

Julius-Maximilians-Universität Würzburg  
Fakultät für Mathematik und Informatik  
Institut für Informatik  
Lehrstuhl für Informatik VI  
Künstliche Intelligenz und angewandte Informatik

# **Regelbasierte Attributierung von Personen in Romanen**

## **Ausarbeitung zum Master-Praktikum**

August 2015

von Titus Dose

Betreuer: Markus Krug

# Inhaltsverzeichnis

<b>1 Einleitung</b>	<b>1</b>
<b>2 Daten</b>	<b>1</b>
<b>3 Der Algorithmus</b>	<b>3</b>
3.1 Beschreibung des Algorithmus . . . . .	3
3.2 Beispiele . . . . .	4
<b>4 Evaluation und Fazit</b>	<b>6</b>
4.1 Test an dem Beispielroman „Mathilde Möhring“ . . . . .	6
4.2 Allgemeine Evaluation und Fazit . . . . .	9
<b>A Im Roman „Mathilde Möhring“ von Theodor Fontane gefundene Attribute</b>	<b>11</b>

# 1 Einleitung

In dem Praktikum, von dem diese Ausarbeitung berichtet, ging es um die regelbasierte Attributierung von Personen in Romanen. Es ging also darum, Eigenschaften wie zum Beispiel Alter, Geschlecht, Haarfarbe aber auch charakterliche Eigenschaften, die in Romantexten Erwähnung finden, zu finden und den Personen zuzuordnen, auf welche sie im Text bezogen werden. Der Ansatz hierfür war regelbasiert. Das heißt es wurden einfache, deterministische Regeln codiert, die auf der Basis von Hilfsannotationen aus beispielsweise Part-of-Speech-Tagging oder Dependency-Parsing arbeiten. Eine genauere Beschreibung der zur Verfügung stehenden Daten ist in Abschnitt 2 zu finden.

Es wird dann weiterhin das Programm und der verwendete Algorithmus in Abschnitt 3 vorgestellt, bevor dieser in Abschnitt 4 evaluiert und diskutiert wird. Auf Basis dessen wird in Abschnitt 4.2 ein kurzes Fazit gezogen.

Die Motivation hinter dem Praktikum ist offensichtlich: Grundsätzlich ist das Ziel, deutschsprachige Romane algorithmisch besser zu verstehen. Ein Baustein dafür ist die Zuordnung von Attributen zu Personen. Dem Versuch, dies mit einem regelbasierten Ansatz zu bewältigen, liegt zugrunde, dass ein Maschinenlernverfahren als Vorersetzung einen großen Umfang an annotierten Daten benötigen würde, worauf bei dem hier gewählten Verfahren verzichtet werden kann. Ferner ist ein regelbasiertes Vorgehen für Menschen wesentlich leichter verständlich, was beispielsweise auch die Fehleranalyse und -korrektur sowie die Evaluation des Programms erleichtert.

Ziel des Praktikums war es somit auch, durch einen ersten einfachen Ansatz, die Aufgabe besser zu verstehen und Erkenntnisse zu gewinnen, die bei weitergehenden und umfassenderen Ansätzen verwendet werden können.

# 2 Daten

Die Romane, anhand derer das in Abschnitt 3 beschriebene Programm entwickelt und evaluiert wurde, entstammen alle dem Projekt TextGrid<sup>1</sup>. Sie lagen in deutscher Sprache vor und stammten weitgehend aus dem Zeitraum 1600-1950.

Auf diesen Daten wurden drei Vorverarbeitungsschritte vorgenommen, ehe die Attributierung von Personen von dem eigentlichen Programm vorgenommen wurde. Diese drei Schritte umfassen

- NamedEntity-Erkennung: Es ist eine NamedEntity-Erkennung notwendig, um festzustellen, welche potentiellen Attribute auf eine Person bezogen sind. Informationen zu dem verwendeten Tool sind in [Jannidis et al., 2015] zu finden.
- Part-of-Speech-Tagging: Der im Abschnitt 3 beschriebene Algorithmus verwendet eine Regel, die insbesondere auf das Finden bestimmter Verben angewiesen ist. Dafür

---

<sup>1</sup><https://textgrid.de/projekt>

wurde der Part-of-Speech-Tagger von Helmut Schmid (vgl. [Schmid, 1995]) verwendet.

- Dependency-Parsing: Es wurde der Dependency-Parser aus dem Mate-Toolkit<sup>2</sup> verwendet. Eine Beschreibung des Parsers ist in [Bohnet, 2010] zu finden.

An dem folgenden Satz, der im Abschnitt 3 erneut diskutiert werden wird, wird beispielhaft gezeigt, wie die einzelnen Annotation aussehen:

*Das heißtt, eine so ganz richtige Mathilde war sie doch nicht, dazu war sie zu hager und hatte einen grisen Teint.*

- Die NamedEntity-Erkennung setzt NamedEntity-Annotationen für die Wörter „Mathilde“, „sie“ und „sie“. Dabei hat jede NamedEntity eine ID, welche von der Korefenzresolution vergeben wird. Diese wird von dem im folgenden Abschnitt beschriebenen Algorithmus jedoch nicht benötigt.
- Die Part-Of-Speech-Annotationen werden besonders für Verben mit der Grundform „haben“ oder „sein“ benötigt. Folgende POS-Annotation liegt für das erste Auftreten von „war“ vor:
  - begin: 7199
  - end: 7202; es sind im gesamten Roman alle Zeichen durchnummieriert und es wird hier angegeben, welche Zeichen, das Wort umfasst.
  - POSTag: VAFIN; das Wort „war“ ist an dieser Stelle ein finites Hilfsverb.
  - Lemma: sein; Es wird also zudem die Grundform eines jeden Wortes angegeben, diese wird insbesondere benötigt, wenn -wie in Abschnitt 3 nach Verben zur Grundform „haben“ oder „sein“ gesucht wird.
- Die Dependency-Annotationen werden am Beispiel des ersten Auftretens des Wortes „sie“ verdeutlicht. Folgende Informationen umfasst die Annotation.
  - begin: 7203
  - end: 7206
  - Headname: war; es wird also der Vater im Dependency-Parse-Tree angegeben
  - WordNumber: 9; die Nummer des Vaters im Parse-Tree macht den Vater erst eindeutig. Zu beachten ist, dass auch ein Komma eine eigene Nummer erhält.
  - DependencyRelation: SB; „sie“ ist ein Subjekt.

---

<sup>2</sup><https://code.google.com/p/mate-tools/>

## 3 Der Algorithmus

Es wird zunächst die Arbeitsweise des Algorithmus zum Attributieren von Personen beschrieben. Danach werden eine Reihe von Beispielen angegeben, die die Funktionsweise veranschaulichen.

### 3.1 Beschreibung des Algorithmus

Als Schwierigkeit beim Erarbeiten und Austesten von Regeln stellte sich die partielle Fehlerhaftigkeit der im vorausgegangenen Abschnitt geschilderten Daten dar. Es gab jedoch eine Regel, die dennoch zum einen brauchbare Informationen lieferte, zum anderen auch einen vertretbaren Recall hervorrief.

Diese stellt die Basis des gesamten Programms dar. Folgende grob geschilderte Vorgehensweise liegt ihr zugrunde. Dabei sind in der Schilderung die Aspekte vernachlässigt, die nur der Reaktion auf fehlerhafte Annotationen aus dem Vorverarbeitungsschrittes dienen.

- Berücksichtige in allen folgenden Schritten lediglich Sätze, welche keine Fragesätze sind.
- Finde alle Subjekte, welche eine NamedEntity darstellen.
- Laufe jeweils von dem NamedEntity-Subjekt im Dependency-Parse-Tree nach oben, bis eine Verbform erreicht wird.
- Ist diese eine Form von „sein“ oder „haben“, führe folgendes aus:
  - Falls kein Partizip Perfekt existiert, welches im Dependency-Parse-Tree ein Nachfolger der genannten Verbform ist und falls nach obiger Verform keine weitere Verbform im gleichen Teilsatz<sup>3</sup> auftritt, welche im Dependency-Parse-Tree Nachfolger der obigen Verbform ist, so speichere diesen Satz als SentencePart-Objekt<sup>4</sup>.

Damit sind die Kandidaten für Attribute generiert. Aus den meisten dieser Kandidaten wird tatsächlich ein Attribut generiert. Dazu sind diese Kandidaten aber zunächst weiterzuverarbeiten. Im Folgenden wird geschildert, wie dies für ein konkretes SentencePart-Objekt abläuft.

Zunächst enthält dieses -wie erwähnt- den gesamten Satz, das Subjekt und das Prädikat. In folgenden Schritten werden beispielsweise Relativsätze und direkte Reden gesondert behandelt. Es werden Relativsätze, die weder Subjekt noch Prädikat enthalten, entfernt. Genauso wird mit direkten Reden vorgegangen. Liegen umgekehrt sowohl Subjekt als auch Prädikat innerhalb einer direkten Rede oder innerhalb eines Relativsatzes, so wird nur die direkte Rede oder der Relativsatz weiterhin gespeichert.

---

<sup>3</sup>d.h. vor dem Auftreten des nächsten Kommas und vor der nächsten Konjunktion

<sup>4</sup>In einem solchen Objekt werden zunächst der gesamte Satz, sowie Subjekt und Prädikat gespeichert.

Zudem wird eine Heuristik angewendet, nach welcher alle Textabschnitte gelöscht werden, die sich vor dem hintersten Komma, das vor Subjekt und Prädikat steht, befinden und desgleichen alle Abschnitte, welche nach dem vordersten Komma, welches sowohl hinter dem Subjekt als auch dem Prädikat steht, auftreten. Diese Heuristik deckt in aller Regel auch alle die Schritte ab, die im Absatz zuvor zur Behandlung von Relativsätze und direkten Reden beschrieben wurden.

Es wird nun ein Attribute-Objekt erzeugt. Dieses enthält neben den verbliebenen Annotationen noch eine Liste von Phrasen, die im Konstruktor berechnet werden. Die Phrasen werden dabei wie folgt erzeugt:

- Berechne die Kinder des Prädikates im Dependency-Parse-Tree.
- Für jedes Kind bestimme alle Nachfolger im Dependency-Parse-Tree. Sortiere diese nach ihrer Reihenfolge, in der sie im Text auftreten. Somit erhalten wir für jedes Kind genau eine Phrase.

Damit ist das Attribut-Objekt vollständig. Wir werden im Folgenden ein Attribut, welches aus  $k$  Phrasen besteht, stets wie folgt durch einen String repräsentieren:

$$\text{verb}(\text{phrase}_1, \text{phrase}_2, \dots, \text{phrase}_k).$$

Zudem wird jedes Attribut als Annotation in das CAS eingefügt. Hierzu wurde im Coref-TypeSystem ein Typ *de.uniue.kalimachos.coref.type.AttributeTypeP1415* angelegt. Dieser ist vom Typ *uima.tcas.Annotation* und hat zwei Features: *verb*, welches das Verb des Attributs als *de.uniue.kalimachos.coref.type.POS* enthält, und *childPhrases*, welches vom Typ *uima.cas.FSArray* ist, wobei jeder Eintrag des Arrays genau eine Phrase kennzeichnet.

## 3.2 Beispiele

Die folgenden Beispiele sind alle dem Roman „Mathilde Möhring“ von Theodor Fontane entnommen.

**1. Beispielsatz:** *Das heißt, eine so ganz richtige Mathilde war sie doch nicht, dazu war sie zu hager und hatte einen grisen Teint.*

Der Algorithmus markiert die folgenden Attribute:

war(eine so ganz richtige Mathilde, sie, doch, nicht)  
war(dazu, sie, zu hager, und hatte einen grisen Teint)

Offenbar findet der Algorithmus die beiden Subjekte „sie“, welche jeweils auch eine Named Entity darstellen. Nun wird nach oben hin im Parse-Tree nach Verben gesucht, welche eine Form von „sein“ oder „haben“ sind. Dabei wird jeweils „war“ gefunden. Das Wort „hatte“ wird so nicht gefunden, da es nicht oberhalb von „sie“ im Parse-Tree liegt, sondern auf

Geschwisterebene. Die Konjunktion „und“ ist ein Kind von „war“ und alles unterhalb von ihr wird lediglich zu einer Phrase, wie auch an der Ausgabe erkenntlich ist: „und hatte einen grisen Teint“.

**2. Beispielsatz:** *Da war er besser dran, er hatte fünf Häuser, und das in der Georgenstraße war beinah schon ein Palais, vorn kleine Balkone von Eisen mit Vergoldung.*

Der Algorithmus markiert die folgenden Attribute:

war(Da, er, besser dran)  
hatte(er, fünf Häuser)

Es werden wiederum die beiden Subjekte, welche Named Entities sind, gefunden und es lieben die Prädikate „war“ und „hatte“ oberhalb dieser im Parse-Tree. Die Abschnitte nach den Komma werden ignoriert.

**3. Beispielsatz:** *Der schöne Mann mit dem Vollbart sah sich um, und wahrnehmend, daß die beiden Dinge fehlten, gegen die er eine tiefe Aversion hatte, Öldruckbilder und Antimakassars, war er sofort geneigt zu mieten, vorausgesetzt, daß er Aussicht hatte, für seine kleinen Bequemlichkeiten seitens der Wirtin gesorgt zu sehen.*

Der Algorithmus markiert die folgenden Attribute:

hatte(er, eine tiefe Aversion)  
hatte(daß, er, Aussicht)

Da der Algorithmus nach einem recht speziellen Muster sucht, entgehen ihm hier einige Attribute<sup>5</sup>. Die Attribute, die von dem Algorithmus gesetzt werden, sind jedoch unvollständig und für sich genommen sinnlos, da der Algorithmus recht großzügig Satzbestandteile entfernt. Im ersten Fall ist das Fehlen wesentlicher Bestandteile aber auch der komplexen Satzstruktur geschuldet. Im zweiten Fall wird ein an sich wertvolles Attribut -die Bequemlichkeit- gefunden, jedoch wieder entfernt, weil wesentliche Informationen erst nach dem Komma kommen, während Subjekt und Prädikat beide davor stehen.

**4. Beispielsatz:** Insgesamt sind jedoch viele im Text vorliegende Attribute genau so formuliert, dass der Algorithmus sie findet. Der folgende Beispielsatz ist wie gemacht für den Algorithmus.

*Weil er bequem ist, weil er keinen Muck hat, weil er ein Schlappier ist.*

Vom Algorithmus werden die folgenden Attribute gesetzt:

ist(er, bequem)  
hat(weil, er, keinen Muck)  
ist(weil, er, ein Schlappier,,)

Die Attribute werden offenbar leicht gefunden.

Eine Liste aller im besagten Roman gefundener Attribute findet sich im Anhang A.

---

<sup>5</sup>Diese werden jedoch teilweise von ihm an anderer Stelle herausgefunden.

## 4 Evaluation und Fazit

Die Evaluation des Algorithmus findet in zwei Schritten statt. Wir messen zunächst die Güte des Algorithmus im Hinblick auf unterschiedliche Gütekriterien für den schon im vorausgehenden Kapitel erwähnten Roman „Mathilde Möhring“. Dies zeigt auf anschauliche Weise einige Probleme und Stärken des Algorithmus.

Im zweiten Schritt werden ein paar allgemeine Aspekte hinsichtlich einer Evaluation angeprochen und zudem ein Fazit gezogen.

### 4.1 Test an dem Beispielroman „Mathilde Möhring“

Um den Algorithmus zahlenmäßig evaluieren zu können, wurde ein Gold-Standard für den Roman „Mathilde Möhring“ anhand zweier Inhaltsangaben beziehungsweise Personenkonstellationen entwickelt. Die Auswahl geeigneter Texte erweist sich jedoch als schwierig, weil in den meisten Zusammenfassungen wenig auf handfeste, äußere Kriterien wie etwa körperliche Merkmale, Alter und ähnliche Dinge eingegangen wird.

Für den vorliegenden Roman war die Suche jedoch einigermaßen erfolgreich. Es wurden zwei Quellen aus dem Internet verwendet<sup>6</sup>.

Im Folgenden sind die Attribute aus dem Gold-Standard aufgeführt und es ist gekennzeichnet, welche dieser Attribute von dem Algorithmus „entdeckt“ wurden, wobei bei der Entscheidung, ob der Algorithmus ein gegebenes Attribut ebenfalls herausgefunden hat, in einigen Fällen Interpretationsspielraum vorliegt.

---

<sup>6</sup>in erster Linie <http://www.studentshelp.de/p/referate/02/1944.htm> und ergänzend [https://de.wikipedia.org/wiki/Mathilde\\_Mhring](https://de.wikipedia.org/wiki/Mathilde_Mhring)

jeweils Version vom 15.6.2015

### **Mathilde Möhring**

jung	
pramatisch	
wenig anziehend	
lebt mit ihrer Mutter	
lebt in Berlin	
lebt in der Georgenstraße	
Vater ist tot	
aus armen Verhältnissen	
tüchtig	
zielstrebig	
fleißig	
stark	
statusbewusst	
nicht sensibel	
nicht emotional	
klug	x
hat Menschenkenntnis	
umgängliche Art	
nicht gelöst von ihrer Mutter	
geduldig	
enorm ehrgeizig	

### **Hugo Großmann**

Jura-Student	
kurz vor dem Examen (am Anfang)	
examiniert/Referendar (später)	x
liebt Lektüre literarischer Werke	
besucht oft das Theater	
vertäumt	x
bequem	x
schön	
schwach	x
kränklich	x
empfindsam	x
amüsiert sich gerne	
kindlich	
nicht konfliktfähig	
nicht selbstbewusst	
groß	
ängstlich	x
breitschultrig	x

### **Mutter Möhring**

arm  
sehr ängstlich  
geizig  
stets auf finanziellen Vorteil bedacht

### **Hans Rybinski**

Künstler

### **die alte Runtschen**

einäugig  
hässlich  
alt

Es ist leicht ersichtlich, dass nahezu alle gefundenen Attribute zu Hugo Großmann gehören. Dies hat seine Ursache unter anderem darin, dass einige Gespräche zwischen Mathilde und ihrer Mutter über Hugo erzählt werden, indem sich diese über Hugos Eigenschaften unterhalten.

Ferner muss beachtet werden, dass viele der Eigenschaften, die im Gold-Standard auftreten, im Roman selbst gar nicht zu finden sind, sondern Interpretationen sind. Insofern ist es für einen nicht unbeträchtlichen Teil der Attribute gar nicht möglich, diese aus dem Text zu extrahieren.

Insgesamt wurden somit 9 von 47 Attributen aus dem Gold-Standard entdeckt, was etwa 19% entspricht. Für Hugo Großmann wurden mit 8 von 18 Attributen sogar etwa 44% entdeckt. Es sei an dieser Stelle aber noch einmal darauf hingewiesen, dass

- einige der Attribute aus dem Gold-Standard von mehreren Attributen aus dem Output des Algorithmus abgedeckt werden.
- in einigen Fällen Attribute aus dem Gold-Standard in etwa getroffen werden, aber nicht genau der gleiche Inhalt auch aus den Attributen des Algorithmus hervorgeht. Dadurch ergeben sich schlechtere Werte, obwohl unter Umständen innerhalb des Textes gar keine Attribute gefunden werden können, die das jeweilige Attribut aus dem Gold-Standard genauer treffen, weil dieses oftmals auch eine Interpretation ist.
- in jeder Inhaltsangabe nur ein kleiner Teil der möglichen Attribute aufgeführt werden können, der Algorithmus jedoch auch einige nicht aufgeführte Attribute annotiert.
- der Algorithmus mit zum Teil stark fehlerbehafteten Daten (Dependency-Parsing, POS-Tagging) arbeitet.

Insgesamt hat der Algorithmus 315 Attribute gekennzeichnet, wie aus dem Anhang A ersichtlich ist.

## 4.2 Allgemeine Evaluation und Fazit

Viele Aspekte, die sich bei der Evaluation anhand des konkreten Romans ergeben haben, lassen sich auf die allgemeine Situation übertragen. Vor allem wird deutlich, dass eine sinnvolle Evaluation nur sehr eingeschränkt möglich ist.

Es ist zu erwarten, dass der Algorithmus bei weiteren Romanen stets nur einen kleinen Teil aller möglichen Attribute findet, da er nach einer sehr speziellen Situation sucht. Dennoch kann davon ausgegangen werden, dass er aus einem Roman von etwa dem gleichen Umfang wie „Mathilde Möhring“ eine nicht vernachlässigbare Menge an brauchbaren Eigenschaften von Personen extrahiert.

Es ist ferner bei der Bewertung des Algorithmus zu beachten, dass er eine äußerst schwierige Aufgabe zu bewältigen hat. Alleine aufgrund der unbegrenzten sprachlichen Möglichkeiten innerhalb der Gattung „Roman“, von denen innerhalb eines Zeitraums von über 300 Jahren durch unterschiedlichste Schriftsteller massiv Gebrauch gemacht wird, ist ein regelbasierter Ansatz sehr schwer umzusetzen.

Bei einem Maschinenlernverfahren kann ausgenutzt werden, dass für einen konkreten zu annotierenden Roman ein Modell anhand eines Korpus aus möglichst ähnlichen Romanen gelernt werden kann und damit spezifischer an die gegebene Situation angepasst werden kann. Ein regelbasierter Ansatz kann diese feinere Unterscheidung nicht vornehmen. Ferner sind sinnvolle Regeln aufgrund der Komplexität der Daten von Menschen nur sehr schwer zu erstellen, während mit den Rechenmitteln eines Computers derartige Regeln leichter erkannt werden können.

Der vorliegende Ansatz ist somit allenfalls in der Lage, einen weiter gefassten regelbasierten Ansatz oder auch ein maschinelles Lernverfahren zu ergänzen, indem er recht konzentriert oftmals sehr „griffige“ Eigenschaften erkennt, wie es der im Anhang befindliche Output für den Roman „Mathilde Möhring“ zeigt.

## Literatur

- [Bohnet, 2010] Bohnet, B. (2010). Very high accuracy and fast dependency parsing is not a contradiction. In *Proceedings of the 23rd International Conference on Computational Linguistics*, COLING '10, pages 89–97, Stroudsburg, PA, USA. Association for Computational Linguistics.
- [Jannidis et al., 2015] Jannidis, F., Krug, M., Reger, I., Weimer, L., Toepfer, M., and Puppe, F. (2015). Automatische Erkennung von Figuren in deutschsprachigen Romanen. [Automatic recognition of Characters in German novels] Digital Humanities im deutschsprachigen Raum (Dhd 2015). Graz, Austria.
- [Schmid, 1995] Schmid, H. (1995). Improvements in part-of-speech tagging with an application to german. In *In Proceedings of the ACL SIGDAT-Workshop*, pages 47–50.

## A Im Roman „Mathilde Möhring“ von Theodor Fontane gefundene Attribute

Die folgende Liste enthält alle Attribute, die von dem Algorithmus für den Roman „Mathilde Möhring“ von Theodor Fontane gesetzt werden, in der Reihenfolge, in der sie im Text auftraten.

Der Algorithmus berechnet ebenfalls eine Liste von Attributen, in der die Attribute jeweils ihrer Bezugsperson zugeordnet sind. Die Güte dieser Liste ist genau so gut wie die Güte der Koreferenzresolution.

war(Da, er, besser dran)  
hatte(er, fünf Häuser)  
war(eine so ganz richtige Mathilde, sie, doch, nicht)  
war(dazu, sie, zu hager, und hatte einen grisen Teint)  
hatte(das, sie)  
hatte(sie, wirklich, ein Gemmengesicht)  
hatten(Sie, einen Glanz)  
hatte(was, sie,)  
war(Unter, auch ein junger Mann von etwa sechsundzwanzig)  
war(Als, er, mit seiner Musterung, des Bildes, fertig)  
war(Er, nämlich, auf der Wohnungssuche)  
hatte(er, eine tiefe Aversion)  
hatte(daß, er, Aussicht)  
ist(er, bequem)  
hat(weil, er, keinen Muck)  
ist(weil, er, ein Schlappier,,)  
hast(Du, so viele Wörter)  
ist(dazu, er, viel zu bequem)  
ist(wenn, er, damit fertig)  
war(Ob, der Mann mit dem Vollbart in den Zelten, gewesen)  
war(daß, er, zwischen fünf und sechs, wieder, oben, bei Möhrings)  
war(daß, er, in den Zelten,,)  
hast(Rückenschmerzen, du, immer, und klagst in einem fort)  
hättest(Und, wenn, du, noch recht,)  
haben(wir, ja, doch, unser Auskommen, und bezahlen unsre Miete mit 'm Glockenschlag)  
bist(du, immer, so sicher)  
sind(Wir, artig und manierlich und immer gefällig, und laufen alle Gänge und sehen bloß)  
haben(wir, kein Klavierspiel)  
sei(er, sehr müde)  
waren(heute, sie, schon, um sechs, auf)  
sei(ob, ihr Mieter, nicht, ein Frühauf,)  
ist(daß, er, ein Frühauf,)

hat(der, einen Bärenschlaf)  
ist(daß, er, noch, so ein halber Student,)  
hast(du, recht)  
ist(weil, er, so breit,)  
ist(er, nicht, älter als sechsundzwanzig)  
ist(Er, bloß faul, und hat kein Feuer im Leibe)  
ist(bloß, weil, er, schläfrig,)  
ist(Und, sentimental, er, auch)  
haben(wir, ja, das Sparkassenbuch)  
bist(du, gut)  
bist(Wie, du, da, vorhin, vor dem Spiegel standst :, von der Seite, du, ganz hübsch)  
bist(Wie, du, da, vorhin, vor dem Spiegel standst :, von der Seite, du, ganz hübsch)  
sind(Da, wir, mal, einig)  
bist(Aber, du, für süß)  
war(Um sieben, Hugo Großmann)  
bin(Nachmittags, ich, immer, unterwegs)  
hast(Du, entschieden mehr vom Siebenschläfer als vom Landbriefträger)  
bin(daß, ich, wieder, da)  
hast(feine, Fühlung, du)  
hätte(er, was)  
hätte(wie, wenn, er, sie, von Wallenstein,,)  
war(Dazu und fror immer, er, viel zu bequem)  
war(als, er, eben Burgemeister geworden)  
war(als, ich, so, 'n, halbwachsner, Junge)  
war(dabei, dein Vater doch eigentlich ein famoser Knopp)  
war(das, er)  
sind(So im meisten, wir, uns gleich)  
war(Fürs, er, auch, nie)  
hatte(als, er, den Referendarius, hinter sich)  
war(seine Familie, doch, gewiß, dagegen, und wollte einen Minister aus ihm backen)  
hatte(mein Vater, doch, die halbe Quälerei, hinter sich)  
bin(ich, für das Übergehn aus dem einen, in den andern)  
bin(wenigstens, solange, ich, unberühmt)  
bist(du, der geborne Karl Moor)  
bin(wo, ich, heute, schon,)  
sind(So weit, wir, doch noch nicht, runter)  
hat(Und, sie, solch Profil)  
hatte(Am, Hugo, sein Repetitorium)  
war(am Abend, er, zu Haus)  
war(wenn, er, wieder, zu Hause)  
hatten(sie, jeden Morgen eine dünne Staubdecke)  
war(Selbst der Herr mit der polnischen Mütze, noch nicht, wieder, da)

ist(sie, doch noch sehr Mutter aus dem Volk)  
haben(wenn, wir, auch, verschiedene, Plätze)  
sind(so, wir, intim)  
habe(weil, er, noch hunderterlei zu tun und zu bedenken,)  
bin(ich, wieder, zu alt)  
bin(Ich, ja, nich)  
hatte(Die Mutter, kleine Bedenken, und sagte :)  
hat(er, soviel Anstand)  
ist(weil, er, kein Leichtfuß, und es ernst nimmt)  
war(Aber, er, nicht, da)  
war(Sie, aber)  
waren(Erst nach zehn, sie, zu Hause)  
hätte(als, ob, ich, nicht, gewollt,)  
ist(Ida, immer, da)  
ist(wenn, er, so)  
bist(du, so sicher)  
war(Als, aber, der Mieter, noch immer nicht, da)  
war(die Alte, immer noch, bei dem Stück)  
ist(Mir, so angst)  
habe(Aber, ich, solch Herzschlagen, und sehe immer den alten Mann)  
hat(der, nu seinen Denkzettel)  
war(Als, sie, wieder, drüben, in ihrem Zimmer)  
sind(daß, Sie, da,)  
bin(Von, ich, drüben, bei Hauptmann Petermann und von acht bis neun bei Kulickes,  
unten)  
ist(Um, er, immer, weg)  
ist(Aber, mitunter, er, auch, noch, da, und sieht so aus 'm Fenster)  
ist(Wenn, er, nicht, da)  
ist(wenn, er, da)  
haben(sonst, wir, einen Toten)  
war(Mathilde, dabei nicht, zugegen)  
ist(erzähl ein, ich, dir, andermal wenn's da,)  
haben(den, wir, mal, wieder, raus)  
haben(Wir, ja, doch, bloß, die eine)  
bist(du, eine alte Frau)  
haben(UND, da, wir ihn, denn)  
war(so was Interessantes, Vater, nich)  
war(er, ein sehr guter Mann)  
war(das, er)  
haben(daß, wir, die Chaiselongue,)  
sind(Kranke, so unruhig, und liegen mal hier und mal da)  
hat(er, was Edles)

war(Der Arzt, mit dieser Umlogierung, sehr zufrieden, und sagte)  
war(als, er, mit Hugo, allein)  
hätte(wenn, er, Lust)  
hatte(so, sie, jetzt, das Gefühl)  
war(so, er, im Leben selbst, doch)  
war(weil, sie, richtig)  
war(schließlich, er, doch, immer, ein Burgemeisterssohn mit Vollbart)  
bin(Ich, fürs Ernste)  
bist(du, noch nicht, fest auf den Füßen)  
bist(was, du)  
habe(wenn, ich, ihn, bloß, erst)  
bin(ich, nich, so dumm)  
ist(weil, er, noch, krank,)  
sei(mit euch < <)  
war(sie, wie verwachsen)  
hatte(sie, nicht Zeit)  
war(weil, sie, früher Schneiderin gewesen)  
war(weil, er, kein Spielverderber)  
sind(wir, sehr d'accord)  
sei(wenn, er, seinerseits als ein Mann des Baus auch die Ehe, als, deren Vorkammer, die Verlobung, anzusehen)  
haben(daß, wir, diese, hier)  
hat(das, er)  
war(Schultze, der einzige)  
hatte(weil, er, die zwei Steine, vorn und weil er Wirt ist, im Chemisette)  
ist(vorn, weil, er, Wirt)  
hat(daß, sie, die dicken blonden Zöpfe)  
war(weil, sie, vom frühen Morgen des andern Tages an, ein Gegenstand besonderer Aufmerksamkeit im ganzen Schultzeschen Hause und in der Nachbarschaft,)  
hatte(früher, sie, auch noch Pickel)  
war(wo, er, noch, wie rapplig,)  
war(Er, so)  
hatte(daß, ich, warm, Wasser, im Ofen,)  
war(Bloß, eine Braut, noch, da)  
war(Bloß, eine Braut, noch, da, und alle sehr drum rum)  
war(Fräulein, auch immer, bloß, um die Schüsseln rum und präsentierte)  
war(wenn, Ulrike, nich, da)  
wäre(als, wenn, er, nich, so recht zufrieden,,)  
wäre(als, ob, er, gar, nich so recht, da,)  
war(Sie, nun, Braut)  
war(Solange, sie, bloß, Fräulein, Thilde)  
hatte(sie, das Recht)

habe(wie, sie, Hugo zu trainieren)  
sind(wenn, wir, mal alt geworden)  
hättest(du, so nicht, den rechten Sinn)  
wärest(du, immer)  
habe(daß, ich, eine heitere lebenslustige Braut,)  
haben(Vielleicht, wir, den Kaiser vis à vis)  
wäre(daß, ich, gerne, mit dir, allein,)  
war(Hugo, glücklich)  
war(Hugo, damit und diesmal auch Thilde, zufrieden)  
war(daß, Hiller, kein Lokal für die Mutter,)  
sei(sie, für das schön Menschliche und in der Liebe für das Übermenschliche)  
war(Als Kosinsky, er, er, selbst.)  
war(Als Kosinsky, er, er, selbst.)  
wären(wir, Millionäre)  
war(weil, die alte Möhring, zwar, schon, wieder, außer Bett)  
hatte(Der Alte, solche Spezialitäten)  
hast(du, recht)  
bin(ich, doch, noch nicht, so alt)  
bin(ich, schon, so taprig)  
bin(ich, schon, so taprig, mitunter denk, ich)  
hast(Du, überhaupt so was Kleines und Ängstliches)  
hat(Daß, sie, das Pflaster)  
ist(Und, so hoch, er, doch, auch nich)  
hast(da ganz unrecht, du)  
hat(zu weiches Herz, er)  
hat(er, ein starkes menschliches Gefühl und beinah männlich)  
ist(ein so guter Mensch, er)  
habe(ich, die Beweise davon)  
habe(was, daß, ich, einen Titel,,)  
seien(wir, keine > > Leute < )  
sind(dann, wir, auch > > Leute < < )  
ist(Aurelie, ein sehr gutes Mädchen und auch nicht eng und nicht kleinlich)  
hat(sie, doch, so 'n sonderbares Honoratiorengefühl)  
hätte(wenn, sie, so was)  
hat(Sie, was Herbes)  
bin(was, ich, mir und andern schuldig,,)  
war(Sie, auch, abgebildet, auf dem Zettel)  
habe(daß, ich, so verwogene, Gedanken,)  
war(Als, er, oben)  
war(Die Alte, nicht, da)  
sind(bloß alt, sie)  
waren(Wir, nu von Heiligabend bis Silvester jeden Tag, aus, oder hatten unsren Punsch)

waren(einmal, wir, in einem ganz feinen Lokal)  
bin(Ich, gar nicht, so schlimm und so schrecklich vernünftig)  
bin(Ich, auch für Sichputzen und für Vergnügen)  
sind(Daß, wir, arme, Leute)  
bist(daß, du, nicht, reich)  
bin(ich, glücklich)  
bin(ich, es)  
war(der, schon immer, um Klock sechse, auf)  
bin(Dafür, ich, nicht)  
habt(ihr, da, ja, solche Fragehefte mit beigeschriebner Antwort)  
war(So schwach, er)  
war(was, sie, lustig)  
war(so uneinsichtig, er, nicht)  
hätte(daß, er, das,)  
war(eigentlich, er, froh)  
seien(Die Japaner, den Chinesen, doch, weit voraus)  
war(Die alte Möhring, immer)  
bin(um zwölf, ich)  
bin(ich, wieder, hier, und bringe dir deinen Kaffee)  
hatte(weil, sie, wie die meisten alten Berlinerinnen, das Bedürfnis der Aussprache)  
war(der, zu Ehren, des Tages, von einer Extrastärke,)  
hätten(weil, sie, das meiste Geld)  
war(Er, nun, Referendar)  
hatte(das schon auf dem Heimwege von der Examinationsstätte bis zur Georgenstraße  
gehabt, er,)  
war(er, verlegen)  
war(daß, sie, minder häuslich als früher und jeden Vormittag, ein paar Stunden in der  
Stadt,)  
bist(du, jetzt, immer, gerade weg)  
hättest(wie, wenn, du's, alles, schon in der Tasche,)  
bist(Du, doch, ein, und aus)  
ist(Um, er, auf der Station und um sechs, in Woldenstein in Westpreußen)  
hat(den Tag nach Johanni, sie, die Karte)  
waren(Da, sie beide, aber, halb polnisch und sehr hübsch)  
ist(Sie, Tragödin)  
hatte(das zum Polterabend ( ( aber ausfiel ), sie, geschickt,)  
waren(Am 26., mittags, sie, da)  
sein(Majestät, eine Freude,, dem Könige der ein Hort der Verfassung ist, zu der wir, alle  
stehn mit Leib und Leben)  
hast(den Justizrat, du nu, schon sicher, auf deiner Seite)  
habt(da, ihr, endlich, mal, einen richtigen Burgemeister)  
ist(er, ein Administrationstalent)

hat(er, Ideen < <)  
bist(Du, bloß zu ängstlich)  
bist(Du, immer wie im Traum)  
hätten(die Stadtverordneten, nicht gewollt)  
ist(Und, für sonst, sie, zu streng)  
ist(Silberstein, gegen das Militär)  
hast(so, du, den Kronenorden, weg)  
war(Er, aber, dagegen)  
war(sie, beinah, die einzige Unzufriedne in der Stadt)  
hat(Moses, die Priorität)  
ist(Hugo die Beurré grise, noch hart)  
ist(Er, ein sehr guter Herr und eigentlich liebenswürdig von Natur)  
ist(Er, ein feiner Herr)  
bin(ich, es)  
warst(du, ja, aus)  
war(ich, in der Ressource)  
hatten(Als, wir, damals, das, Gespräch)  
bist(wenigstens, wenn, du, dabei,)  
sei(daß, sie, sehr klug, und immer wisse)  
haben(Hugo dazu, wir's, denn doch noch nicht)  
ist(dazu, mir Mutter, zu schade)  
sind(Wir, doch)  
hat(obschon, Woldenstein, bloß dreitausendfünfhundert)  
bin(was, ich,)  
haben(Sie, eine famose Frau)  
war(Er, im Frack mit weit ausgeschnittner Weste)  
war(er, im Fieber, und hüstelte)  
war(er, ein sehr guter Schlittschuhläufer als solcher, und wollte sich in den Pausen zeigen)  
war(er, der reichste und angesehnste Mann der ganzen Gegend)  
war(er, enchantiert)  
war(er, enchantiert als sie seine Aufforderung)  
hatten(Da, wir, Maskenball)  
hatten(da, wir, Maskenball)  
habe(ich, gesehn)  
waren(Als, sie, fort)  
war(Der Arzt)  
war(Der Arzt, ] und hatte gegen Thildes Behandlung des Kranken : Brotrinde mit Essig-wasseraufguß)  
war(Einer, aus Breslau)  
war(wie geschickt, sie, und wie sie jeden einzelnen Nagel rausholte)  
hatten(als, wir, alles, raus)  
war(Sie, aber, noch, nich ganz zufrieden)

sind(sie, immer so sehr nach Fleisch)  
bist(daß, Du, ganz sicher,)  
bin(Ich, immer so sehr fürs Sichre)  
war(sie, immer)  
war(wenn, er, nicht, da)  
hat(sie, immer solche neuen Sätze)  
bin(Ich, für einen)  
war(so, kurze, Zeit, er, in der Stadt)  
waren(der, weil, Osterferien)  
war(der, er, immer,)  
war(Denn, er, aus einem sehr guten Hause)  
wäre(als, ob, ich, wunder, was,)  
hatte(er, so was Edles)  
war(Daß, er, bloß schwächlich)  
war(daß, er, sehr gut)  
war(So, daß, selbst, Pastor Lämmel, zufrieden, und ihm die Hand gab)  
bist(damit, Du, ganz sicher, und die Runtschen zu rechter Zeit bestellen kannst)  
hat(sie, bloß schwarze)  
hatte(ich, den Wind ins Gesicht und wollte nicht zumachen)  
hat(Silberstein, mir alles besorgt, und hatte alles auf Lager)  
hat(Silberstein, es, doch, auch nich, umsonst)  
habe(Denn, wenn, ich, auch die Pension)  
ist(dann, Mutter, die klügste Frau)  
bin(von Natur, ich, gradeso, wie Mutter)  
hat(daß, er, mehr Einfluß auf mich gehabt als ich)  
war(weil, die Runtschen, seine einzige Renonce,)  
hatten(was, wir, heute)  
hättest(das, du, gewiß)  
hast(du, ja, Mitleid mit jedem und mit mir auch)  
bist(du, gut)  
war(er, katholsch)  
hatte(sie, ja, nun, einen Titel)  
warst(du, ja, heute, wieder, da)  
hatte(solange, sie, den Wiederverheiratungsplan)  
bist(Und, nun, du, gradeso, wie die andern Menschen)  
bin(Ich, kein Fräulein, und habe)  
hätte(wenn, ich, dich, nicht,,)  
haben(wir, es)