

Übung 1 -Empirische Bewertung in der Informatik

Aufgabe 1-1: Definitionen von Forschung (engl. Research)

Definition 1:

„Forschung und Entwicklung ist die systematische Suche nach neuen Erkenntnissen unter Anwendung wissenschaftlicher Methoden und in geplanter Form. Forschung ist der generelle Erwerb neuer Kenntnisse, Entwicklung deren erstmalige konkretisierende Anwendung und praktische Umsetzung. Die neuen Kenntnisse können sich sowohl auf Produkte als auch auf (Herstellungs-)Verfahren und Produkt- sowie Verfahrensanwendungen erstrecken. Entbehrt Forschung noch eines realen Verwertungsaspekts, so handelt es sich um Grundlagenforschung. Die angewandte Forschung ist dagegen bereits auf konkrete Anwendungsmöglichkeiten hin ausgerichtet.“

Quelle:

Springer Gabler Verlag (Herausgeber), Gabler Wirtschaftslexikon, Stichwort: Forschung und Entwicklung (F&E), online im Internet: <http://wirtschaftslexikon.gabler.de/Archiv/9701/forschung-und-entwicklung-f-e-v10.html>

Autoren: Prof. Dr. Kai-Ingo Voigt, Prof. Dr. Günter Müller-Stewens, Prof. Dr. Martin G. Möhrle, Prof. Dr. Dieter Specht

Definition 2:

„Alle systematischen Bestrebungen und Bemühungen, um neue Erkenntnisse für Wissenschaft oder Industrie zu erlangen; man unterscheidet zwischen Grundlagenforschung und angewandter Forschung; die Grundlagenforschung soll lediglich neue Erkenntnisse bringen, wird meist an Universitäten durchgeführt und bildet die Basis für die angewandte Forschung; die angewandte Forschung soll bestimmte praktische Anliegen oder Probleme lösen, wird auch in der freien Wirtschaft betrieben und liefert der Grundlagenforschung neue Ideen; bevor man mit der Forschung beginnt, setzt man sich Ziele, welche erreicht werden sollen und bestimmt jenes Wissen, welches derzeit zur Verfügung steht; danach wird der Versuch geplant und anschließend durchgeführt; zum Schluss wird der gesamte Forschungsprozess dokumentiert und interpretiert; die Finanzierung erfolgt entweder über staatliche Ressourcen (staatlich finanzierte Forschung, zum Beispiel die Hochschulforschung) oder nicht staatliche Zuschüsse (nicht staatlich finanzierte Forschung, zum Beispiel durch Stiftungen); Gelder für die Forschung kommen zum größten Teil aus der Wirtschaft.“

Quelle: wissen.woxikon.de – Begriff: Forschung

Link: <http://wissen.woxikon.de/forschung>

Definition 3:

„Systematic investigative process employed to increase or revise current knowledge by discovering new facts. It is divided into two general categories: (1) Basic research is inquiry aimed at increasing scientific knowledge, and (2) Applied research is effort aimed at using basic research for solving problems or developing new processes, products, or techniques.“

Quelle: businessdictionary.com - Begriff: research

Link: <http://www.businessdictionary.com/definition/research.html#ixzz308BlNIIR>

Zusammenfassung von Konzept der Forschung:

Bei Forschung handelt es sich also um die Absicht systematisch neue Erkenntnisse zu erlangen. Dort ist zwischen Grundlagenforschung und angewandter Forschung zu unterscheiden. Die Grundlagenforschung hat zum Ziel neue wissenschaftliche Erkenntnisse hervorzubringen und angewandte Forschung nutzt die Grundlagenforschung, um Probleme zu lösen oder neue Produkte und Prozesse zu entwickeln.

Aufgabe 1-2: R - Introduction

Erklärungen zu Indexing:

- in „The R Language Definition“ (<http://www.r-project.org>):
 - <http://127.0.0.1:20648/doc/manual/R-lang.html#Indexing>
- in "An Introduction to R" (<http://www.r-project.org>):
 - <http://127.0.0.1:20648/doc/manual/R-intro.html#index-Indexing-vectors>
 - <http://127.0.0.1:20648/doc/manual/R-lang.html#Indexing-matrices-and-arrays>
- Nach der Eingabe von "?" in R werden die Indizes auch erklärt:
 - <http://127.0.0.1:20648/library/base/html/Extract.html>
- in der kurzen „R Reference Card“ von Tom Short in Abs. „Slicing and extracting data“:
 - <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>

Notizen:

- $a**b$ - Operator $\Leftrightarrow a^b$
- $a\%b$ = Modulo-Operator
- `mode(x)` - gibt Typ von x zurück
- **c(a,b,...,z)** – **erstellt Vektor**, `c()` gibt NULL zurück
- `rep(x,b)` – Wiederholt Objekt x b mal.
- **Sequenzen**: $x:y \Leftrightarrow \text{seq}(\text{from}=x, \text{to}=y)$, `along=x` - Länge von x, `length` – gewünschte Länge, `by=x` – Inkrementieren von x (x Schritte)
- `length(x)` – gibt gröÙe von Vektor zurück
- **Vektorindizierung** (x und y sind Vektoren):
 - `x[y]` = gibt Index `x[y1], x[y2], ..., x[yn]` (`yn = y[n]`)
 - `y >= 3` – gibt aus, ob `y[1..n] >= 3` ist. (TRUE || FALSE)
 - `y[y >= 3]` – gibt die Elemente von y zurück, welche ≥ 3 sind, also bei `y >= 3` TRUE sind
 - `x[-2]` – Vektor x ohne das 2. Element, `x[-y]` – Vektor x ohne Element `y[1..n]`
 - Jedes 2. Element: `x[2*(1:(length(x)/2))]` \Leftrightarrow `x[c(2,4,6,...,n)]`
 - Jedes 2. Element: `x[c(F,T)]` – Wenn `length(c(F,T)) < length(x)` dann wird `c(F,T)` immer wiederholt. An der stelle, wo TRUE vorkommt wird `x[i]` zurückgegeben
- **Matrizen** :
 - `cbind(x,y)`, `rbind(x,y)`: verbindet Vektoren column-Weise/row-Weise
 - `dim(m)`, gibt dimension wieder, `nrow(m)`, `ncol(m)` – Anzahl Zeilen/Spalten
 - `m[i]` gibt i-te Position wieder. Bei 3x3-Matrix: Erst Zeilen nach unten zählen, dann nächste Spalte. Die 6. Position ist also das Feld `m[3,2]`
- **Listen**
 - `l=list(one=x,two=y)` – Erzeugt Liste mit Elementen one und two Variablen.
 - Zugriff: `l[2]` – Name von Element + Element, `l[[2]]` – nur 2. Element \Leftrightarrow `l$two`
- Sonstiges: `Inf` = NaN, `NA` = not assigned, `NULL` = Leer
- Funktionen auf Vektoren und Matrizen an Stelle von Schleifen
 - `lapply(l, FUN)` - FUN auf Listenelemente
 - `apply(m,(1|2),FUN)` – FUN auf matrix, 1 für zeilen, 2 für Spalten
- **Funktionen**:
 - `jitter(m)` fügt noise zu Werten hinzu
 - Bsp:
`legal.level = function(dat, levels=5) dat >= 1 && dat <= levels && dat == trunc(dat)`
Aufruf: `legal.level(4)`
 - `apply(xy, 1, function(x, lev) all(legal.level(x,lev)), lev=7)` –
`lev=7` fügt initialisierte Variable zu FUN hinzu, `all()` prüft, ob alle Werte TRUE sind

Aufgabe 1-3: In R (ohne Loops)- 100000 mal würfeln. Wie oft kommt die Sequenz 3,4,5 vor?

Idee: Generiere aus den 100000 Würfelwürfen alle aufeinanderfolgenden 3-elementigen Teilmengen und vergleiche jede mit der Sequenz {3,4,5}. Zähle die Gleichnisse.

R-Code :

```
repNum=100000 # number of Repitions
diceRun = sample(1:6,repNum,replace=T) #create 100000 dice throws
cpDice = diceRun # copy of dice throws
wishedSeq = c(3,4,5) # wished sequence

matr1=matrix(cpDice,nrow=3)
# delete last column because it is invalid (recycled last column)
matr1=matr1[,-dim(matr1)[2]]
# count all equal Sequences, in sum T equals 1 and F equals 0
counter= sum(apply(matr1,2, function(x) all(x==wishedSeq)))
cpDice <- cpDice[-1] # delete first element
matr2=matrix(cpDice,nrow=3)
counter= counter + sum(apply(matr2,2, function(x)
all(x==wishedSeq)))
cpDice <- cpDice[-1] # delete first element
matr3=matrix(cpDice,nrow=3)
matr3=matr1[,-dim(matr3)[2]] # delete last column because it is
invalid (recycled last column)
counter= counter + sum(apply(matr3,2, function(x)
all(x==wishedSeq)))
counter
```