Simulation und Modellierung I WS 1999 / 2000

Common Examination Questions

Lösung anhand des Vorlesungsbegleitbuches:
Banks et al. : "Discrete Event System Simulation"

Christian Meusel, FSI Maschinenbau,
Oktober 2000

Die FSI übernimmt keine Gewähr für die Vollständigkeit und Korrektheit aller Antworten!!!

1 Random number generation

1.1 Wie werden Zufallszahlen erzeugt?

Ziel:

- Folge von Zahlen zwischen 0 und 1
- möglichst gute Erfüllung der idealen Eigenschaften Uniformität und Unabhängigkeit

Lineare Kongruenzmethode:
$\Longrightarrow$ Produktion von ganzen Zahlen zwischen 0 und m-1 über die rekursive Beziehung

$$X_{i+1} = (aX_i + c) \mod m$$

X_0: Startwert oder "seed"
a: konstanter Multiplikator ($\neq 0$)
c: Inkrement
m: Modul ($\neq 0$)
$\Longrightarrow$ Einfluß auf statische Eigenschaften, Zykluslänge (so groß wie möglich!)
$\Longrightarrow$ Umwandlung in Zufallszahlen über $R_i = \frac{x_i}{m}$

Beachte: Das sind diskrete (nicht kontinuierliche!) Werte von R_i aus [0,1] !

1.2 Eigenschaften von Zufallszahlen (vgl. [1] Kapitel 8.1

1. Uniformität ($\Longrightarrow$ gleichverteilt) pdf: $f(x) = \begin{cases} 1 & : \quad 0 \leq x \leq 1 \\ 0 & : \quad \text{sonst} \end{cases}$

   Erwartungswert: $E(R) = \int_0^1 x dx = \frac{1}{2}$
   Varianz: $Var(R) = \int_0^1 x^2 dx - [E(R)]^2 = \frac{1}{12}$

2. Unabhängigkeit

1.3 Was ist so "pseudo" an manchen Zufallszahlen?

sie sind unbrauchbar, also nur pseudos, wenn ...

- nicht gleichmäßig verteilt $\neq$ Uniformität
  (nicht kontinuierlich, sondern Häufungen)
- Mittelwert zu hoch oder zu niedrig
- Varianz zu hoch oder zu niedrig
- Autokorrelation zwischen Zahlen $\neq$ Unabhängigkeit

1.4 Anforderungen an Zufallszahlengeneratoren

- schnell (Kosten!)
- portabel
  (plattformunabhängig $\Longrightarrow$ gleiches Simulationsergebnis für den Simulationslauf egal wo er ausgeführt wird)
- ausreichende Zykluslängen ($\geq$ 10.000 Ereignisse)
- reproduzierbar bei gleichen Anfangsbedingungen
  - Hilfe bei debugging
  - Leichterer Vergleich alternativer Systeme möglich
- "gleichverteilt" und "unabhängig" möglichst gut angenähert

1.5 Wie erreicht man längere Perioden?

a) richtige Wahl der Größen a, c, m und X_0:

$$\begin{array}{llll}
m = 2^b & \text{und } c \neq 0 \text{ und } a = 1 + 4k & ;\text{ggt}(c,m) = 1 & \rightarrow \quad P = m = 2^b \\
m = 2^b & \text{und } c = 0 \text{ und } a = \begin{cases} 3 + 8k \\ 5 + 8k \end{cases} & ; \begin{array}{l} X_0 \text{ ungerade} \\ X_0 \text{ was anderes} \end{array} & \rightarrow \quad \begin{array}{l} P = \frac{m}{4} = 2^{b-2} \\ P < \frac{m}{4} \end{array} \\
m = \text{Primzahl} & \text{und } c = 0 \text{ und } \frac{a^{m-1}-1}{m} \text{ teilbar;} & & \rightarrow \quad P = m - 1
\end{array}$$

z. B. $m = 2^3 1 \Longrightarrow P \approx 10^9$

b) "Combined Linear Congruential Generator":

$$W_i = \left(\sum_{j=1}^{k} W_{i,j} \right) \mod m_1 - 1$$

mit
$W_{i,1}...W_{i,k}$: i unabhängige Zufallsgeneratoren, die j = 1...k diskrete unabhängige Zufallszahlen liefern
$W_{i,1}$ gleichverteilt zwischen 0 und $m_1 - 2$
nochmal:
$W_{i,j} = X_{i,j} - 1$: auf $[0; m_j - 2]$ gleichverteilte ganze Zahlen
i: i-ter Generator mit Modul m_j und Multiplikator a_j
j: j-ter Output

$\rightarrow$ max. Periode:

$$P = \frac{(m_1 - 1)(m_2 - 1)...(m_k - 1)}{2^{k-1}}$$

für

$$X_i = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) \mod m_1 - 1$$

und

$$R_i = \begin{cases} \frac{x_i}{m_1} & ; \quad x_i > 0 \\ \frac{x_1 - 1}{m_1} & ; \quad x_i = 0 \end{cases}$$

Beispiel:

$$\begin{aligned} X_{1,j+1} &= 40014 X_{1,j} \mod 2147483563 \\ X_{2,j+1} &= 40692 X_{2,j} \mod 2147483399 \\ X_{j+1} &= X_{1,j+1} - X_{2,j+1} \mod 2147483562 \\ P &= (m_1 - 1)(m_2 - 1)/2 \approx 2x10^{18} \end{aligned}$$

1.6 Wie können Zufallszahlen getestet werden? ([1] Kapitel 8.4)

Übersicht

1. Frequency-Test auf Uniformität
   - Kolmogorov-Smirnov
   - Chi-Square

   $\rightarrow$ Vgl. des zu testenden Satzes von Nummern mit der Gleichverteilung

2. Runs-Test
3. Autokorrelationstest
4. Gap-Test
5. Poker-Test

Kolmogorov-Smirnov-Test

Geg. :
Serie von N Zufallszahlen
Nullhypothese $H_0 : R_i$ sind gleichverteilt auf [0;1]
Angestrebtes Signifikanzniveau $\alpha = P(\text{Ablehnung von} H_0 | H_0 \text{wahr})$
typischerweise 1 - 5 %

Vorgehen:

1. Ordne R_i aufsteigend
2. Berechne größte und kleinste Abweichung:
   $\max \mid F(x) - S_N(x) \mid$
   $D^+ = \max\limits_{1 \leq i \leq N} \left\{ \frac{i}{N} - R_i \right\}$
   $D^- = \max\limits_{1 \leq i \leq N} \left\{ R_i - \frac{i-1}{N} \right\}$
3. Ermittle $\max(D^+, D^-) = D$
4. Ermittle kritischen Wert D_α aus Tabellenwerk für α und N
5. $D \leq D_\alpha \rightarrow H_0$ wahr, sonst falsch

Chi-Square (χ^2-)Test

Sortieren der N Zahlen und anschließende Einteilung in n Klassen

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

O_i Anzahl der beobachteten ("observed") Zahlen im Intervall i
$E_i = \frac{N}{n}$ Anzahl der erwarteten ("expected") Zahlen im Intervall i
Wähle Intervallgröße so, dass $E_i \geq 5$!
H_0 wird verworfen, wenn $\chi_o^2 > \chi^2_{\alpha,n-s-1}$, wobei s für die Anzahl der für die Beschreibung der Zufallsverteilung geschätzten Parameter steht
hier: Gleichverteilung: s = 0 (keine Freiheitsgrade)

Merke:
Kolmogorov-Smirnov in der Regel genauer, daher v. a. für kleine Sample-Größen;
ab $N \geq 50$ nimm Chi-square-Test!

Anmerkung:
Wird der χ^2-Test zum Überprüfen, ob Verteilungen richtig angenommen wurden, verwendet, ist die Anzahl der Freiheitsgrade s i. allg. nicht null:
POISSON-Verteilung: s = 1 ($\hat{\alpha} = \bar{X} = np$),
GAUSS-Verteilung: s = 2 (σ^2 und μ), siehe S. 370 und Bsp. S. 377

Autokorrelations-Test

Gibt es spezifische Muster?
Gibt es zyklische Variationen?
→ Test auf Abhängigkeiten, regelmäßigen Querbezügen zwischen den erzeugten Zahlen

Gap-Test

Überprüfung der Intervalllängen bis zum Wiedererscheinen der gleichen Zahl unter Anwendung des Kolmogorov-Smirnov Tests

Runs-Test($\Rightarrow \chi^2$)

run
: Folge von ähnlichen Ereignissen, denen ein anders geartetes Ereignis vorausgeht und folgt.

up run
: Folge von Zahlen, die jeweils von einer größeren Zahl gefolgt werden

down run
: Folge von Zahlen, die jeweils von einer kleineren Zahl gefolgt werden

Man unterscheidet:

- Runs up and down:
  In einer wirklich zufälligen Sequenz von N Zahlen errechnet sich der Mittelwert und die Varianz für die Anzahl a der runs wie folgt:

$$\mu_a = \frac{2N-1}{3}$$

$$\sigma_a^2 = \frac{16N-29}{90}$$

  Für $N > 20$ ergibt sich die Test-Statistik zu:

$$Z_0 = \frac{a-\mu_a}{\sigma_a}$$

  Die Nullhypothese kann nicht verworfen werden, wenn gilt:

$$-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$$

- Runs above and below the mean (über bzw. unter dem Erwartungswert)
  Hier gibt es ähnliche Beziehungen $\mu_b(n_1, n_2, N)$ und $\sigma_b^2(n_1, n_2, N)$, die von der Anzahl der Läufe N mit den individuellen Beobachtungen n_1 und n_2 für über bzw. unter dem Durchschnitt liegende Werte abhängen. Mit der tatsächlich beobachteten Anzahl b von Läufen ergibt sich wieder eine Beziehung für Z_0.

- Length of runs
  Hier findet der χ^2-Test eine weitere Anwendung (vgl. S. 307f.)
  $\chi^2_{0_{\alpha,\#Klassen-1}}$

Poker-Test

Basiert auf Häufigkeit mit der gewisse Ziffern in einer Folge von Zahlen sich wiederholen. Chi-square-test mit den beobachteten und erwarteten Häufigkeiten

$$\chi_0^2 = \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} \leq \chi^2_{0_{\alpha,n-1}}$$

i=1: 3 unterschiedliche Ziffern;
i=2: 2 unterschiedliche Ziffern;
i=3: 3 gleiche Ziffern;

1.7 Wie können Zufallszahlen einer beliebigen Verteilung generiert werden?

Ausgangspunkt: Quelle von gleichverteilten Zufallszahlen in (0,1)

$$\text{pdf:} \quad f_R(x) = \begin{cases} 1\,; & 0 \leq x \leq 1 \\ 0\,; & \text{sonst} \end{cases}$$

$$\text{cdf:} \quad F_R(x) = \begin{cases} 0\,; & x < 0 \\ x\,; & 0 \leq x \leq 1 \\ 1\,; & x > 1 \end{cases}$$

Inverse Transformationstechnik

einfach aber nicht unbedingt effizient

1. Berechne cdf (Verteilungsfunktion)
   z.B. für EXPO: $F(c) = 1 - e^{-\lambda X}; X \geq 0$
2. Setze $F(X) = R = 1 - e^{-\lambda X}$ $(0 \leq R \leq 1)$
3. Löse $F(X) = R$ nach X auf $\rightarrow X = F^{-1}(R)$ "Random variate generator"
4. durch Einsetzen von gleichverteilten Zufallszahlen R erhält man exponentiell verteilte Zufallszahl X

2 Random Variables

2.1 Wie kann man Zufallsvariablen beschreiben?

- Wertebereich
- Parameter (z. B. Mittelwert, Varianz ...)
- Art der zugrundeliegenden Dichte- (pdf) oder Verteilungsfkt. (cdf)
- kontinuierlich oder diskret (pm(ass)f statt pdf)

2.2 Was sind die Haupteigenschaften von "pdfs" und "cdfs"? (S. 188 ff.)

probability density function (pdf)

- $f(x) > 0$ für alle x in R_x
- $f(x) = 0$ wenn x $\notin R_x$
- $\int_{R_x} f(x)dx = 1$
- $\int_{x_0}^{x_0} f(x)dx = 0$

cumulative distribution function (cdf)

- $F(x) = \int_{\infty}^{x} f(t)dt$ für x kontinuierlich; $F(x) = \sum_{\text{all } x_i \leq x} p(x_i)$ für x diskret
- F ist nicht abnehmend ($a < b \rightarrow F(b) \geq F(a)$)
- $\lim_{x->\infty} F(x) = 1; \quad \lim_{x->-\infty} F(x) = 0$
- $P(a < x \leq b) = F(b) - F(a)$ für $a < b$

2.3 Einige Verteilungen mit Beispielen für deren Anwendung

<table>
<tr><th>Bezeichnuhlineng</th><th>pdf</th><th>cdf</th></tr>
<tr><td>Gleichverteilung</td><td>$f(x) = \begin{cases} \frac{1}{b-a} & : \quad a \leq x \leq b \\ 0 & : \quad \text{sonst} \end{cases}$</td><td>$F(x) = \begin{cases} 0 & : \quad x < a \\ \frac{x-a}{b-a} & : \quad a \leq x \leq b \\ a & : \quad x \geq b \end{cases}$</td></tr>
<tr><td colspan="3">wird verwendet, wenn die zu beschreibende Variable jeden Wert auf einem Intervall mit gleicher Wahrscheinlichkeit annehmen kann oder man gar nichts über die Zufallsgröße weiß.</td></tr>
<tr><td>Dreiecksverteilung</td><td>$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & : a \leq x \leq b \\ \frac{2(c-x)}{(c-b)(c-a)} & : b < x \leq c \\ 0 & : \text{sonst} \end{cases}$</td><td>$F(x) = \begin{cases} 0 & : x \leq a \\ \frac{(x-a)^2}{(b-a)(c-a)} & : a < x \leq b \\ 1 - \frac{(c-x)^2}{(c-b)(c-a)} & : b < x \leq c \\ 1 & : x > c \end{cases}$</td></tr>
<tr><td colspan="3">Beispiel: Sensor misst Qualität von Speicherchips (o.k. od. Ausschuß);
man erhält Minimum a, Maximum c und Mittelwert $E(X)$ des Ausschusses über einen bestimmten Zeitraum. Daraus errechnet sich die Spitze der Verteilung (Mode) zu $b = 3E(X) - (a + c)$; der Wert für X, bei dem die Gesamtdreiecksfläche genau halbiert würde heißt Median und ergibt sich logischerweise aus der cdf.
Merke: Die Werte für Mode, Median und Mittelwert sind nicht identisch!</td></tr>
<tr><td>Exponentialverteilung</td><td>$f(x) = \begin{cases} \lambda e^{-\lambda x} & : \quad x \geq 0 \\ 0 & : \quad \text{sonst} \end{cases}$</td><td>$F(x) = \begin{cases} 0 & : \quad x < 0 \\ 1 - e^{-\lambda x} & : \quad x \geq 0 \end{cases}$</td></tr>
<tr><td colspan="3">Wird oft verwendet, um "interarrival times" zu modellieren, wenn diese komplett zufällig sind und für äußerst unterschiedliche Service-Zeiten.
λ ist eine Rate: arrivals pro h, Services pro min ...
Weiteres Einsatzfeld: Lebenszeiten von Komponenten, die plötzlich versagen, wie eine Glühbirne, dann ist λ die Ausfallrate.
"memoryless property": $P(X > s + t | X > s) = P(X > t)$
mit anderen Worten: zukünftiges Verhalten unabhängig von der Vorgeschichte</td></tr>
<tr><td>Normalverteilung</td><td>$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$</td><td>$F(x) = P(X \leq x) = \int\limits_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt$</td></tr>
<tr><td colspan="3">wird beschrieben durch Mittelwert ν und Varianz σ^2; cdf kann nicht in geschlossener Form angegeben werden, inverse Transformationstechnik nur mit Polygonzugannäherung anwendbar</td></tr>
<tr><td>Weibullverteilung</td><td colspan="2">Eigenschaften: drei Parameter α, β, ν,
Exponentialverteilung ist Spezialfall mit $\lambda = 1/\alpha$
Beispiele:
Zeit bis zum Versagen eines Flachbildschirms,
Zeit bis Landebahn wieder frei ist</td></tr>
<tr><td>Gammaverteilung</td><td colspan="2">Eigenschaften:
Enthält Gammafunktion mit den zwei Parametern β, θ,
Exponentialverteilung ist Spezialfall mit $\beta = 1$, $\theta = \lambda$</td></tr>
<tr><td>Erlangverteilung</td><td colspan="2">Kombination mehrerer Exponentialverteilungen</td></tr>
</table>

(siehe auch Kapitel 6.3 und 6.4)

3 Data Analysis

3.1 Phasen

1. collect data -> sometimes need to guess!
2. identify probability distribution
   - histogram / frequency distribution (#classes $=\sqrt{N}$)
   - chose family of distributions
3. chose parameters to adjust distribution
4. evaluate by goodness-of-fit tests

3.2 What is Q-Q plot?

usefull tool for evaluating distribution fit

- sortiere die sample data von klein nach groß $(Y_i, i = 1...n)$
- $Y_i \approx F^{-1}\left(\frac{i-0{,}5}{n}\right)$ Plot ergibt Gerade mit Steigung 1, falls richtige Parameter gewählt wurden

4 Output Analysis

4.1 How are results obtained using a simulation program?

Eine stochastische Diskret-Ereignis-Simulation ist ein statistisches Experiment
$\rightarrow$ statistische Analyse nötig
Zweck der Simulation ist es, eine Abschätzung für die interessierende Systemantwort zu finden.
Unterscheidung in

- terminierende Simulation
  $\rightarrow$ von t=0 bis t=T_E (vorher bekannte Endzeit, oder bei Eintritt eines bestimmten Ereignisses)
- steady-state Simulation
  schwieriger wegen Anfangsbedingung und Wahl der Lauflänge
  Gefahr:
  - Löschen von Daten
  - zu lange Läufe

bei steady-state Sim.:

$$\theta = \lim_{n->\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i$$

batch means (S. 454 ff):
Unterteilt man den beobachteten Zeitbereich in Intervalle, ergeben sich für jeden Zeitbereich j gesonderte "batch means" $Y_{rj} = \frac{1}{\Delta t} \int\limits_{(j-1)\Delta t}^{j\Delta t} L_Q(t,r)dt$.
Bildet man für verschiedene Läufe die Mittelwerte über die Zeitbereiche mit jeweils gleichem j-Zähler, hat man die sog. "ensemble averages" $Y_{\cdot j} = \frac{1}{R} \sum\limits_{r=1}^{n} Y_{rj}$
Man sieht, dass für j =1 und 2 deutlich geringere Werte als für den Rest vorliegen ("initialization phase of length T_0")
Diese ($d = 2$) Werte sollten bei der Mittelwertbildung herausgenommen werden:

$$\overline{Y}_{\cdot\cdot}(n,d) = \frac{1}{n-d} \sum_{j=d+1}^{n} Y_{\cdot j}$$

Der zweite Teil der steady-state Simulation mit $t \geq T_0$ ist die "data collection phase" und sollte mindestens 10mal so lange dauern wie T_0

Methode der unabhängigen Replikationen (Richtwert: n= 25)

$\rightarrow$ statistisch unabhängige Observationen werden erhalten
Beispiel: Geschäft
Gesucht:
a) Auslastung des Verkäufers
b) Verweildauer des Kunden
für die Auslastung ergeben sich aus Sicht der Statistik 2 Hauptfragen, wenn $\hat{\rho}_r$ mit $r = 1 \ldots n$ die ermittelten Auslastungen für n unabhängige Replikationen (=Läufe) sind:

1) Was ist eine Abschätzung für die wahre Auslastung?
   $\rho = E(\hat{\rho}_r)$ "point estimate"

2) Wie groß ist der Fehler in unserem "point estimate"?
   "standard error", "confidence intervall"

Bias: Abweichung des "point estimators" vom wahren Wert der gesuchten Größe:
$E(\hat{\theta}) = \theta + b$
Ziel der Simulation: Finde $\theta = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)$

4.2 Name two ways of obtaining data sets!

Simulation output data	Estimator	Calculation
$\{Y_1, Y_2 \ldots Y_n\}$ = in der Zeit diskrete Daten z. B. Service-Zeit der Kunden 1, 2, ... n	θ: ordinary mean	$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i$
$\{Y(t); 0 \leq t \leq t_E\}$ = in der Zeit kontinuierliche Daten z. B. Schlangenlänge zur Zeit t	Φ: time weighted mean	$\hat{\Phi} = \frac{1}{T_E}\int_{t_0}^{T_E} Y(t)dt$

4.3 What is a confidence intervall? (Beliebte Frage!!!!)

genaueres nachzulesen in [1] Kapitel 12 !

$$\hat{\theta} \pm t_{\alpha/2,f} \cdot \hat{\sigma}(\hat{\theta})$$

4.4 How can two system designs be compared?

Beispiele:

- queuing system: Anzahl Queues, Server
- inventory system: versch. Bestellintervalle, Lagerbuffer

Man schaut sich das Konfidenzintervall an: Gibt es einen signifikanten Unterschied zwischen zwei Systemen? (vgl. Achsenskizze mit 0, S. 478!)

Beachte: Unterschied: statistische und praktische Signifikanz (S. 479)

sampling methods:

- independent with equal variances
- independent with unequal variances
- correlated sampling, common random numbers

5 General

5.1 What is the method of independent replications?

Y_{ri}: i-te Observation (z. B. Servicezeit von Kunde i) in der r-ten Replikation (Simulationslauf)
Beachte: Für ein fixes r ist $Y_{r1}, Y_{r2}, Y_{r3}, Y_{r4}...$ eine autokorrellierende Sequenz!
Aber: Y_{ri} und Y_{si} sind statistisch unabhängig
Sample mean für die Replikation r: $\hat{\theta}_r = \frac{1}{n_r} \sum\limits_{i=1}^{n_r} Y_{ri}$
daraus folgt: "overall point estimate" $\hat{\theta} = \frac{1}{n} \sum\limits_{i=1}^{R} \hat{\theta}_r$
Schätzung der Varianz: $\hat{\sigma}^2(\hat{\theta}) = \frac{1}{R(R-1)} \sum\limits_{r=1}^{R} (\hat{\theta}_r - \hat{\theta})^2$
$\rightarrow$ Konfidenzintervall

5.2 Was ist Simulation und wann setzt man sie ein?

Def. laut Skript: "... imitation of the operation of a real-world process or system over time"

- Betrachtung eines komplexen Systems oder eines Subsystems innerhalb eines komplexen Systems
- Simulilerung von Wechseln informeller, organisatorischer Art bzw. in der Umgebung -> Einflüsse auf das Verhalten des Modells werden klar.
- Wissensgewinn beim Aufstellen des Modells -> Verbesserung des Systems durch Analyse der Wirkungsmechanismen/ Stukturen und Optimierung anhand der gewonnenen Erkenntnisse
- Variation von Eingabeparametern und Auswertung der zugehörigen Simulationsergebnisse -> Einsicht, welche Variablen wichtig sind und wie die Variablen zusammenwirken
- Pädagogisches Hilfsmittel um analytische Methoden zu verstärken
- Herumexperimentieren mit alternativen Designs -> "Was würde passieren, wenn.... ?"
- Verifizieren analytischer Ergebnisse

5.3 Vor- und Nachteile von Simulation

Vorteile:

- Neues ausprobieren ohne in laufende Vorgänge eingreifen bzw. diese unterbrechen zu müssen
- Hypothesen über das "wie" und "warum" von gewissen Phänomenen können auf ihre Wahrscheinlichkeit getestet werden.
- Zeit kann komprimiert oder gedehnt werden, je nachdem, was untersucht werden soll
- Einsicht in das Zusammenwirken von Variablen
- Verständnis der Bedeutung von Variablen für das Verhalten des Systems
- "Bottleneck"- Analysen -> Indiz, wo viel Arbeit im Prozess geleistet wird, wo sich Information und Material usw. besonders aufstauen
- Simulationen helfen zu verstehen, wie ein System wirklich funktioniert, was manchmal anders ist als ein einzelner Mensch sich das vorstellt

Nachteile:

- Modellbildung erfordert Erfahrung
- Simulationsergebnisse sind oft schwer zu interpretieren
- Modellbildung und Analyse incl. Auswertung der Ergebnisse erfordert Zeit (und damit auch Geld!)
- Simulation wird manchmal (unsinnigerweise) auch dann eingesetzt, wenn eine analytische Lösung möglich wäre, was allerdings bei komplexen realen Systemen äußerst selten der Fall ist

5.4 Describe each stage of a simulation project!

Problemformulierung

Festlegung der Ziele (Anforderungsliste) / Gesamtplan

- Was für Fragen interessieren?
- Entscheidung, ob Sim. die angemessene Methode ist für das formulierte Problem
- alternative Systeme
- Methode zur Beurteilung der Alternativen (Gewichtung von Ergebnisgrößen, Index...)
- Personalbedarf, Kostenvoranschlag

Konzeptuelles Modell

- mit einfachen Modell beginnen
- Idealisierung - Abstrahierung

Datensammlung

- hängt vom konzeptuellen Modell ab
- Daten für spätere Validierung nicht vergessen

Programmierung

Verifizierung

Tut das Programm das, was es soll? ("debugging")

Validierung

Ist das, was aus dem Programm rauskommt, das gleiche, wie das wirkliche System produzieren würde?
Iterativer Vorgang: ggf. Anpassen der Parameter und abermaliger Vergleich mit Realität

- face validity
  - Expertenrat einholen:
    "Kann das so überhaupt stimmen, was da rauskommt?")
  - sensitivity analysis:
    Was passiert, wenn Eingabegrößen geringfügig verändert werden? Explodiert die Lösung? Kommt was ganz anderes raus?
- Annahmen nochmals überprüfen (ist statistisches Modell wirklich richtig -> goodness-of-fit Tests mit den Daten)
- Vergleich der von der Simulation produzierten Daten mit den Referenzdaten die für die Valdierung aus dem realen System genommen wurden. Stimmen sie überein?

Experimental design

- Festlegen der möglichen Alternativen
- Auswahl derer von Interesse

Durchführung der Simulationsläufe und Analyse

Entscheidung ob mehr Läufe nötig

Sind Ergebnisse akurat oder Redesign nötig?

Dokumentierung und Bericht

- Beschreibung des Modells und der durchgeführten Experimente
- Aufführen aller Annahmen und Vereinfachungen!

Umsetzung der Ergebnisse in reale Lösungen

5.5 Validierung und Verifizierung

Verifizierung = Building the model right (Vergleich Modell - Simulationsprogramm)
Validierung = Building the right model (Vergleich Modell - Realität)

Methoden der Verifizierung:

- Debugger (schrittweiser Abgleich)
- Trace (für jeden Schritt Ausgabe von interessierenden Variablenwerten
- Dokumentatin (Gedankengang nachvollziehbar)

5.6 Beschreibe folgende Begriffe!

entity:	Jedes Objekt oder jede Komponente im System, die eine explizite Repräsentierung im Modell erfordert.
event / Ereignis:	Ein instantanes (plötzliches) Ereignis, das den Zustand des Systems verändert
system:	Eine Sammlung von Entities, die über eine gewisse Zeit miteinander interagieren, um ein oder mehrere Ziele zu erreichen
activity:	Zeitdauer mit genau spezifizierter Länge, die bekannt ist (z. B. Service- oder Ankunftszeiten, wenn sie (die activity) beginnt (auch wenn diese Zeitdauer einmal als statistische Verteilung definiert wurde!)
delay:	Im Gegensatz zur activity: Zeitdauer mit unspezifizierter Länge, die man nicht von vorne herein kennt (z. B. die Zeit die ein Kunde in der Warteschlange verbringen wird)
state / Zustand:	Satz von Variablen, die genug Infos enthalten, ein System zu jeder beliebigen Zeit komplett zu beschreiben
boundary / Systemgrenze:	Was wird in der Simulation noch berücksichigt, was nicht mehr?
event notice:	Ein Aufzeichnung eines Ereignisses, das sich in Zukunft ereignen wird, oder gerade stattfindet. Hierbei sind sämtliche Daten assoziiert, die für die Ausführung des events notwendig sind (Minimum: Typ des Ereignisses und Zeit, wann es stattfindet)
event list:	Eine Liste von event notices für zukünftige Ereignisse, sortiert nach der Zeit ihres Eintreffens (-> F(uture)E(vent)L(ist))

Literatur

[1] Banks, Carson, Nelson: Discrete-Event System Simulation