Kompakttutorium Theorie: Modelle für Rechensysteme, Theoretische Grundlagen der Informatik

Michael Köhler

6. Oktober 2000

1 Netz-Theorie

1. Erklären Sie folgende Begriffe: Präzedenzrelation, Kausalnetz, li, co, unabhängige Relation, Klique, Elementarüberdeckung, Bezirk, Linie, Schnitt, sequentielles Kausalnetz!
2. Wie kann ein endlicher Automat durch ein Petrinetz beschrieben werden?
3. Die Relationen li und co sind für Kausalnetze definiert. Warum nicht für ST-Netze?
4. Geben Sie die formalen Definitionen wieder: S/T-Netz, B/E-Netz, Markierung, Aktivierung, Schalten, Erreichbarkeitsmenge, Feuerfolgen, schlichtes Netz!
5. Definieren Sie die Begriffe und beschreiben Sie die Zusammenhänge zwischen ihnen: Erreichbarkeits-, Fall- und Überdeckungsgraph!
6. Geben Sie den Überdeckungsgraph des folgenden Netzes in Abb. 1 an!

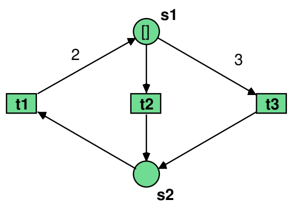

Abbildung 1: Zum Überdeckungsgraph

7. Ist in folgenden Netzen in Abb. 2, die Transition t aktiviert nach (a) der Schaltregel nach Valk? (b) nach der Schaltregel von Reisig, wenn die Stelle s_2 die Kapazität $K(s_2) = 1$ besitzt?
8. Unterscheiden sich die beiden Schaltregeln für Netze ohne Nebenbedingungen?
9. Sei $N = (S, T, F, W, K, m)$ ein S/T Netz mit der Kapazitätsbeschränkung $K(s) = 1$ für alle Stellen $s \in S$. Was gilt dann für Transitionen $t \in T$ mit der Eigenschaft $t^\circ \cap {}^\circ t \neq \emptyset$, d.h. Transitionen mit Nebenbedingungen?

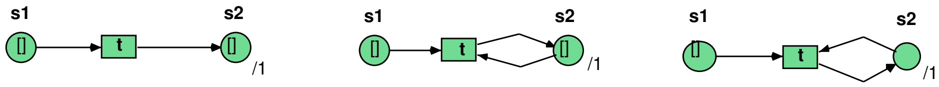

Abbildung 2: Zur Schaltregeln

10. Ist es entscheidbar, ob eine Markierung m von einem B/E-Netz $N = (S, T, F, m_0)$ erreichbar ist, d.h. ob $m \in R(N)$ gilt? Wenn ja: warum?

11. Ist entscheidbar, ob $m \in R(N)$ für ein S/T-Netz $N = (S, T, F, W, m_0)$ ohne Kapazitätsbeschränkungen gilt? Wenn ja: Warum?

12. Ist es mit Hilfe des Überdeckungsgraphen entscheidbar, ob eine Markierung m von einem S/T-Netz $N = (S, T, F, W, m_0)$ erreichbar ist, d.h. ob $m \in R(N)$ gilt? Wenn ja: warum? Wenn nein: was ist dann am Überdeckungsgraphen abzulesen?

13. Erklären Sie folgende Begriffe: asynchroner / synchroner Prozeß!

14. Welche Bedeutung haben S-Schnitte für Prozesse?

15. Was bedeutet es, daß zwei Transitionen eines asynchronen Prozesses (a) nebenläufig, (b) sequentiell sind?

16. Was bedeutet es, daß zwei Transitionen eines synchronen Prozesses (a) kollateral, (b) seriell sind?

17. Konstruieren Sie einen Prozeß des folgenden Netzes in Abb. 3! Welche Ereignisse sind nebenläufig?

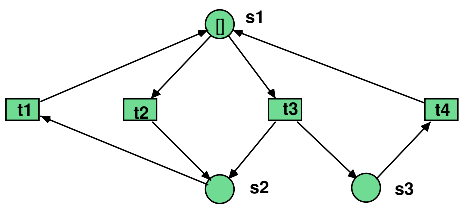

Abbildung 3: Zur Prozeßkonstruktion

18. Wann ist ein Netz kontaktfrei? Beweisen oder widerlegen Sie folgende Aussage! „Ein S/T-Netz ohne Kapazitäten ist stets kontaktfrei.“

19. Was ist eine Komplementärstelle? Konstruieren Sie ein S/T Netz $\bar{N}$ ohne Kapazitäten das äquivalent zu folgendem Netz in Abb. 4 ist!

20. Beschreiben Sie die Zusammenhänge zwischen folgenden Begriffen: S/T-Netz mit Transitionsinschrift, FES, passendes Auftragsystem, kompatible Feuerfolgen!

    Beschreibe, wie im Gegensatz dazu ein Auftragssystem in einem AV-Netz bearbeitet wird!

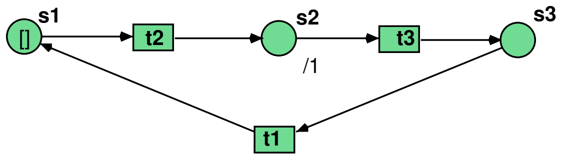

Abbildung 4: Zur Konstruktion der Komplementärstelle

21. Beschreiben Sie folgende Aspekte zur Strukturierung von Rechensystemen: Algorithmus von Kruskal, Bindungsnetz, konfliktfreie Partition!

2 Nebenläufigkeit

1. Erklären Sie das Konzept des kritischen Abschnitts und geben Sie ein Beispiel an! Setze hierzu folgende Begriffe in Beziehung: busy waiting, deadlock livelock, atomare Handlung, Semaphor, Speicher-, Rendezvous-Synchronisation!

2. Beschreiben Sie den Decker-Algorithmus!

3. Wie läßt sich der Wechselseitige Ausschluß mit Hilfe eines Petrinetzes (einfach) darstellen?

4. Erkläre: schematisches Auftragssystem, vollständiges AS, Interpretation eines AS, Zustandsfolge des interpretierten AS, Äquivalenz zweier Ausführungsfolgen!

5. Wann heißt eine Ausführungsfolge (a) seriell, (b) Serialisierung, (c) serialisierbar oder (d) schwach serialiserbar?

6. Was ist der Unterschied zwischen der „liest-von"-Relation und der Werteübertragungsrelation?

7. Was bedeutet es anschaulich, das ein Auftrag relevant (bzw. nutzlos) ist?

8. Geben Sie folgendes Theorem auswendig wieder und erläutern Sie es!

   Theorem 1 *Zwei Ausführungsfolgen $w_1, w_2 \in F(AS)$ eines schematischen Auftragssystems sind genau dann* äquivalent, *wenn für ihre Vervollständigungen w'_1, w'_2 gilt:*

   (a) w'_1, w'_2 haben die gleichen Mengen von relevanten Aufträgen und

   (b) für alle relevanten Aufträge a_i, a_j und jede Variable $v \in V$ gilt: a_j liest v von a_i in w'_1 gdw. a_j liest v von a_i in w'_2.

9. Zwei Ausführungsfolgen heißen äquivalent, falls sie für *alle* Interpretationen das gleiche Ergebnis liefern (s.o). Äquivalenz ist also *semantisch* definiert. Es existieren jedoch unendliche viele Interpretationen, so daß man die Äquivalenz zweier Folgen nicht direkt beweisen kann, indem man testet, ob sie für jede Intepretation gleich sind. Allenfalls kann man so beweisen, daß sie nicht äquivalent sind.

Theorem 1 liefert jedoch ein rein *syntaktisches* Kriterium, mit dem man Äquivalenz überprüfen kann. Erklären Sie, (a) warum das Theorem nur relevante Aufträge betrachtet, (b) warum die Gleichheit der „liest-von“-Relation die Gleichheit des Ergebnisses bezüglich der Herbrand-Interpretation bedeutet und (c) warum die Gleichheit bezüglich der Herbrand-Interpretation die Gleichheit bezüglich aller Interpretationen impliziert!

10. Erkläre: virtuell serialisierbar, Serialisierungsgraph, zweiphasige Ausführungsfolge!

11. Geben Sie folgendes Theorem auswendig wieder:

    Theorem 2 *Eine Ausführungsfolge w_1 ist genau dann serialisierbar, wenn mindestens ein Serialisierungsgraph aus $MSG(w_1)$ ein Auftragssystem ist.*

12. Ist es entscheidbar, ob eine Ausführunsgfolge serialisierbar ist?

13. Definiere: Funktionalität, Spur, spuräquivalent, -funktional, störungsfrei, Bernsteinrelation, verlustfrei, maximal nebenläufig für Funktionalität.

14. Serialisierbarkeit fordert, daß eine Ausführungsfolge äquivalent zu einer Serialisierung ist. Für welche Anwendungen ist dies ausreichend?

15. Funktionalität fordert, daß alle Ausführungsfolge äquivalent sind. Für welche Anwendungen ist dies sinnvoll?

16. Spurfunktionalität fordert, daß alle Spuren äquivalent sind. Für welche Anwendungen ist dies notwendig?

17. Geben Sie folgendes Theorem wieder und erklären Sie es!

    Theorem 3 *Sei AS ein vollständiges Auftragssystem.*

    (a) AS ist genau dann funktional, *wenn alle relevanten Aufträge paarweise störungsfrei sind.*

    (b) AS ist genau dann spurfunktional, *wenn alle verlustfreien Aufträge paarweise störungsfrei.*

18. Warum ist es für Funktionalität ausreichend, nur die relevanten Aufträge zu betrachten?

19. Warum ist es für Spurfunktionalität ausreichend, nur die verlustfreien Aufträge zu betrachten?

3 Invarianten

1. Definiere: Inzidenz-, Wirkungsmatrix eines Netzes!

2. Welchen Wertebereich besitzen S-Invarianten, welchen T-Invarianten?

3. Was ist eine Invariantengleichung? Was ist eine lineare Invariantengleichung?

4. Geben sie folgendes Theorem wieder!

Theorem 4 (Satz von Lautenbach) *Sei i eine Stellen-Invariante. Dann gilt für alle erreichbaren Markierungen $\forall m \in R(m_0)$: $i^T m = i^T m_0$.*

5. Beschreiben Sie, wie man mit Hilfe des Satzes von Lautenbach aus einer S-Invariante eine Invariantengleichung gewinnt!
6. Folgt jede Netzinvariante aus einer S-Invariante?
7. Geben Sie folgendes Theorem wieder!

Theorem 5 *Sei w eine Schaltfolge, die m_1 in m_2 überführt und j eine T-Invariante. Kommt jede Transition t in der Schaltfolge w genau $j(t)$-mal vor, so sind beide Markierungen gleich:*

$$(m_1[w\rangle m_2 \wedge \forall t \in T : j(t) = |w|_t) \Longrightarrow m_1 = m_2$$

8. Ist folgende Aussage richtig, z.B. als Implikation des obigen Theorems? Begründen Sie!

„Sei j eine T-Invariante. Dann überführt die Schaltfolge w

$$w = t_1^{j(t_1)} \cdots t_n^{j(t_n)}$$

die Markierung m_0 wieder in m_0, d.h. $m_0[w\rangle m_0$.“

9. Was ist die Bedeutung des Parikhvektors einer Schaltfolge?
10. Geben Sie zwei Invarianten des Netzes aus Abb. 4 an!

4 Verklemmungen

1. Erkläre: Verklemmung, sichere Markierung, ω-sicher, ω-fortsetzbar, Residuum, abgeschlossene Mengen, K-eingeschränkte Feuerfolgen, K-gesteuertes Netz
2. Beschreiben Sie das Bankiersproblem und seine Lösung nach Dikstra! Benutzen Sie dabei die Begriffe der vorangegangenen Aufgabe!
3. Gebe die Definitionen für Fortsetzbarkeit, Lebendigkeit, Fairness wider!
4. Ist Fortsetzbarkeit eines Netzes abgeschlossen gegen eine Markierungserhöhung? Und Lebendigkeit?
5. Es existieren zwei Definitionen für Lebendigkeit: Ein Netz N heißt lebendig in der Markierung m_0, falls

$$\forall t \in T : \forall m \in R(N, m_0) : \exists m' \in R(N, m) : m'[t\rangle$$

bzw.

$$\forall t \in T : \forall m \in R(N, m_0) : \exists w \in T^\omega m_0[w\rangle \wedge |w|_t = \infty$$

Zeige die Äquivalenz beider Definitionen!

6. Diskutiere die Begriffe anhand der Ähnlichkeiten in den Definitionen!

   Lebendigkeit:

   $$\forall t \in T : \forall m \in R(N, m_0) : \exists w \in T^\omega m_0[w\rangle \wedge |w|_t = \infty$$

   Fairness:

   $$\forall t \in T : \forall m \in R(N, m_0) : \forall w \in T^\omega m_0[w\rangle \Longrightarrow |w|_t = \infty$$

7. Beweise folgendes Theorem!

   Theorem 6 *In einem S/T-Netz ist eine Markierung m genau dann lebendig, wenn alle von m erreichbaren Markierungen fortsetzbar sind.*

8. Beweise mit Hilfe des obigen Theorem folgende Aussage!

   Theorem 7 *Das CONT-gesteuerte Netz N' zu einem Netz N verhält sich wie der lebendige Teil des Netze N.*

9. Definiere: verschleppungsfreie Schaltregel, faire Schaltregel. Welche ist leichter zu implementieren? Warum?

10. Ist ein Netz, das nach der fairen Schaltregel feuert fair? Ist ein lebendiges Netz unter der der fairen Schaltregel fair? Geben Sie einen Beweis oder ein Gegenbeispiel an!

11. Erklären sie den Algorithmus für den verteilten, wechselseitigen Ausschluß!

12. Definiere das Lebendigkeits- und das Erreichbarkeitsproblem! Ist das Lebendigkeitsproblem entscheidbar? Ist das Erreichbarkeitsproblem entscheidbar? Stehen die Probleme in Beziehung zueinander?

13. Welche Mindestlaufzeit besitzt ein Algorithmus für das Erreichbarkeitsproblem?

14. Das Beschränktheitproblem benötigt mindestens $O(2^{c\sqrt{n}})$ und höchstens $O(2^{c|S|\log|S|})$ viel Platz für S/T-Netze. Mit welchem Problem ist es aufgrund der Laufzeit vergleichbar?

5 Funktionseinheiten

5.1 Allgemeine Definitionen

1. Rekapituliere folgende Definitionen!

Füllung	f
Kapazität	k
rel. Füllung	$\phi := f/k$
Verweilzeit	$y := t_2 - t_1$
Durchsatz	$d := \frac{n}{t_2 - t_1}$, wobei n = Anz. der Auftragsende-Ereignisse
Grenzdurchsatz	$c = 1/b$
Auslastung	$\rho := d/c = d \cdot b$
Bedienzeit	$b :=$ Verweilzeit bei Füllung $f = 1$
Besuchszahl	$v_{ij} :=$ Zahl der Aufträge der Klasse j in der FE i

Spezialfall: Einfache Funktionseinheit

Für den Fall einer einfachen FE ergibt sich:

Belegzeit	$x := \sum_1^n y_i$
mittlere Verweilzeit	$\bar{y} = x/n = \bar{b}$
Grenzdurchsatz	$c = 1/\bar{y} = n/x = 1/\bar{b}$
Auslastung	$\rho := d/c = d \cdot b = \frac{x}{t_2 - t_1}$
mittlere Bedienzeit	$\bar{b} = x/n$
KDV	$KDV_{FE} := K_{FE} \cdot \bar{b}$
mittlere Kosten je Auftrag	$\bar{q} = \frac{K_{FE} \cdot \bar{b}}{\rho} = \frac{K_{FE} \cdot \bar{y}}{\rho}$

5.2 Stochastische Grundlagen

1. Erkläre die grundlegenden stochastischen Definitionen: Zufallsvariable T, Dichte f, Verteilung F, Erwartungswerte $E[T]$, Streuung σ, Varianz σ^2, Momente $E[T^n]$.

2. Erkläre die Gleichungen!

$$f(t) = \frac{\mathrm{d}F}{\mathrm{d}t} \qquad \Longleftrightarrow \qquad F(t) = \int_{-\infty}^{t} f(\tilde{t})\mathrm{d}\tilde{t}$$

$$\begin{aligned} E[T] &:= \int_{-\infty}^{\infty} t f(t)\mathrm{d}t \\ E[T^n] &:= \int_{-\infty}^{\infty} t^n f(t)\mathrm{d}t \\ \sigma^2 &:= \int_{-\infty}^{\infty} (t - E[T])^2 f(t)\mathrm{d}t \\ &= E[T^2] - E^2[T] \end{aligned}$$

5.3 Stochastisches Modell für Ausfallwahrscheinlichkeiten

Rekapitulieren wir einige grundlegende Ergebnisse.

- Die Wkt., daß ein Ereignis zum Zeitpunkt T im Intervall $(t, t + \Delta t]$ stattfindet, beträgt:

$$P(T \leq t + \Delta t | T > t) = \frac{F(t + \Delta t) - F(t)}{1 - F(t)}$$

- Enderate $\theta(t)$:

$$\theta(t) := \frac{\mathrm{d}}{\mathrm{d}t} P(T \leq t + \Delta t | T > t) = \frac{f(t)}{1 - F(t)}$$

- Eine konstante Enderate, d.h. eine gedächtnislose Vtlg., erzeugt die Exp.-Vtlg.:

$$\theta(t) = \theta = \text{const.} \Longrightarrow F(t) = 1 - e^{-\theta t}$$

- Für die Exp.-Vtlg. gilt, daß die Anzahl der Ereignisse im Zeitraum d Poisson verteilt ist:

$$P(M = m) = \frac{(\theta d)^m}{m!} e^{-\theta d} = \pi(\theta d)$$

- Erwartungswert der Restzeit / Rekurrenzzeit:

$$E[RT] = \frac{E[T^2]}{2E[T]} = \frac{E[T]}{2} + \frac{\sigma^2}{2E[T]}$$

- Der Erwartungswert der Intervallänge zwischen zwei Ereignissen ist größer als der Erwartungswert des Endereignisses:

$$E[T'] = 2E[RT] = ET + \frac{\sigma^2}{ET} \geq E[T]$$

5.3.1 Asynchrone Poissonprozesse

Mehrere unabhängige, neg-exp. verteilte Ereignisse sind insgesamt ebenfalls neg-exp. verteilt:

$$\begin{aligned} P(T \leq t) &= 1 - P(T > t) = 1 - P(\bigcap T_i > t) \\ \text{unabhängig} &= 1 - \prod P(T_i > t) = 1 - \prod e^{-\theta_i t} \\ &= 1 - e^{\sum -\theta_i t} = 1 - e^{-\theta t} \end{aligned}$$

5.4 Zuverlässigkeiten von Strukturen

Sei L die Lebensdauer der Struktur, $r(t)$ die Überlebenswkt. und $a(t))$ die Ausfallwkt. einer Struktur:

$$r(t) = P(L > t) \qquad a(t) = P(L \leq t) = 1 - r(t)$$

Für den Spezialfall der Exp-vtlg. $r_i(t) = e^{-\lambda_i t}$ ergibt sich:

- Serienstruktur:

$$r(t) = \prod_i r_i(t) = e^{-\sum \lambda_i t}$$

- Parallelstruktur:

$$r(t) = 1 - \prod_i a_i(t) = 1 - \prod_i 1 - r_i(t) = 1 - \prod_i (1 - e^{-\lambda_i t})$$

  Für den Sonderfall gleicher Ausfallraten λ ergibt sich

$$r(t) = 1 - (1 - e^{-\lambda_i t})^n$$

  und für den Erwartungswert

$$E[L] = \int_0^\infty r(t)\mathrm{d}t = \frac{1}{\lambda} \sum_{i=1}^n \frac{1}{i}$$

- k aus n Struktur:

$$r(t) = \sum_{j=k}^n \binom{n}{j} r(t)^j (1 - r(t))^{n-j}$$

  Für den Fall der 2 aus 3 Struktur ergibt sich:

$$r(t) = 3r(t)^2(1 - r(t)) + r(t)^3 = 3e^{-2\lambda_i t} - 2e^{-3\lambda_i t}$$

  und für den Erwartungswert

$$E[L] = \int_0^\infty r(t)\mathrm{d}t = \frac{3}{2\lambda} - \frac{2}{3\lambda} = \frac{5}{6\lambda}$$

6 Markov-Ketten

1. *Markov-Eigenschaft*: Die Wkt.-Vtlg. für den Folgezustand hängt nur vom derzeitigen Zustand ab:
$$P(S(t+\Delta t)=j|S(t)=i \wedge S(t-a)=k) = P(S(t+\Delta t)=j|S(t)=i)$$
2. Ein *homogener* Markov-Prozeß liegt vor, falls die Übergangswkt. nicht vom Zeitpunkt t abhängt:
$$P(S(t+\Delta t)=j|S(t)=i) = P_{ij}(\Delta t)$$
3. *Zustandsübergangsrate* λ_{ij} (im [Bause/Kritzinger] als q_{ij} bezeichnet):
$$\lambda_{ij} = \lim_{\Delta t \to 0} \frac{P_{ij}(\Delta t)}{\Delta t}$$

6.1 Homogener, zustandsdiskreter Markov-Prozeß

$$\begin{aligned}
p_i(t+\Delta t) &= \sum_j p_j(t) P_{ji}(\Delta t) \\
&= \sum_{j\neq i} p_j(t) P_{ji}(\Delta t) + p_i(t) P_{ii}(\Delta t) \\
&= \sum_{j\neq i} p_j(t) P_{ji}(\Delta t) + p_i(t)(1 - \sum_{j\neq i} P_{ij}(\Delta t)) \\
&= \sum_{j\neq i} p_j(t) P_{ji}(\Delta t) + p_i(t) - p_i(t) \sum_{j\neq i} P_{ij}(\Delta t))
\end{aligned}$$

Dies ergibt mit der Grenzwertbildung eine Beziehung zur Ableitung $\dot{p}(t)$:

$$\begin{array}{rrcl}
 & \frac{p_i(t+\Delta t)-p_i(t)}{\Delta t} & = & \sum_{j\neq i} p_j(t) \frac{P_{ji}(\Delta t)}{\Delta t} - p_i(t) \sum_{j\neq i} \frac{P_{ij}(\Delta t))}{\Delta t} \\
\Delta t \to 0: & \dot{p}_i(t) & = & \sum_{j\neq i} p_j(t)\lambda_{ji} - p_i(t) \sum_{j\neq i} \lambda_{ij}
\end{array}$$

Für eine konvergierende Zustandsverteilung, d.h. falls $\lim p_i(t) = \pi_i$ existiert, läßt sich die obige Differentialgleichung einfacher lösen, da sie sich wegen $\dot{p}_i(t) = 0$ wie folgt vereinfacht:

$$\sum_{j\neq i} \pi_j \lambda_{ji} = \pi_i \sum_{j\neq i} \lambda_{ij}$$

1. Beschreibe, wann ein solche konvergierende Zustandsverteilung eintritt! Erkläre dazu: Absorptionszustand, irreduzibler Zustandsraum, stationärer Grenzprozeß.

Machine-repairman-Modell

Ein einfaches Beispiel für eine derartig zu analysierende Markovkette ist in Abb. 5 dargestellt.

Es ergibt sich für diese konkrete Markov-Kette folgendes Gleichungssystem:

$$\begin{aligned}
p_0 n\lambda &= p_1 \beta \\
p_1((n-1)\lambda + \beta) &= p_0 n\lambda + p_2\beta \\
&\vdots \\
p_i((n-i)\lambda + \beta) &= p_0(n-i+1)\lambda + p_{i+1}\beta \\
&\vdots \\
p_n\beta &= p_{n-1}\lambda
\end{aligned}$$

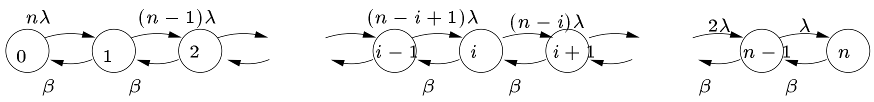

Abbildung 5: Machine-repairman-Modell

Diese Gleichungen besitzen eine anschauliche Beschreibung: Im Gleichgewicht ist der Fluß aus einem Knoten heraus genauso groß wie der Fluß in den Knoten hinein.

Mit der Abkürzung $\phi = \lambda/\beta$ erhalten wir durch Umformungen:

$$p_i = \frac{n!}{(n-i)!}\phi^i p_0$$

Aufgrund der stochastischen Normierung ergibt sich p_0:

$$1 = \sum_0^n p_j = \sum_0^n \frac{n!}{(n-j)!}\phi^j p_0 \Longrightarrow p_0 = \left(\sum_0^n \frac{n!}{(n-i)!}\phi^i\right)^{-1}$$

1. Erkläre die Bedeutung der Kanteninschriften in Abb. 5 : Was bedeutet λ, was β? Welche Bedeutung haben die Zustände?
2. Was bedeutet Flußgleichgewicht für Markovketten?
3. Wieso ist es zulässig, die Überlagerung des Ausfalls von n Komponenten (jeweils mit der Rate λ) durch den Ausfall *einer* Komponente mit der Rate $n\lambda$ zu beschreiben?

7 Auftragsverkehr

Zugänge in $[t_0, t)$	$n_z(t)$
Abgänge in $[t_0, t)$	$n_a(t)$
Füllung	$f(t) := n_z(t) - n_a(t)$
Zugänge	$z(t_1, t_2) := \frac{n_z(t_2) - n_z(t_1)}{t_2 - t_1}$
Durchsatz	$a(t_1, t_2) := \frac{n_a(t_2) - n_a(t_1)}{t_2 - t_1}$
Flußgleichgewicht	gdw. $z(t_1, t_2) = a(t_1, t_2)$

Theorem 8 (Gesetz von Little) *Der Mittelwert der Füllung ist das Produkt aus Durchsatz und mittlerer Verweilzeit:*

$$\bar{f} = d\bar{y}$$

Beweis: Jeder Auftrag A_i leistet einen Beitrag von $f \cdot t = 1 \cdot y_i$ zur Füllung, so daß gilt:

$$\begin{aligned}
\bar{y}(t_1, t_2) &= \frac{1}{n_z(t_2) - n_z(t_1)} \sum_{n_z(t_1)+1}^{n_z(t_2)} y_i \\
&= \frac{1}{n_z(t_2) - n_z(t_1)} \int_{t_1}^{t_2} f(t)\mathrm{d}t \\
&= \frac{t_2 - t_1}{n_z(t_1) - n_z(t_2)} \bar{f}(t_1, t_2)
\end{aligned}$$

8 Wartenetze und -systeme

1. Erkläre: Wartesystem, Verlustsystem, $M/G/m/k$-Systeme, geschlossenes System, offenes System!

2. Warum gilt für $-/-/m/k$-Systeme $m \leq k$?

Für Exp.-Vtlg mit den Parameter λ bzw. μ ergeben sich folgende Erwartungswerte:

Zwischenankunfszeit A	$E[A] = \frac{1}{\lambda}$
Zugang Z	$E[Z] = \frac{t_2 - t_1}{E[A]} \frac{1}{t_2 - t_1} = \lambda$
Durchsatz D	$E[D] = \lambda$
Grenzdurchsatz C	$E[C] = \mu$
Bedienzeit B	$E[B] = \frac{1}{\mu} = \frac{1}{E[C]}$

8.1 Bedienstategien

1. Erläutere, wann eine Bedienstrategie: verdrängend, produktiv, fair oder overhead-frei ist!

Für ein Wartesystem, bestehend aus einer Bedienstation mit m Bedieneinheiten und einem Wartepool, existieren folgende Erwartungswerte:

Bedienstation BS / -einheit BE	$c_{BS} = mc_{BE} = \frac{m}{EB}$	
	$E[F_{BS}] = E[D]E[B]$	
Wartepool WP	$E[F_{WP}] = E[D]E[W]$	mit der Wartezeit W
Gesamtsystem WS	$E[F_{WS}] = E[D]E[Y] = E[D](E[W] + E[B])$	

Theorem 9 (Allgmeines Durchsatzgesetz) *Sei v_{ik} die Besuchsanzahl des Auftrags k in der FE i. Dann gilt:*

$$d_i = \sum_k d_{ik} = \frac{1}{t_2 - t_1} \sum_k v_{ik} = \frac{n}{t_2 - t_1} \bar{v}_i = d_{FES} \, \bar{v}_i$$

Theorem 10 (Bottleneck) *Der Verkehrsengpaß VE wir durch die FE mit der größten Auslastung bestimmt:*

$$\rho_{VE} = \max\{\rho_i\}$$

Notwendig, aber nicht hinreichend:

$$\rho_{VE} \leq 1$$

8.2 Sättigungsverhalten

1. Betrachte Abb. 6, die das Sätigungsverhalten von Durchsatz $d(f)$ und Verweilzeit $y(f)$ in Abhängigkeit der Füllung f beschreibt! Erkläre die Schranken und Asymptoten!

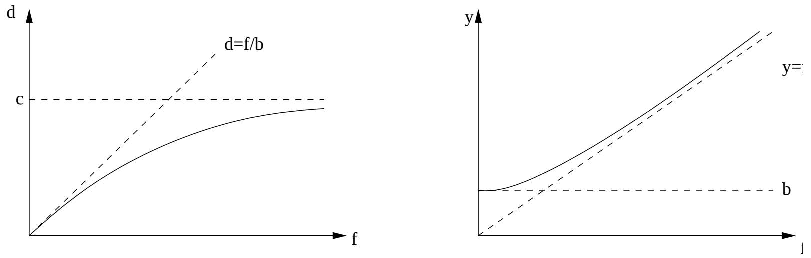

Abbildung 6: Sättigungsverhalten von Durchsatz und Verweilzeit

8.3 Elementare Wartesysteme

Ein Wartenetz besteht aus einer Menge von elementaren Wartesystemen. Es existieren mehrere große Klassen elementarer Wartesystemene. Mit $\rho := \lambda/(m\mu)$ ergibt sich:

$$\begin{aligned} p_i &= \frac{(\rho m_i)^i}{i!} p_0 \qquad f_i \in \{1..m\} \\ p_i &= \frac{\rho^i m^m}{m!} p_0 \qquad f_i \in \{m+1..k\} \\ p_0 &= \left(\sum_{j=0}^{m} \frac{(\rho m)^j}{j!} + \sum_{j=m+1}^{k} \frac{(\rho m)^m}{m!} \right)^{-1} \end{aligned}$$

Daraus ergibt sich der Erwartungswert der Füllung:

$$E[F] = \sum_{f=1}^{k} f p_f$$

Unbegrenzte Kapazität ($M/M/m$) Da ein $M/M/m/k$-System streng genommen, kein Wartesystem ist – es können Aufträge verloren gehen - betrachten wir den Fall mit unbegrenztem Warteraum. Falls $\rho < 1$ gilt, konvergiert p_0 für $k \to \infty$.

Unbeschränkte Anzahl an Bedienprozessen($M/M/\infty$) Für $m = k$ und $m \to \infty$ konvergiert p_0:

$$p_0 = \left(\sum_{j=0}^{\infty} \frac{(\rho m)^j}{j!} \right)^{-1} = e^{-\lambda/\mu}$$

Da die Verweilzeit gleich der Bearbeitungszeit ist, gilt für die Füllung:

$$E[Y] = E[B] = 1/\mu \quad \Longrightarrow \quad E[F] = E[D]\, E[Y] = \lambda/\mu$$

Eine Bedienprozeß ($M/M/1$)

$$\begin{aligned} E[F] &= \frac{\rho}{1-\rho} \\ E[Y] &= \frac{E[F]}{E[D]} = \frac{\rho}{1-\rho} \frac{1}{E[D]} = \frac{E[B]}{1-\rho} \\ E[W] &= E[Y] - E[B] = \frac{\rho}{1-\rho} E[B] \end{aligned}$$

1. Skizzieren sie E[F], E[Y] und E[W] in Abhängigkeit von ρ!

8.4 Wartenetze

Jackson-Netze Eigenschaften (u.a): das Netz ist offen mit exp-vert. Zugängen, alle Knoten haben unendliche Kapazität, jeder Knoten hat m_i Bedieneinheiten, die exp.-vert. Bedienstrategie ist produktiv und restbedienzeitunabhängig, es existiert eine Auftragsklasse.

Theorem 11 *Die Wkt., daß der Knoten i die Füllung f_i beträgt wie im Falle eines einfachen $M/M/m$-Systems:*

$$p_i(f_i) = \frac{(\rho m_i)^{f_i}}{f_i!} p_i(0) \qquad f_i \in \{1..m_i\}$$

$$p_i(f_i) = \frac{\rho^{f_i} m_i^{m_i}}{m_i!} p_i(0) \qquad f_i \in \{m_i..\infty\}$$

$$p_i(0) = \left(\sum_{j=0}^{m_i-1} \frac{(\rho m_i)^j}{j!} + \frac{(\rho m_i)^{m_i}}{m_i!(1-\rho)} \right)^{-1}$$

Dies gilt, obwohl die Rückkehr von Aufträgen die Unabhängigkeit der Ankünfte zerstört.

Gordon-Newell-Netze Eigenschaften (u.a): das Netz ist geschlossen mit n Aufträgen, alle Knoten haben endliche Kapazität $> n$, jeder Knoten hat m_i Bedieneinheiten, die exp.-vert. Bedienstrategie ist produktiv und restbedienzeitunabhängig, es existiert eine Auftragsklasse.

1. Erkläre: Faltungs- und Mittelwertverfahren ($E[Y_i(f)] = E[B_i](1 + E[F_i(f-1)])$) zur Analyse.

BCMP-Netze Allgemeine Klasse: mehrere Auftragsklassen, offene u. geschlossene Subketten, verschiedene Bedienstrategien.

9 Nichtverdrängende Bedienstrategien

Wir setzen im folgenden $M/G/1$ voraus.

- Bedienphase BP:

$$E[BP] = \frac{E[B]}{1-\rho} \qquad E[BP^2] = \frac{E[B^2]}{(1-\rho)^3}$$

- Die Mindest-Wartezeit erntspricht der gewichteten Restbedienzeit GRB eines Auftrags:

$$\begin{aligned} E[W]_{min} &= E[GRB] = P(F=0)0 + P(F>0)E[RB] \\ &= \rho\frac{E[B^2]}{2E[B]} = \frac{\lambda}{2}E[B^2] \end{aligned}$$

- Die maximale Wartezeit entspricht der gewichteten Restbedienphase $GRBP$ eines Auftrags:

$$\begin{aligned} E[W]_{max} &= E[GRBP] = P(F=0)0 + P(F>0)E[RBP] \\ &= \rho\frac{E[BP^2]}{2E[BP]} = \rho\frac{\frac{E[B^2]}{(1-\rho)^3}}{2\frac{E[B]}{1-\rho}} = \frac{\lambda}{2}\frac{E[B^2]}{(1-\rho)^2} \end{aligned}$$

1. Skizzieren die Schranken für die Wartezeit in Abhängigkeit von ρ!

$$\frac{\lambda}{2}E[B^2] = E[W]_{min} \leq E[W] \leq E[W]_{max} = \frac{\lambda}{2}\frac{E[B^2]}{(1-\rho)^2}$$

9.1 FCFS = FIFO

FCFS ist eine faire Strategie. Ein neu ankommender Auftrag wartet die Bearbeitung aller im Wartepool befindlichen Aufträge, sowie die Restbedienzeit des aktuell bearbeiteten Auftrags ab:

$$E[W] = E[F_{WP}]\ E[B] + E[GRB] = \lambda\ E[W]\ E[B] + \frac{\lambda}{2}E[B^2]$$

Dies ist mit $\lambda E[B] = \lambda/\mu = \rho$:

$$E[W] = \frac{\lambda E[B^2]}{2(1-\rho)} = \frac{\rho}{1-\rho}E[B]\frac{1+C_B^2}{2} = \frac{E[GRB]}{1-\rho}$$

9.2 LCFC = LIFO

LCFS ist für $\rho < 1$ eine faire Strategie. Ein neu ankommender Auftrag wartet die Bearbeitung aller während des Wartens eintreffenden Aufträge – d.h. $E[W]\,\lambda$ Aufträge mit einer Bedienzeit $E[B]$ –, sowie die Restbedienzeit des aktuell bearbeiteten Auftrags ab:

$$E[W] = E[W]\lambda E[B] + E[GRB]$$

Also ergeben sich die gleichen Ergebnisse wie für FCFS.

9.3 Kleinrocks Erhaltungsgestz

FCFS und LCFS sind Strategien, die bedienzeitunabhängig sind. Welche Ergebnisse ergeben sich für die bedienzeitabhängigen Wartezeiten $E[W|B=x]$, wenn eine Strategie die Zeiten kennt oder schätzt, so wie z.B. SPT (shortest processing time next)?

Theorem 12 (Kleinrocks Erhaltungsgesetz) *Es gilt:*

$$\int_0^\infty E[W|B=x]xf_B(x)dx = \frac{1}{2}\frac{\rho}{1-\rho}E[B^2]$$

Folgerungen:

- Bei vorgegebener Bedienzeitvtlg. ist $xf_B(x)$ und $E[B^2]$ vorgegeben, – eine konkurrierende nichtverdrängende, lasterhaltende Strategie kann die Bedienzeiten einer Klassen nur senken, indem sie die Bedienzeiten anderer Klassen erhöht.

- Alle bedienzeitunabhängigen Strategien (z.B. FCFS, LCFS) liefern gleiche Wartezeiten.

1. Ist SPT eine faire Strategie?

10 Verdrängende Bedienstrategien

Verdrängende Bedienstrategien erlauben eine bedienzeitabhängige Diskriminierung *ohne* Kenntnis auftragsindividueller Bedienzeiten. Es existiert keine untere Schranke für die Wartezeit:

$$E[W]_{min} = 0$$

Die maximale Wartezeit beträgt:

$$E[W]_{max} = E[W|B = x] = \frac{\lambda E[B^2]}{2(1-\rho)^2} + \frac{\rho}{1-\rho}x$$

1. Warum existieren keine untere Schranke für die Wartezeit?

10.1 Entartete neg-exp. Vtlg.

Sei auch die Bedienzeit exp.-verteilt, dann ändert eine Verdrängung den Zustand des Systems nicht. Es gelten die bekannten Formeln für $M/M/1$, falls eine restbedienzeitunabhängige Strategie vorliegt (z.B. LCFS-p, RR, SET):

$$E[F] = \frac{\rho}{1-\rho} \qquad E[Y] = \frac{E[B]}{1-\rho} \qquad E[W] = \frac{\rho}{1-\rho}E[B]$$

LCFS-p und RR-ps unterscheiden sich aber bezüglich der Streuung der Wartezeiten, RR-ps hat eine niedrige Streuung, aber einen hohen Overhead.

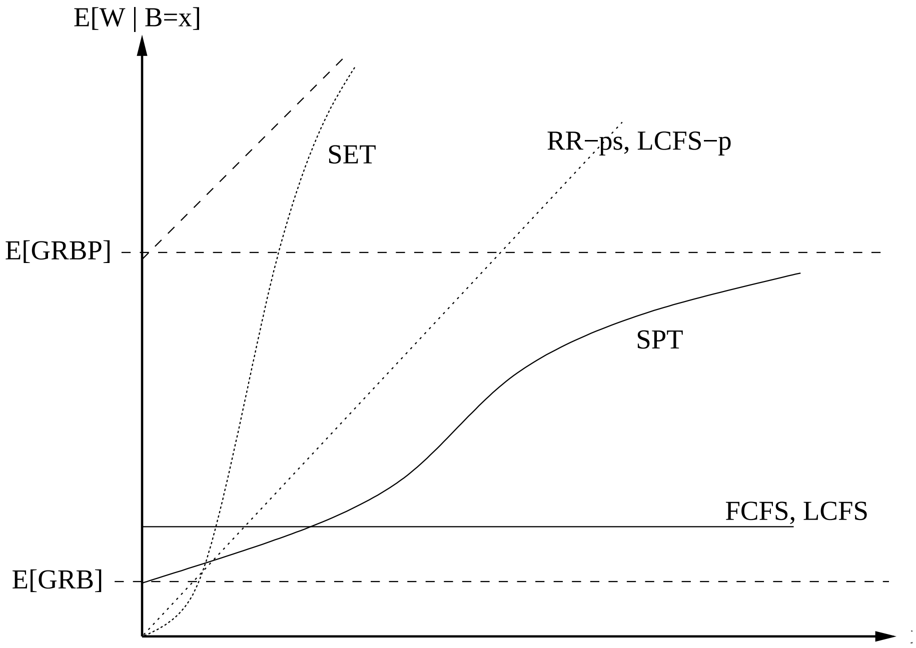

Abbildung 7: Bedienzeitabhängige Strategien

1. Skizzieren sie $E[w|B = x]$ als Funktion von x, so wie in Abb. 7 selbst! (a) Zeichnen Sie zuerst die Schranken für verdrängende und nicht verdrängende Strategien ein! (b) Skizzieren sie dann die nicht verdrängenden Strategien FCFS und LCFS, (c) dann bedienzeitabhängige Strategie SPT! (d) Skizzieren Sie dann die verdrängenden Strategien LCFS-p, RR und SET!