnost $s_0s_1s_2s_2s_4$ **je cesta** v tomto Markovském řetězci
 - posloupnost $s_0s_1s_3s_2s_4$ **není cesta** v tomto Markovském řetězci
 - není zde přechod ze stavu s_3 do stavu s_2 , tedy $P(s_3, s_2) = 0$



- **pravděpodobnost cesty v Markovském řetězci Pr**

- o **konečné cesty**

- pro konečnou cestu $s_0s_1s_2\dots s_n$ platí
 - $Pr(s_0s_1s_2\dots s_n) = \prod_{i=0}^{n-1} P(s_i, s_{i+1})$
 - násobíme pravděpodobnosti hran na této cestě
- pro konečnou cestu s_0 platí
 - $Pr(s_0) = 1$
- uvažujme o následujícím Markovském řetězci



- mějme konečnou cestu $s_0s_1s_2s_2s_4$
- její pravděpodobnost činí $Pr(s_0s_1s_2s_2s_4) = P(s_0, s_1) \cdot P(s_1, s_2) \cdot P(s_2, s_2) \cdot P(s_2, s_4) = 1/2 \cdot 1/3 \cdot 1/3 \cdot 1/2 = 1/36$

- o **nekonečné cesty**

- pokud má nekonečná cesta $s_0s_1s_2\dots$ nekonečnou příponu $s_k s_{k+1} s_{k+2} \dots$, kde platí, že $\forall i \in \mathbb{N} : i \geq k \Rightarrow P(s_i, s_{i+1}) = 1$, pak
 - $Pr(s_0s_1s_2\dots) = \prod_{i=0}^{k-1} P(s_i, s_{i+1})$
 - tedy najdeme produkt pouze konečného množství pravděpodobností, pokud všechny následující hrany mají pravděpodobnost 1

- jinak $\Pr(s_0s_1s_2\dots) = 0$

Tranzientní analýza Markovských řetězců

Na Markovské řetězce lze rovněž nahlížet jako na nekonečné posloupnosti diskretních náhodných veličin. Ať X_0, X_1, X_2, \dots jsou diskretní náhodné veličiny nad nějakými pravděpodobnostními prostory $(\Omega_t, \mathcal{S}_t, P_t)$, tedy ať pro každé $t \in \mathbb{N}$ platí, že X_t je diskretní náhodná veličina nad nějakým pravděpodobnostním prostorem $(\Omega_t, \mathcal{S}_t, P_t)$.

Náhodná veličina X_{t+1} bude popisovat chování Markovského řetězce v diskretním čase $t+1$. Pomocí náhodné veličiny X_{t+1} jsme schopni popsat pravděpodobnost, že se Markovský řetězec bude v čase $t+1$ nacházet v nějakém stavu s_{t+1} za předpokladu, že se v čase t nacházel ve stavu s_t , tedy díky vlastnosti bezpaměťovosti se nezabýváme tím, ve kterých stavech se Markovský řetězec nacházel v minulosti.

$$P(X_{t+1} = s_{t+1} | X_t = s_t, X_{t-1} = s_{t-1}, \dots, X_0 = s_0) = P(X_{t+1} = s_{t+1} | X_t = s_t)$$

Pozn.: Tento vztah platí, pokud jsou jednotlivé složky podmíněné pravděpodobnosti dobře definované, zde nenulové.

Ať $M = (S, s_0, P)$ je Markovský řetězec. Uvažujme o náhodných veličinách definovaných výše. Definujeme **tranzientní pravděpodobnost** pro Markovský řetězec M jako zobrazení $t : \mathbb{N} \times S \rightarrow <0, 1>$ (psáno $t_n(s)$), kde:

$$t_n(s) = P(X_n = s | X_0 = s_0),$$

tedy jako pravděpodobnost, že se **v diskretním čase n Markovský řetězec nachází ve stavu s** za předpokladu, že v čase 0 byl ve stavu s_0 (což je naše iniciální podmínka).

Tuto pravděpodobnost lze určit pomocí enumerace všech cest délky n v Markovském řetězci, nebo lze využít dopřednou propagaci pravděpodobnosti pomocí vzorců:

$$\begin{aligned} t_0(s_0) &= 1 \\ t_0(s) &= 0, s \neq s_0 \end{aligned}$$

$$t_n(s) = \sum_{v \in S} t_{n-1}(v) \cdot P(v, s), n > 0$$

Vycházíme tedy z iniciální pravděpodobnosti stavu s_0 , což je tranzientní pravděpodobnost v čase 0 . Postupně počítáme tranzientní pravděpodobnosti v časech $1, 2, \dots$, dokud nedospějeme do času n , přičemž využíváme výsledky z předchozí časové jednotky.

Symbolem T_n značme vektor tranzientních pravděpodobností všech stavů v čase n (vektor o délce $|S|$). Máme-li k dispozici vektor T_n , lze vypočítat vektor T_{n+1} za použití přechodové matice. To lze vyjádřit jako násobení vektoru s maticí P :

$$T_{n+1} = T_n \cdot P$$

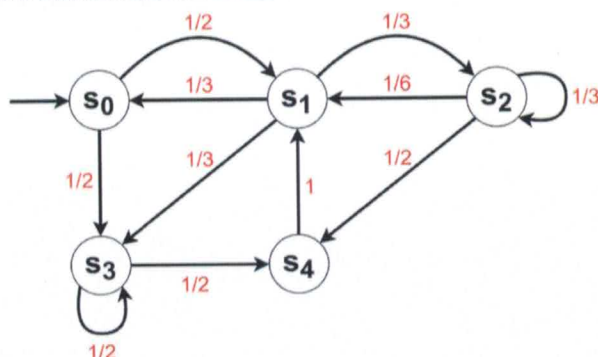
Tranzientní pravděpodobnosti v čase n lze tedy vypočítat pomocí opakovaného umocňování matice P následovně:

$$T_n = T_0 \cdot P^n$$

Tento výpočet tedy lze realizovat v čase $O(|S|^3 \cdot n + |S|^2)$. Využijeme-li při umocňování vhodné asociativitu násobení matic, lze dosáhnout času $O(|S|^3 \cdot \log(n) + |S|^2)$.

Příklad.

Uvažujme o následujícím Markovském řetězci:



Proveďme tranzientní analýzu. Zkusme najít hodnoty $t_4(s)$ pro všechny stavy, tedy najděme vektor T_4 .

stav	s_0	s_1	s_2	s_3	s_4
T_0	1	0	0	0	0

Iniciální distribuce, jistě se nacházíme ve výchozím stavu s_0 .

stav	s_0	s_1	s_2	s_3	s_4
T_1	0	1/2	0	1/2	0

Zde počítáme:

- $$t_1(s_0) = t_0(s_0) \cdot P(s_0, s_0) + t_0(s_1) \cdot P(s_1, s_0) + t_0(s_2) \cdot P(s_2, s_0) + t_0(s_3) \cdot P(s_3, s_0) + t_0(s_4) \cdot P(s_4, s_0) = 1 \cdot 0 + 0 \cdot 1/3 + 0 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 = 0$$
- $$t_1(s_1) = t_0(s_0) \cdot P(s_0, s_1) + t_0(s_1) \cdot P(s_1, s_1) + t_0(s_2) \cdot P(s_2, s_1) + t_0(s_3) \cdot P(s_3, s_1) + t_0(s_4) \cdot P(s_4, s_1) = 1 \cdot 1/2 + 0 \cdot 0 + 0 \cdot 1/6 + 0 \cdot 0 + 0 \cdot 1 = 1/2$$
- $$t_1(s_2) = t_0(s_0) \cdot P(s_0, s_2) + t_0(s_1) \cdot P(s_1, s_2) + t_0(s_2) \cdot P(s_2, s_2) + t_0(s_3) \cdot P(s_3, s_2) + t_0(s_4) \cdot P(s_4, s_2) = 1 \cdot 0 + 0 \cdot 1/2 + 0 \cdot 1/2 + 0 \cdot 0 + 0 \cdot 0 = 0$$
- $$t_1(s_3) = t_0(s_0) \cdot P(s_0, s_3) + t_0(s_1) \cdot P(s_1, s_3) + t_0(s_2) \cdot P(s_2, s_3) + t_0(s_3) \cdot P(s_3, s_3) + t_0(s_4) \cdot P(s_4, s_3) = 1 \cdot 1/2 + 0 \cdot 0 + 0 \cdot 1/3 + 0 \cdot 1/3 + 0 \cdot 0 = 1/2$$

- $$t_1(s_4) = t_0(s_0) \cdot P(s_0, s_4) + t_0(s_1) \cdot P(s_1, s_4) + t_0(s_2) \cdot P(s_2, s_4) + t_0(s_3) \cdot P(s_3, s_4) + t_0(s_0) \cdot P(s_4, s_4)$$

$$= 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 1/2 + 0 \cdot 1/2 + 0 \cdot 0 = 0$$

Další iterace proběhnou obdobně.

stav	s ₀	s ₁	s ₂	s ₃	s ₄
T ₂	1/6	0	1/6	5/12	1/4
T ₃	0	13/36	1/18	7/24	7/24
T ₄	13/108	65/216	5/36	115/432	25/144

Spočítali jsme vektor T₄, tedy pravděpodobnosti, že v čase 4 budeme v jednotlivých stavech Markovského řetězce.

Pokud bychom tuto skutečnost počítali maticově, postupovali bychom následovně:

$$T_4 = T_0 \cdot P \cdot P \cdot P \cdot P$$

$$T_4 = T_0 \cdot (P \cdot P) \cdot (P \cdot P)$$

$$T_4 = T_0 \cdot (P^2 \cdot P^2)$$

$$T_4 = T_0 \cdot P^4$$

Tedy s využitím asociativity maticového násobení stačí počítat generace P², P⁴ a znovu přitom použít mezivýsledky z předchozích iterací.

P	s ₀	s ₁	s ₂	s ₃	s ₄
s ₀	0	1/2	0	1/2	0
s ₁	1/3	0	1/3	1/3	0
s ₂	0	1/6	1/3	0	1/2
s ₃	0	0	0	1/2	1/2
s ₄	0	1	0	0	0

P ²	s ₀	s ₁	s ₂	s ₃	s ₄
s ₀	1/6	0	1/6	5/12	1/4
s ₁	0	2/9	1/9	1/3	1/3
s ₂	1/18	5/9	1/6	1/18	1/6
s ₃	0	1/2	0	1/4	1/4
s ₄	1/3	0	1/3	1/3	0

P^4	s_0	s_1	s_2	s_3	s_4
s_0	13/108	65/216	5/36	115/432	25/144
s_1	19/162	5/18	25/162	89/324	19/108
s_2	2/27	79/324	25/162	31/108	13/54
s_3	1/12	17/72	5/36	5/16	11/48
s_4	2/27	19/54	1/9	13/54	2/9

Všimněme si, že $T_4 = T_0 \cdot P^4$ je vlastně pouze projekcí prvního řádku matice P^4 . Matice P^n tedy pro každý svůj řádek s_k říká, jaká je pravděpodobnost, že po provedení n kroků se budeme nacházet ve stavu daném konkrétním sloupcem za předpokladu, že právě s_k byl iniciální stav.

Neohraňčená dosažitelnost

Ať $M = (S, s_0, P)$ je Markovský řetězec. Dále ať $T \subseteq S$ je množina stavů Markovského řetězce M (množina cílových stavů). Upravme zobrazení P tak, že $P(s, s) = 1$, pokud $s \in T$, tedy stavy z množiny stavů T jsou nově absorbující. Rovněž uvažujme o množině $N \subseteq S$ stavů Markovského řetězce M , která obsahuje stavy, z nichž není možné dosáhnout žádný stav z množiny T .

Pozn.: Množinu N lze zkonstruovat pomocí prohledávání do šířky/do hloubky grafu, který odpovídá Markovskému řetězci M . To se nám povede v čase lineárním vzhledem k velikosti reprezentace tohoto grafu.

Zobrazení $x : S \rightarrow \langle 0, 1 \rangle$ je **neohraňčená dosažitelnost** množiny T , pokud

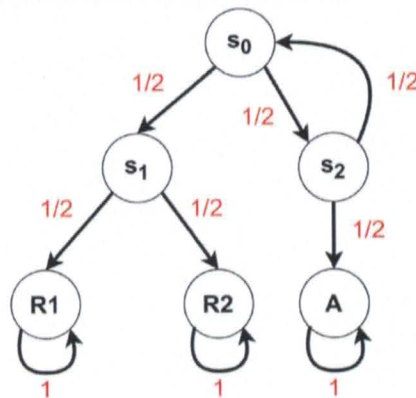
- $s \in T \Rightarrow x(s) = 1$
 - množina T je triviálně dosažitelná z prvku z množiny T
- $s \in N \Rightarrow x(s) = 0$
 - množina T je triviálně nedosažitelná z prvku z množiny N
- $s \in S \setminus (T \cup N) \Rightarrow x(s) = \sum_{u \in S} P(s, u) \cdot x(u)$
 - tedy uvažujeme o všech stavech u , které jsou jedнокrokově dosažitelné z s
 - pravděpodobnost tohoto přechodu pronásobíme neohraňčenou dosažitelností množiny T ze stavu u

Tímto způsobem vytvoříme soustavu $|S|$ lineárních rovnic, jde tedy o zpětnou propagaci pravděpodobností.

Příklad. Mějme dvě karty se stejným rubovým značením (*rub [revers] je ta část karty, na níž není poznat hodnota karty, část s hodnotou je líc [avers]*). První z karet je eso, zatímco druhá je špatně vytištěná, tedy i na lícové straně má stejný obrázek jako na rubové (má rub z obou stran). Jedna z karet je náhodně vytažena z váčku a položena na stůl tak, že není vidět její spodní strana. Pokud je na viditelné straně karty symbol esa (*vylosována byla první karta a byla položena lícem vzhůru*), je karta vrácena do balíčku a pokus se opakuje. Pokud na viditelné straně karty není symbol esa, karta je otočena. Pokud se po otočení ukáže, že bylo vylosováno eso, vyhrávám, jinak vyhrává soupeř. Jaká je pravděpodobnost, že ve hře zvítězím? Je hra spravedlivá?



Situaci můžeme modelovat pomocí Markovského řetězce:



Pokus začíná ve stavu s_0 . Buď vylosojeme špatně vytištěnou kartu (stav s_1) nebo eso (stav s_2). Pokud je na stole eso položené lícem vzhůru, vracíme se do stavu s_0 (opakujeme pokus novým losováním karty). Jinak kartu otočíme a odhalíme výsledek (stav R_1 a R_2 - bylo odhaleno, že vybraná karta je špatně vytištěná, stav A - bylo odhaleno, že karta je eso).

Chceme zjistit, jaká je pravděpodobnost mé výhry, tedy **počítáme neohraničenou dosažitelnost stavu A ze stavu s_0** , což značí mou výhru. Proto $T = \{A\}$ je množina cílů. Konstruujeme dále množinu $N = \{s_1, R_1, R_2\}$, tedy množinu stavů, z nichž není možné dosáhnout žádný stav v T . Sestavíme rovnice:

$$\begin{aligned}
 x(A) &= 1 \\
 x(s_1) &= 0 \\
 x(R_1) &= 0 \\
 x(R_2) &= 0 \\
 x(s_2) &= P(s_2, A) \cdot x(A) + P(s_2, s_0) \cdot x(s_0) \\
 x(s_0) &= P(s_0, s_1) \cdot x(s_1) + P(s_0, s_2) \cdot x(s_2)
 \end{aligned}$$

Dosazujeme:

$$x(s_2) = \frac{1}{2} + \frac{1}{2} \cdot x(s_0)$$

$$x(s_0) = \frac{1}{2} \cdot x(s_2)$$

Dále dosadíme $x(s_2)$ do druhé rovnice:

$$x(s_0) = \frac{1}{2} \cdot \left(\frac{1}{2} + \frac{1}{2} \cdot x(s_0) \right)$$

$$x(s_0) = \frac{1}{4} + \frac{1}{4} \cdot x(s_0)$$

$$\frac{3}{4} \cdot x(s_0) = \frac{1}{4}$$

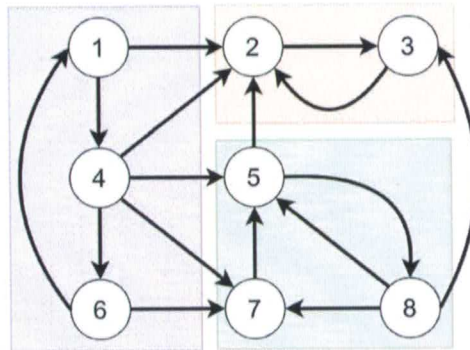
$$x(s_0) = \frac{1}{3}$$

Pravděpodobnost, že bude v libovolném (neomezeném) počtu kroků dosažen stav A , odpovídá $1/3$. Pravděpodobnost mé výhry tedy činí pouze $1/3$. Spravedlivost je subjektivní koncept, ovšem dle všeobecně uznávaných měřítek bychom řekli, že hra spravedlivá není.

Analýza ustáleného stavu

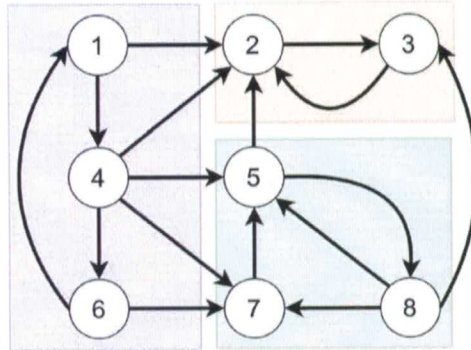
Definujeme následující pojmy:

- **silně souvislá komponenta grafu**
 - at' $G = (V, E)$ je orientovaný graf
 - **silně souvislá komponenta** (SCC) grafu G je taková množina vrcholů $U \subseteq V$, kde
 - $\forall u, v \in U : uE^*v$
 - tedy každé dva uzly z množiny U jsou v G vzájemně dosažitelné
 - U je \subseteq -maximální taková množina
 - tedy neexistuje vlastní nadmnožina U splňující podmínku o vzájemné dosažitelnosti všech uzlů
 - každý uzel grafu vždy leží právě v jedné silně souvislé komponentě grafu, tedy lze
 - vytvořit rozklad množiny uzlů V na silně souvislé komponenty grafu
 - definovat relaci ekvivalence *leží ve stejné silně souvislé komponentě* na množině uzlů
 - žádné dvě silně souvislé komponenty v grafu nejsou nikdy vzájemně dosažitelné
 - uvažujme o níže vyobrazeném orientovaném grafu

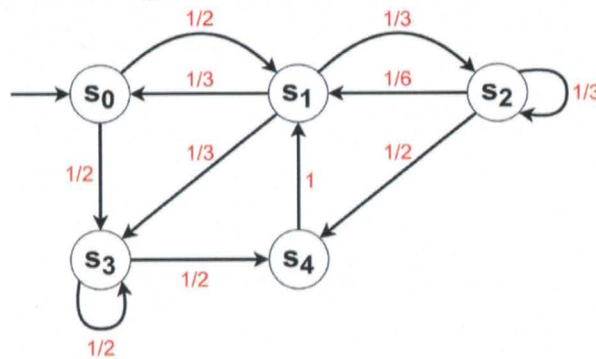


- graf obsahuje silně souvislé komponenty $\{1, 4, 6\}$, $\{2, 3\}$, $\{5, 7, 8\}$
 - všimněme si například, že všechny uzly v množině $\{5, 7, 8\}$ jsou vzájemně dosažitelné
 - všimněme si, že současně všechny uzly v množině $\{5, 8\}$ jsou vzájemně dosažitelné, ovšem množina $\{5, 8\}$ není silně souvislou komponentou, protože **není maximální** množinou splňující podmínku vzájemné dosažitelnosti, neboť $\{5, 8\} \subset \{5, 7, 8\}$
- **spodní silně souvislá komponenta grafu**
 - ať $G = (V, E)$ je orientovaný graf
 - ať $U \subseteq V$ je jeho silně souvislá komponenta
 - řekněme, že U je **spodní silně souvislá komponenta** (BSCC) pokud
 - $\forall u \in U : \forall v \in V \setminus U : \neg(uE^*v)$
 - tedy z vrcholů v silně souvislé komponentě U není dosažitelný žádný vrchol, který leží mimo tuto silně souvislou komponentu
 - pokud pomocí přechodů v grafu dospějeme do této silně souvislé komponenty, již neexistuje možnost, jak ji opustit, jsme v pasti
 - v každém orientovaném grafu musí vždy existovat alespoň jedna spodní silně souvislá komponenta
 - pokud je v grafu jediná silně souvislá komponenta, je triviálně spodní
 - pokud je v grafu více silně souvislých komponent, alespoň jedna musí být spodní, jinak by musely být některé silně souvislé komponenty vzájemně dosažitelné, což je spor
 - spodní silně souvislé komponenty mohou být klidně všechny silně souvislé komponenty grafu
 - např. u grafu bez jakýchkoliv hran
 - uvažujme o níže vyobrazeném orientovaném grafu

11/2

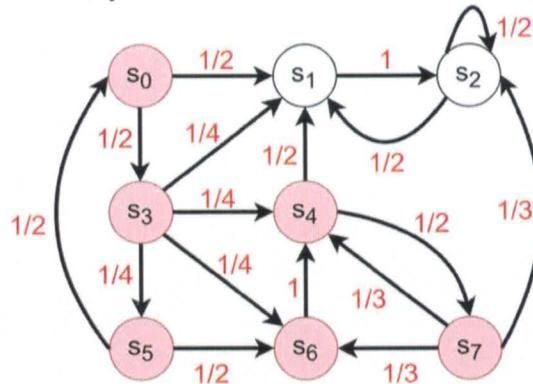


- tento graf má jedinou spodní silně souvislou komponentu {2, 3}
- z této silně souvislé komponenty není možné dosáhnout žádný uzel, který není součástí této silně souvislé komponenty
- z ostatních silně souvislých komponent {1, 4, 6}, {5, 7, 8} jsou dosažitelné uzly 2, 3, které v nich neleží, tedy tyto silně souvislé komponenty nejsou spodní
- v Markovských řetězcích bude v každém nekonečném běhu s pravděpodobností 1 někdy dosažena nějaká spodní silně souvislá komponenta
- **ireducibilní Markovský řetězec**
 - ať $M = (S, s_0, P)$ je Markovský řetězec
 - Markovský řetězec je **ireducibilní**, pokud celá množina S je jedinou silně souvislou komponentou tohoto Markovského řetězce
 - uvažujme o následujícím Markovském řetězci

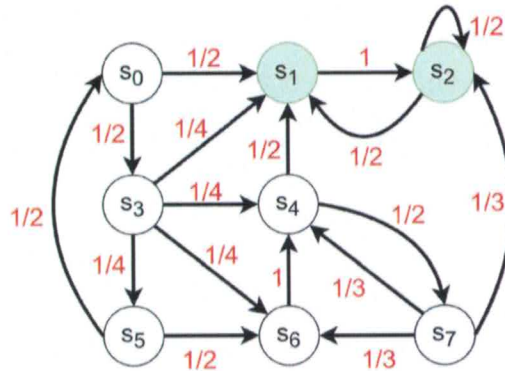


- všechny stavy v něm jsou vzájemně dosažitelné, tedy tvoří jedinou silně souvislou komponentu
- Markovský řetězec na obrázku je tedy ireducibilní
- **tranzientní stav Markovského řetězce**
 - ať $M = (S, s_0, P)$ je Markovský řetězec
 - stav $s \in S$ tohoto Markovského řetězce nazveme **tranzientním stavem**, pokud platí (následující podmínky jsou ekvivalentní):
 - pokud by v M byl stav s iniciální, pak $\lim_{n \rightarrow \infty} t_n(s) = 0$
 - tedy pravděpodobnost, že se "v nekonečno" budeme nacházet ve stavu s , je nulová
 - stav s bude navštíven nejvýše konečně mnohokrát

- s leží v některé silně souvislé komponentě Markovského řetězce M, která není spodní
- při posuzování, zda je stav s tranzientní, je skutečně třeba se zamyslet nad jeho chováním v případě, že by byl iniciální
- mohlo by se totiž stát, že existuje stav, který není z s_0 vůbec dosažitelný, ale přesto je součástí spodní silně souvislé komponenty, takže tranzientní být nemůže, ačkoliv by platilo $\lim_{n \rightarrow \infty} t_n(s) = 0$, neboť nikdy není z s_0 dosažen
- uvažujme o následujícím Markovském řetězci



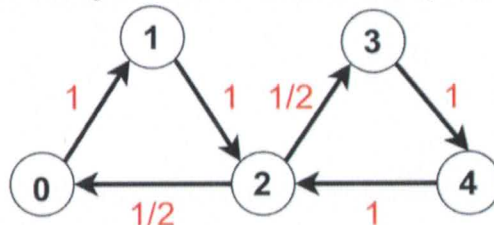
- stavy $s_0, s_3, s_4, s_5, s_6, s_7$ jsou zde tranzientní, neboť leží mimo spodní silně souvislé komponenty
- pravděpodobnost, že se v těchto stavech budeme nacházet v nekonečnu, je nulová
- "v nekonečnu z tohoto stavu odejde veškerá pravděpodobnost"
- **rekurentní stav Markovského řetězce**
 - ať $M = (S, s_0, P)$ je Markovský řetězec
 - stav $s \in S$ tohoto Markovského řetězce nazveme **rekurentním stavem**, pokud (následující podmínky jsou ekvivalentní):
 - pokud by v M byl stav s iniciální, pak $\lim_{n \rightarrow \infty} t_n(s) > 0$
 - tedy pravděpodobnost, že se "v nekonečnu" budeme nacházet ve stavu s, je určitě nenulová
 - stav může být navštíven nekonečně mnohokrát
 - s leží v některé spodní silně souvislé komponentě Markovského řetězce M
 - opět si uvědomme, že stav s teoreticky nemusí být dosažitelný z iniciálního stavu, tedy by platilo $\lim_{n \rightarrow \infty} t_n(s) = 0$, ovšem to ještě neznamená, že není rekurentní, protože podmínku s limitou vyhodnocujeme pro případ, že by stav s iniciální byl
 - uvažujme o následujícím Markovském řetězci



- stavy s_1, s_2 jsou zde rekurentní, neboť leží ve spodní silně souvislé komponentě
- pravděpodobnost, že se v těchto stavech budeme nacházet v nekonečno, je nenulová
- "ani v nekonečno z tohoto stavu neodteče veškerá pravděpodobnost"

• **periodický stav Markovského řetězce**

- ať $M = (S, s_0, P)$ je Markovský řetězec
- stav $s \in S$ je **periodický**, pokud $\text{NSD}\{n \in \mathbb{N}^+ \mid P^n(s, s) > 0\} > 1$
- P^n je zde mocnina matice P , $\text{NSD}(\cdot)$ je největší společný dělitel všech čísel v dané množině
- to lze vysvětlit následovně:
 - řekněme, že bych stav s považoval za iniciální v M a začal bych provádět tranzientní analýzu
 - zaznamenával bych si číslo každého kroku, v němž byla pravděpodobnost stavu s nenulová
 - takto bych vytvořil posloupnost všech diskretních časových okamžiků, v nichž existuje kladná pravděpodobnost, že se nacházíme ve stavu s
 - nyní bych našel největší společný dělitel celé této posloupnosti
 - pokud by byl vyšší než 1, pak pro libovolnou cestu v Markovském řetězci platí, že stav s se na této cestě indexované od nuly může nacházet pouze na násobcích tohoto největšího společného dělitele, nikdy ne mimo
 - pokud by se rovnal 1, pak vždy pro každé přirozené číslo n najdu v tomto Markovském řetězci takovou cestu, že se na ní stav s vyskytuje na pozici, která není násobkem n
- uvažujme o následujícím Markovském řetězci vyobrazeném níže

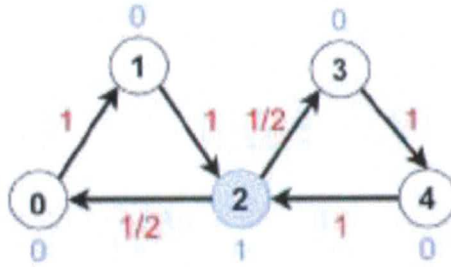


- stanovme $s_0 = 0$

- veškeré cesty v tomto Markovském řetězci budou zřejmě odpovídat řetězcům typu $012(012 + 342)^{\omega}$
 - zdá se, že na všech cestách bude stav 0 vždy na pozicích odpovídajících násobku čísla 3, nikdy jinde
 - např. máme cesty
 - 012012012012012012...
 - stav 0 je zde na pozicích 0, 3, 6, 9, 12, 15, 18,...
 - 012342342012342012342...
 - stav 0 je zde na pozicích 0, 9, 15,...
 - 012012342342342342012...
 - stav 0 je zde na pozicích 0, 3, 18,...
 - ať se tedy vydávám jakoukoliv cestou, nikdy se mi nestane, že by stav 0 byl na této cestě na pozici, jejíž index není dělitelný třemi
 - *toto samozřejmě bude platit jen tehdy, pokud se na stav, jehož periodicitu zkoumáme, díváme jako na iniciální stav*
- zkusme provést transientní analýzu pro několik prvních kroků

stav	0	1	2	3	4
T_0	1	0	0	0	0
T_1	0	1	0	0	0
T_2	0	0	1	0	0
T_3	1/2	0	0	1/2	0
T_4	0	1/2	0	0	1/2
T_5	0	0	1	0	0
T_6	1/2	0	0	1/2	0
T_7	0	1/2	0	0	1/2
T_8	0	0	1	0	0

- vektory pravděpodobností se od T_2 dále periodicky opakují s periodou 3
 - tento Markovský řetězec v čase nespěje k žádnému ustálenému stavu

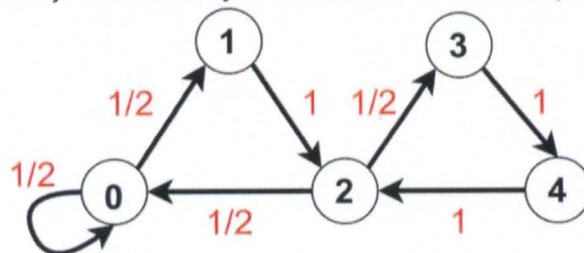


■ **stav 0 je v tomto Markovském řetězci periodický**

- platí totiž $\text{NSD}\{n \in \mathbb{N}^+ \mid P^n(0, 0) > 0\} = \text{NSD}\{0, 3, 6, 9, 12, 15, \dots\} = 3 > 1$

• **aperiodický stav Markovského řetězce**

- ať $M = (S, s_0, P)$ je Markovský řetězec
- stav $s \in S$ je **aperiodický**, pokud $\text{NSD}\{n \in \mathbb{N}^+ \mid P^n(s, s) > 0\} = 1$
- stav je aperiodický právě tehdy, pokud není periodický
- uvažujme o následujícím Markovském řetězci vyobrazeném níže



■ stanovme $s_0 = 0$

■ zřejmě pro každé přirozené číslo n najdeme cestu takovou, že stav 0 bude na této cestě právě na pozici n

■ stav 0 je tedy aperiodický

- všimněme si, že pokud nad stavem s existuje smyčka s nenulovou pravděpodobností, pak tento stav nikdy nemůže být periodický, neboť $P^1(s, s) > 0$, tedy vyhodnocujeme NSD pro množinu obsahující i číslo 1, což si jako největší společný dělitel vynutí právě číslo 1

• **periodický Markovský řetězec**

- ať $M = (S, s_0, P)$ je Markovský řetězec
- Markovský řetězec M je **periodický**, pokud alespoň jeden jeho stav je periodický

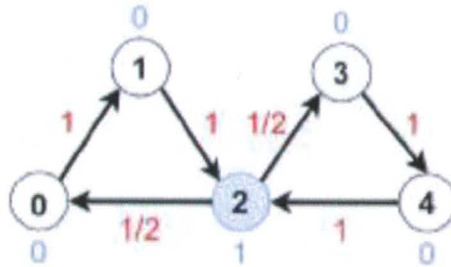
• **aperiodický Markovský řetězec**

- ať $M = (S, s_0, P)$ je Markovský řetězec
- Markovský řetězec M je **aperiodický**, pokud jsou všechny jeho stavy aperiodické

Ať $M = (S, s_0, P)$ je Markovský řetězec. Při analýze ustáleného stavu (*steady-state*) zkoumáme, jakým způsobem se ustálí pravděpodobnosti jednotlivých stavů v limitním případě. Ustálený stav by se tedy dal chápat jako vektor T_∞ , tedy vektor hodnot $t_\infty(s) = \lim_{n \rightarrow \infty} t_n(s)$ pro $s \in S$. Průběh takové analýzy závisí na typu daného Markovského řetězce.

• **analýza ustáleného stavu periodických Markovských řetězců**

- o jak jsme viděli výše, u periodických Markovských řetězců nemusí ustálený stav vůbec existovat



- o viděli jsme, že vektory T_i se stále opakovaly, tedy limitní vektor T_∞ neexistoval
- o nebudeme dále řešit 🙄🙄

• **analýza ustáleného stavu neperiodických (aperiodických) ireducibilních Markovských řetězců**

- o ať $M = (S, s_0, P)$ je Markovský řetězec, který je neperiodický a ireducibilní
- o takový Markovský řetězec tedy tvoří jedinou silně souvislou komponentu
- o ustálený stav v něm nijak nezávisí na tom, který stav je iniciální
- o lze řešit pomocí soustavy bilančních rovnic a normalizační rovnice

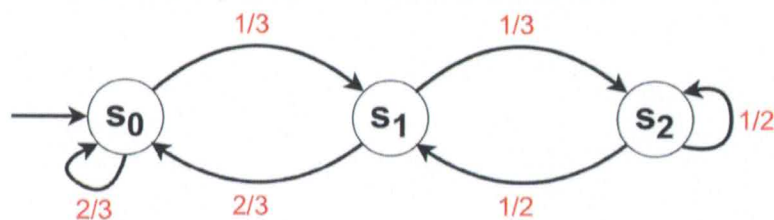
■ $t_\infty(s) = \sum_{v \in S} P(v, s) \cdot t_\infty(v)$

- bilanční rovnice
- tímto způsobem je sestaveno $|S|$ rovnic, kde $t_\infty(s)$ pro stavy s jsou jednotlivé neznámé
- rovnice vyjadřuje, že pravděpodobnost stavu s v nekonečnu závisí na pravděpodobnosti ostatních stavů v nekonečnu a na pravděpodobnostech, že stav s bude z ostatních stavů jednokrokově dosažen
- vyjadřujeme tím, "kolik pravděpodobnosti může v nekonečnu do stavu s vtéct"

■ $\sum_{v \in S} t_\infty(v) = 1$

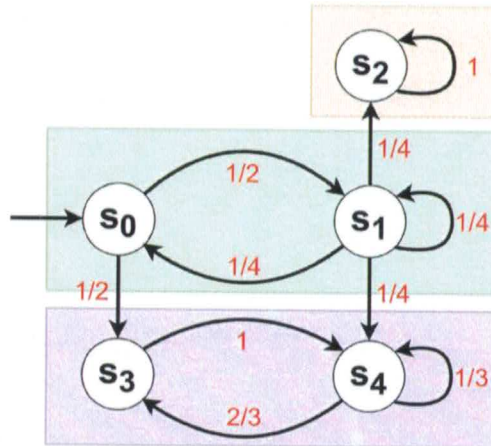
- normalizační rovnice, která vyjadřuje, že i v ustáleném stavu musí být pravděpodobnost v Markovském řetězci rozprostřena očekávatelným způsobem respektujícím axiomatickou definici pravděpodobnosti
- ostatní rovnice jsou na sobě lineárně závislé, tedy bez této rovnice typicky není možné najít právě jedno řešení soustavy

- o uvažujme o níže vyobrazeném Markovském řetězci



- vypočítejme jeho ustálený stav

- sestavíme následující soustavu rovnic:
 - $t_x(s_0) = 2/3 \cdot t_x(s_0) + 2/3 \cdot t_x(s_1)$
 - $t_x(s_1) = 1/3 \cdot t_x(s_0) + 1/2 \cdot t_x(s_2)$
 - $t_x(s_2) = 1/3 \cdot t_x(s_1) + 1/2 \cdot t_x(s_2)$
 - $t_x(s_0) + t_x(s_1) + t_x(s_2) = 1$
 - soustava má jediné řešení $[t_x(s_0), t_x(s_1), t_x(s_2)] = [6/11, 3/11, 2/11]$
- **analýza ustáleného stavu obecných neperiodických Markovských řetězců**
 - takové Markovské řetězce nemusí být složeny z jediné silně souvislé komponenty
 - tentokrát se mohou výsledky lišit v závislosti na iniciálním stavu
 - při analýze ustáleného stavu je nejprve nutné rozlišit jednotlivé silně souvislé komponenty a detekovat spodní silně souvislé komponenty, tedy rozlišit tranzientní a rekurentní stavy
 - **v ustáleném stavu bude pravděpodobnost tranzientních stavů vždy nulová**
 - z tranzientních stavů v nekonečnu vždy veškerá pravděpodobnost odečte
 - tak to dopadne vždy, a to neohledně na to, který stav byl v daném Markovském řetězci iniciální
 - **pokud je některý rekurentní stav nedosažitelný z iniciálního stavu, v ustáleném stavu bude jeho pravděpodobnost nulová**
 - **pokud je některý rekurentní stav dosažitelný z iniciálního stavu, jeho pravděpodobnost v ustáleném stavu bude nenulová**
 - stav je součástí spodní silně souvislé komponenty, do níž "vteče nějaká pravděpodobnost" a už ji nikdy neopustí
 - pro každou spodní silně souvislou komponentu provedeme analýzu neohraničené dosažitelnosti této komponenty
 - pro každou spodní silně souvislou komponentu provedeme analýzu ustáleného stavu, přičemž se na tuto komponentu díváme jako na samostatný aperiodický ireducibilní Markovský řetězec
 - zjištěné pravděpodobnosti ustáleného stavu spodních silně souvislých komponent pronásobíme pravděpodobností, že bude tato komponenta dosažena z iniciálního stavu
 - mějme níže zobrazený Markovský řetězec



- v Markovském řetězci jsou již vyznačeny jednotlivé silně souvislé komponenty
 - $\{s_2\}$ je spodní silně souvislá komponenta
 - $\{s_0, s_1\}$ není spodní silně souvislá komponenta
 - $\{s_3, s_4\}$ je spodní silně souvislá komponenta
- provedme nyní analýzu ustáleného stavu spodních silně souvislých komponent
 - uvažujme o komponentě $\{s_2\}$ jako o **samostatném** Markovském řetězci
 - $t_\infty(s_2) = t_\infty(s_2)$
 - $t_\infty(s_2) = 1$
 - z dvou výše uvedených rovnic plyne $t_\infty(s_2) = 1$
 - uvažujme o komponentě $\{s_3, s_4\}$ jako o **samostatném** Markovském řetězci
 - $t_\infty(s_3) = 2/3 \cdot t_\infty(s_4)$
 - $t_\infty(s_4) = t_\infty(s_3) + 1/3 \cdot t_\infty(s_4)$
 - $t_\infty(s_3) + t_\infty(s_4) = 1$
 - soustava má jediné řešení $[t_\infty(s_3), t_\infty(s_4)] = [2/5, 3/5]$
- analyzujme neohraničenou dosažitelnost komponenty $\{s_2\}$
 - $T = \{s_2\}$
 - $N = \{s_3, s_4\}$
 - $x(s_2) = 1$
 - $x(s_3) = x(s_4) = 0$
 - $x(s_0) = P(s_0, s_3) \cdot x(s_3) + P(s_0, s_1) \cdot x(s_1)$
 - $x(s_1) = P(s_1, s_0) \cdot x(s_0) + P(s_1, s_1) \cdot x(s_1) + P(s_1, s_2) \cdot x(s_2) + P(s_1, s_4) \cdot x(s_4)$
 - zde bychom dospěli k hodnotě $x(s_0) = 1/5$
- analyzujme neohraničenou dosažitelnost komponenty $\{s_3, s_4\}$

- $T = \{s_3, s_4\}$
- $N = \{s_2\}$
- $x(s_2) = 0$
- $x(s_3) = x(s_4) = 1$
- $x(s_0) = P(s_0, s_3) \cdot x(s_3) + P(s_0, s_1) \cdot x(s_1)$
- $x(s_1) = P(s_1, s_0) \cdot x(s_0) + P(s_1, s_1) \cdot x(s_1) + P(s_1, s_2) \cdot x(s_2) + P(s_1, s_4) \cdot x(s_4)$
- zde bychom dospěli k hodnotě $x(s_0) = 4/5$
 - je zřejmé, že neohraničená dosažitelnost všech spodních silně souvislých komponent musí sumovat do jedničky, $1/5 + 4/5 = 1$

- ustálený stav Markovského řetězce tedy odpovídá vektoru T_* , který lze vyjádřit jako

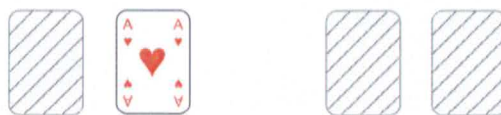
$$[t_*(s_0), t_*(s_1), t_*(s_2), t_*(s_3), t_*(s_4)] = [0, 0, 1/5 \cdot 1, 4/5 \cdot 2/5, 4/5 \cdot 3/5] = [0, 0, 1/5, 8/25, 12/25]$$

Analýza očekávaného počtu kroků

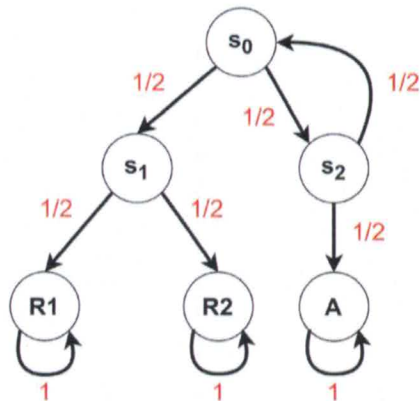
Ať $M = (S, s_0, P)$ je Markovský řetězec, $T \subseteq S$ je množina stavů. Upravme zobrazení P tak, že $P(s, s) = 1$, pokud $s \in T$, tedy stavy z množiny stavů T jsou nově absorbující. **Očekávaný počet kroků** k dosažení některého ze stavů z množiny T ze stavu s popisuje zobrazení $e : S \rightarrow \mathbb{R} \cup \{\infty\}$ definované následovně:

- $s \in T \Rightarrow e(s) = 0$
 - pokud je stav s z množiny T , není třeba z něj vykonat žádný krok, abychom se v množině T ocitli
- $s \in S \setminus T \Rightarrow e(s) = 1 + \sum_{v \in S} P(s, v) \cdot e(v)$
 - pokud stav s není z množiny T , přičítáme 1 ke každému vykonanému kroku, přičemž zpětně propagujeme očekávaný počet kroků z dalších stavů

Příklad. Uvažujme o hře se dvěma kartami popsané v sekci o neohraničené dosažitelnosti. Spočtěme očekávaný počet úkonů nutných k dokončení hry (mezi úkony počítáme vylosování karty, vrácení karty, otočení karty).



Situaci modelujeme pomocí Markovského řetězce:



Mějme množinu $T = \{R_1, R_2, A\}$. Zavedeme rovnice:

$$e(R_1) = 0$$

$$e(R_2) = 0$$

$$e(A) = 0$$

$$e(s_1) = 1 + P(s_1, R_2) \cdot e(R_2) + P(s_1, R_1) \cdot e(R_1)$$

$$e(s_2) = 1 + P(s_2, A) \cdot e(A) + P(s_2, s_0) \cdot e(s_0)$$

$$e(s_0) = 1 + P(s_0, s_1) \cdot e(s_1) + P(s_0, s_2) \cdot e(s_2)$$

Výsledkem je vektor $[e(s_0), e(s_1), e(s_2), e(R_1), e(R_2), e(A)] = [8/3, 1, 7/3, 0, 0, 0]$. Pro ukončení hry bude třeba provést průměrně $8/3$ úkonů.

Pokud v textu najdete chybu, nebudete něčemu rozumět nebo budete mít dojem, že by bylo vhodné něco doplnit, kontaktujte na discordu uživatele kocotom.