

Eberhard Karls Universität Tübingen
Mathematisch-Naturwissenschaftliche Fakultät
Fachbereich Informatik

Titel

Masterarbeit Informatik

Vor- und Nachname

Datum

Betreuer

Prof. Dr. Oliver Bringmann
Lehrstuhl für Eingebettete Systeme
Fachbereich Informatik
Universität Tübingen

Name Zweitgutachter
Arbeitsbereich
Fachbereich Informatik
Universität Tübingen

Nachname, Vorname:

Titel der Arbeit

Masterarbeit Informatik

Eberhard Karls Universität Tübingen

Bearbeitungszeitraum: von-bis

Zusammenfassung

Hier kommt die Zusammenfassung hin!!!

Danksagung

Hier kommen die Danksagungen hin (falls gewünscht)!!!

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich diese schriftliche Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe.

Ort, Datum

Unterschrift

Inhaltsverzeichnis

1	Einleitung	1
1.1	Arbeit	2
2	Grundlagen	5
2.1	Abschnittsüberschrift	5
2.1.1	Unterabschnittsüberschrift	5
2.2	Stichproben	5
2.2.1	Häufigkeiten und Histogramm	6
2.2.2	Wichtige Verteilungen	6
2.3	Schätzung von Parametern	6
2.3.1	Eigenschaften von Punktschätzungen	7
3	Stand der Technik	8
4	Konzept	9
5	Ergebnisse	10
6	Zusammenfassung und Ausblick	11
	Literaturverzeichnis	12
	Abkürzungsverzeichnis	13
	Abbildungsverzeichnis	14

Tabellenverzeichnis

15

Kapitel 1

Einleitung

Beginnen sollte die Arbeit mit einer kurzen Einführung und Motivation in das Themengebiet. Aus der Motivation lassen sich ggf. bestehende Probleme ableiten bzw. Probleme aufzeigen, die man in dieser Arbeit lösen möchte. Danach erfolgt die Beschreibung der Aufgabenstellung.

Die Arbeit gliedert sich dazu wie folgt: Die Grundlagen von BlaBlaBla werden in Kapitel 2 erarbeitet.

Einige LaTeX-Hinweise

Im folgenden wird das Einbinden einer Abbildung als ‘pdf-Datei’ in ein \LaTeX -Dokument gezeigt.

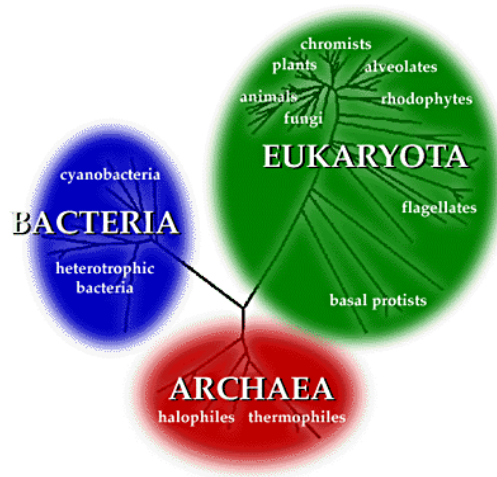


Abbildung 1.1: Three Domains

Abbildung 1.1 zeigt ...

Tabellen können wie folgt erstellt werden:

Spalte 1	Spalte 2	Spalte 3	Spalte 4
xxx1111	xxxxxxx2222222	xxxxxxx333333	xxxxxxxxxxx444444
...

Tabelle 1.1: Beispieltabelle mit einer langen Legende, damit man sieht, dass in der Legende der Zeilenabstand verringert wurde. Außerdem soll auch der Font etwas kleiner gewählt werden. So sieht die ganze Umgebung kompakter aus.

Eine Aufzählung geht wie folgt:

- ...
- ...

Eine nummerierte Aufzählung:

1. ...
2. ...

Betonungen sollen *kursiv* gedruckt werden. **Fettdruck** ist auch möglich.

Referenzen: [SST97, TCS96, STS98]

1.1 Umfang der Arbeit

„Dies ist eine der meistgestellten Fragen. Natürlich verbirgt sich dahinter die Vermutung, die erzielbare Note sei – gutachterabhängig – mit der Seitenzahl korreliert (vgl. Abb.1.2). Nur wie? Linear, normalverteilt, nach dem Gesetz vom abnehmenden Grenznutzen? Tatsächlich kommt es auf die Qualität Ihrer Resultate an. Wenn Sie mit Ihrer Arbeit das Collatz-Problem, auch bekannt als Ulams Vermutung, widerlegen können, genügt eine Seite Inhalt mit dem Hinweis, die Zahl, welche die Vermutung widerlegt, befinde sich auf der beige-fügten CD.

Für alle, die nicht so viel Glück haben, soll die folgende Tabelle 1.2 auf Seite 3 als Richtschnur dienen. Dabei wurden Anhänge, Inhalts- und Abbildungsverzeichnisse sowie Stichwortverzeichnisse (sofern überhaupt vorhanden, da nicht üblich) nicht gerechnet. Bedenken Sie, dass Ihr Gutachter das alles gründlich lesen soll, der Zweitgutachter es vielleicht auszugsweise lesen muss. Formulieren Sie deshalb knapp und auf den Punkt, vermeiden Sie Wiederholungen („Wir kommen nochmal auf das schwierige Problem der Softwareauswahl aus Kapitel zwei zu sprechen, wo wir feststellten, dass ...“). Längliche Passagen, etwa Programmstücke, Teile der Dokumentation, sehr lange Zitate (etwa ein Beweis, ein Gerichtsurteil, ein Zeitschriftenartikel im Wortlaut), Messreihen usw. verbanen Sie in den Anhang (mit der Gewissheit, dass das kaum jemand gründlich

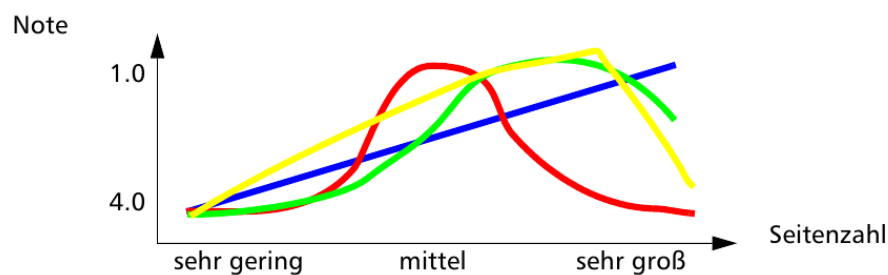


Abbildung 1.2: Welcher Verteilung folgt die Note als Funktion der Seitenzahl? (aus [Weg09a])

Art der Arbeit	Untergrenze	Obergrenze	Anmerkung
Bachelor	35	65	ideal ≤ 50
Master	50	85	ideal ≤ 70

Tabelle 1.2: Empfehlung zur Seitenanzahl der Arbeit

lesen wird). Aber auch bei den Anhängen ist weniger oft mehr. Noch umfangreichere Teile lassen sich auf eine CD brennen, die der Arbeit beigelegt wird; allerdings ist umstritten, ob ein Gutachter sich diese anschauen muss.

Weil das Vorwort, der erste Abschnitt der Einleitung und die abschließende Zusammenfassung mit Ausblick immer gründlich gelesen werden, sollten Sie darauf besonderes Augenmerk legen. In der Regel schreibt man die Einleitung und das Vorwort auch erst, wenn der restliche Teil einschließlich Zusammenfassung (Fazit) steht, Spötter nennen das die Anpassung des Anforderungsprofils an das tatsächlich erzielte Resultat. Zuletzt ein Rat, wenn der Umfang der Arbeit erkennbar zu groß wird. So wie bei Seminarvorträgen Schnellersprechen das Problem eines zu umfangreichen Folienprogramms nicht lösen kann, so wenig lässt sich mit typografischen Mitteln (kleinerem Font, engeren Zeilenabständen, breiteren Spalten) wesentlich Platz ohne Verlust an Lesbarkeit gewinnen. Sie kommen nicht umhin, größere Teile der Arbeit zu streichen oder wesentlich zu straffen. Dafür bieten sich oft die Kapitel an, in denen Sie den mühsamen Prozess der Lösungsfindung einschließlich aller notwendigen Vorarbeiten und Diskussionen mit dem Anwender dokumentiert haben. Hinter solchen längeren Beschreibungen steckt der verständliche Wunsch, der Gutachter möge honorieren, dass Sie unglaublich mit dem Auftraggeber, der undurchsichtigen Software, dem abstürzenden Computer u.a.m. kämpfen mussten und vieles zunächst nicht so funktionierte, wie gedacht. Leser sind aber wie Restaurantgäste, Gutachter ähneln Gourmetkritikern. Sie sind mitleidslos und schauen nur auf den Teller vor sich. Sie wollen nichts davon wissen, dass frische Seezunge heute enorm schwierig zu beschaffen war und der Jungkoch sich am Gratin die Finger verbrannt hat. Halten Sie Ihre Schwierigkeiten in einem

ehrlich geschriebenen 10-Zeilen-Abschnitt der Zusammenfassung fest, als Teil der Selbstreflektion, die immer zu einer Abschlussarbeit gehört, und streichen Sie schweren Herzens Teile der Entwicklungssaga.” (aus [Weg09b])

Kapitel 2

Grundlagen

Ziel dieses Kapitels ist eine Einführung in die Thematik BlaBlaBla ...

2.1 Abschnittsüberschrift

BlaBlaBla ...

2.1.1 Unterabschnittsüberschrift

BlaBlaBla ...

Bevor wir uns der Auswertung bzw. Bewertung der gewonnenen Primärdaten zuwenden, wollen wir zunächst einige grundlegende Begriffe der deskriptiven Statistik wiederholen.

2.2 Stichproben

Grundsätzlich haben wir es bei Microarrayexpressionsdaten mit einer *Stichprobe* aus einer *Population (Grundgesamtheit)* zu tun. Wir bezeichnen nun im allgemeinen mit $X = \{x_1, x_2, \dots, x_n\}$ die Beobachtungsdaten vom Umfang n . Diese Daten sollen mit statistischen Kenngrößen beschrieben werden. Aus diesen will man möglichst zuverlässig auf die zugrundeliegende Verteilung in der Grundgesamtheit schließen. Hierzu verwenden wir die **Lage-** und **Streuparameter**. Zunächst wenden wir uns aber der Häufigkeits- und Summenhäufigkeitsverteilung zu, die sowohl graphisch als auch numerisch einen Eindruck über die Verteilung von X bieten. Dafür betrachten wir diskrete Verteilungen.

Gegeben sei eine Stichprobe (X_1, X_2, \dots, X_n) . Eine Funktion $Z_n = Z(X_1, \dots, X_n)$ wird als *Stichprobenfunktion* bezeichnet. Sie ist selber eine Zufallsgröße.

2.2.1 Häufigkeiten und Histogramm

In der Stichprobe X trete der Wert x_i genau n_i mal auf, $i = 1, 2, \dots, m$. Dann ist $\sum_i n_i = n$. Der Quotient n_i/n ist die *relative Häufigkeit* für das Eintreten des Ereignisses " $X = x_i$ ". Die Menge der relativen Häufigkeiten $\{n_1/n, n_2/n, \dots, n_m/n\}$ heißt *Häufigkeitsverteilung* von X . Ferner heißt die Menge $\{s_1, \dots, s_m\}$ mit $s_i = \sum_{k=1}^i n_k/n$ die *Summenhäufigkeitsverteilung* von X .

Für die graphische Darstellung der Häufigkeitsverteilung wird das *Histogramm* gewählt. für die Summenhäufigkeitsverteilung die *Treppenfunktion*.

Wenn wir natürlich Zahlen mit Komma haben, so sieht das in Deutsch irgendwie seltsam aus, z.B. 0,7. Ich nehme dafür den SIstyle: 0.7 usf. Oder nett ist auch: 9.3×10^5 km

Mit diesem Kommando kann man eigene Kommentare in der Textspalte hinzufügen.

Mit diesem Kommando kann man eigene Kommentare am Spaltenrand hinzufügen.

2.2.2 Wichtige Verteilungen

Die Normalverteilung

Die Dichte der Normalverteilung ist gegeben durch

$$g(x) = \frac{1}{2\pi\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.1)$$

wobei μ (Lage) der Mittelwert und σ (Breite) die Standardabweichung der Normalverteilung ist. Durch die z -Transformation lässt sich die Normalverteilung auf die Standardnormalverteilung mit $\mu = 0$ und $\sigma = 1$ transformieren.

Die Normalverteilung bildet die Basis fast der gesamten statistischen Theorie.¹ Auch bei der Analyse der Microarraydaten werden wir sehr oft von der Annahme der Normalverteilung Gebrauch machen. Allerdings sollten wir uns klarmachen, dass rein experimentell zahlreiche Untersuchungen gezeigt haben, dass die echten Fehler selten, wenn überhaupt normal verteilt sind.

2.3 Schätzung von Parametern

Allgemein erhofft man sich beim Ziehen einer Stichprobe, einen unbekannten Parameter γ der Grundgesamtheit, z.B. den Mittelwert, aus der Stichprobe zu schätzen.

¹„Everyone believes in the normal law, the experimenters because they imagine it is a mathematical theorem, and the mathematicians because they think it is an experimental fact.“ (Gabriel Lippman, in Poincar’s Calcul de probabilité, 1896)

2.3.1 Eigenschaften von Punktschätzungen

Kapitel 3

Stand der Technik

Während im Grundlagenkapitel notwendige Begrifflichkeiten, Datenstrukturen, Basisalgorithmen oder Hardware-Architekturen vorgestellt werden, befasst sich dieser Abschnitt mit einer kurzen Diskussion existierender Ansätze und deren Probleme.

Je nach Themenstellung kann dieser Abschnitt auch entfallen. Eine kurze Diskussion kann in diesem Fall entweder in der Aufgabenstellung (Kapitel 2) oder zu Beginn des eigenen Konzepts (Kapitel 4) erfolgen.

Kapitel 4

Konzept

Mit das Wichtigste natürlich!

Hier wird der eigene Ansatz vorgestellt. Der Titel sollte natürlich nicht einfach Konzept heißen, sondern konkret den eigenen Ansatz benennen.

Kapitel 5

Ergebnisse

Hier sollen die erreichten Ergebnisse vorgestellt werden. Hierzu zählt die Vorstellung des Versuchsaufbaus sowie die geeignete Aufbereitung und Diskussion der Ergebnisse. Mehrwert oder Nutzen benennen!

Kapitel 6

Zusammenfassung und Ausblick

Zusammenfassung der Arbeit unter Verwendung der eingeführten Begrifflichkeiten. Hier darf davon ausgegangen werden, dass der Leser die Arbeit gelesen hat bzw. kennt. Ferner sollte ein Ausblick auf Erweiterungsmöglichkeiten oder sich ergebende Forschungsfragen gegeben werden.

Literaturverzeichnis

- [SST97] G. Saake, I. Schmitt, and C. Türker. *Objektdatenbanken — Konzepte, Sprachen, Architekturen*. International Thomson Publishing, Bonn, 1997.
- [STS98] K. Schwarz, C. Türker, and G. Saake. Specifying Advanced Transaction Models as Transaction Closures with Special Transaction Dependency Combinations. Preprint 5, Fakultät für Informatik, Universität Magdeburg, February 1998.
- [TCS96] C. Türker, S. Conrad, and G. Saake. Dynamically Changing Behavior: An Agent-Oriented View to Modeling Intelligent Information Systems. In Z. W. Raś and M. Michalewicz, editors, *Foundations of Intelligent Systems, Proc. of the 9th Int. Symposium on Methodologies for Intelligent Systems, ISMIS'96, Zakopane, Poland*, volume 1079 of *Lecture Notes in Artificial Intelligence*, pages 572–581, Berlin, June 1996. Springer-Verlag.
- [Weg09a] L. Wegner. Ratgeber für die gute Abschlussarbeit. Universität Kassel, 2009. Online unter: https://www.uni-kassel.de/eecs/fileadmin/datas/fb16/Zentrale_Ebene/Studienservice/Gem.Dokumente/RatgeberAbschlussarbeit.pdf (abgerufen am 2016-10-24).
- [Weg09b] L. Wegner. Ratgeber für die gute Abschlussarbeit. Universität Kassel, 2009. Online available: https://www.uni-kassel.de/eecs/fileadmin/datas/fb16/Zentrale_Ebene/Studienservice/Gem.Dokumente/RatgeberAbschlussarbeit.pdf (accessed on 2016-10-24).

Abkürzungsverzeichnis

DFG

...

Datenflussgraph

...

Abbildungsverzeichnis

1.1	Three Domains	1
1.2	Verteilung: Seitenanzahl-Note	3

Tabellenverzeichnis

1.1	Beispieltabelle mit langer Legende	2
1.2	Empfehlung zur Seitenanzahl der Arbeit	3