

Todo list

Struktur vom Dokument erläutern	3
Definition: 'Gemeinsam' := "etwas, was beide anwender haben und gleich ist"	3
Definition: Vorwissen := "Wissen aus vorherigen Phasen"	3
Definition: Entscheidungsbaum	3
Definition: Attribut	3
Schaltkreis	3
Annahme: ehrliche anwender := "handeln nach protokoll"	3
Vision der Anwendung	3
Festlegen, wie die E-Mails ins Programm kommen	3
Festlegen, wie das Programm verteilt wird und die Teile kommunizieren .	3
Kurz die Einzelnen phasen der Anwendung beschreiben	4
Einleitung, Verweisen auf Figure für gemeinsame Wortliste	5
Für section-Titel besseren Begriff für "Vorkommnisse der Worte in eigenen Spam/Nicht Spam E-Mails" finden	5
Content	5
Content	5
Content	5
Einleitung, auf Figure für Schwellwerte verweisen	6
Für section-Titel besseren Begriff für "Vorkommnisse der Worte in eigenen Spam/Nicht Spam E-Mails" finden	6
Content	6
Content	6
Content	6
Einleitung, Figure referenzieren	7
Content	7
verteiltes ID3 beschreiben	8
Yaos algorithmus grundlegend zusammenfassen (garbled decision table, garbled gate, garbled circuit	8
Feststellung der benoetigten Bytezahl beschreiben	8
Schaltkreis designen: Maximum von Summen von positiven Zahlensequen- zen	8
Schaltkreis designen: Gleichheit.	8
Schaltkreis für $x * \log x$ -Protokoll aus dem Paper zusammenfassen	8
Vorghehen zusammenfassen, Schaltkreis aus dominierender Ausgabe wiederver- wenden	8
Einleitung: Wir brauchen ein Programm, was den Klassifikator auf eine MAil oder Mails anwendet	9
Anhand der Definition von Attributen und Entscheidungsbäumen beschrei- bungssprache fuer Entscheidungsbaum herleiten	9
Arbeitsweise des Klassifikators erklären	9

Contents

1	Einleitung	3
1.1	Begriffe	3
1.2	Annahmen	3
2	Grundlagen der Anwendung	3
2.1	Form der Benutzereingabe	3
2.2	Interaktion der verteilten Programme	3
2.3	Phasen der Anwendung	4
3	Finden der gemeinsamen Wortliste	5
3.1	Berechnung der Vorkommnisse	5
3.2	Sortierung der Worte nach Informationsheuristik	5
3.3	Syncronisierung der Wortlisten	5
4	Finden der gemeinsamen Schwellwerte	6
4.1	Berechnung der Vorkommnisse	6
4.2	Bestimmung der eigenen Schwellwerte	6
4.3	Syncronisierung der Schwellwerte	6
5	Diskretisieren der eigenen E-Mails	7
6	Lernen der gesamten E-Mails	8
6.1	Feststellen der dominierenden Ausgabe	8
6.2	Feststellen ob Ausgabe eindeutig	8
6.3	Das Entropien-Protokoll	8
6.4	Attribut mit maximalem Informationsgewinn finden	8
7	Verwenden des Klassifikators	9
7.1	Eingabe des Klassifikators	9
7.2	Arbeitsweise des Klassifikators	9

1 Einleitung

Struktur vom Dokument erläutern

1.1 Begriffe

Definition: Eigenes, Gesamtes

M sei eine Menge von Elementen, die in zwei Teilmengen M_A und M_B zerfällt, sodass $M = M_A \cup M_B$ ist. Wir nehmen desweiteren an, dass Alice M_A kennt, aber weder M noch M_B und dass Bob M_B kennt, aber weder M noch M_A . Dann bezeichnen wir:

- M als **gesamtes** Wissen
- M_A als das **eigene** Wissen von Alice
- M_B als das **eigene** Wissen von Bob
- M_B als das **andere** Wissen von Alice
- M_A als das **andere** Wissen von Bob

Definition: 'Gemeinsam' := "etwas, was beide anwender haben und gleich ist"

Definition: Vorwissen := "Wissen aus vorherigen Phasen"

Definition: Entscheidungsbaum

Definition: Attribut

Schaltkreis

1.2 Annahmen

Annahme: ehrliche anwender := "handeln nach protokoll"

2 Grundlagen der Anwendung

Vision der Anwendung

2.1 Form der Benutzereingabe

Festlegen, wie die E-Mails ins Programm kommen

2.2 Interaktion der verteilten Programme

Festlegen, wie das Programm verteilt wird und die Teile kommunizieren

2.3 Phasen der Anwendung

Kurz die Einzelnen phasen der Anwendung beschreiben

3 Finden der gemeinsamen Wortliste

Einleitung, Verweisen auf Figure für gemeinsame Wortliste

Für section-Titel besseren Begriff für "Vorkommnisse der Worte in eigenen Spam/Nicht Spam E-Mails" finden

3.1 Berechnung der Vorkommnisse

Content

3.2 Sortierung der Worte nach Informationsheuristik

Content

3.3 Synchronisierung der Wortlisten

Content

4 Finden der gemeinsamen Schwellwerte

Einleitung, auf Figure für Schwellwerte verweisen

Für section-Titel besseren Begriff für "Vorkommnisse der Worte in eigenen Spam/Nicht Spam E-Mails" finden

4.1 Berechnung der Vorkommnisse

Content

4.2 Bestimmung der eigenen Schwellwerte

Content

4.3 Synchronisierung der Schwellwerte

Content

5 Diskretisieren der eigenen E-Mails

Einleitung, Figure referenzieren

Content

6 Lernen der gesamten E-Mails

verteiltes ID3 beschreiben

Yaos algorithmus grundlegend zusammenfassen (garbled decision table, garbled gate, garbled circuit)

Feststellung der benoetigten Bytezahl beschreiben

6.1 Feststellen der dominierenden Ausgabe

Schaltkreis designen: Maximum von Summen von positiven Zahlensequenzen

6.2 Feststellen ob Ausgabe eindeutig

Schaltkreis designen: Gleichheit.

6.3 Das Entropien-Protokoll

Schaltkreis für $x \cdot \log x$ -Protokoll aus dem Paper zusammenfassen

6.4 Attribut mit maximalem Informationsgewinn finden

Vorghehen zusammenfassen, Schaltkreis aus dominierender Ausgabe wiederverwenden

7 Verwenden des Klassifikators

Einleitung: Wir brauchen ein Programm, was den Klassifikator auf eine MMail oder Mails anwendet

7.1 Eingabe des Klassifikators

Anhand der Definition von Attributen und Entscheidungsbäumen beschreibungssprache fuer Entscheidungsbaum herleiten

7.2 Arbeitsweise des Klassifikators

Arbeitsweise des Klassifikators erklären

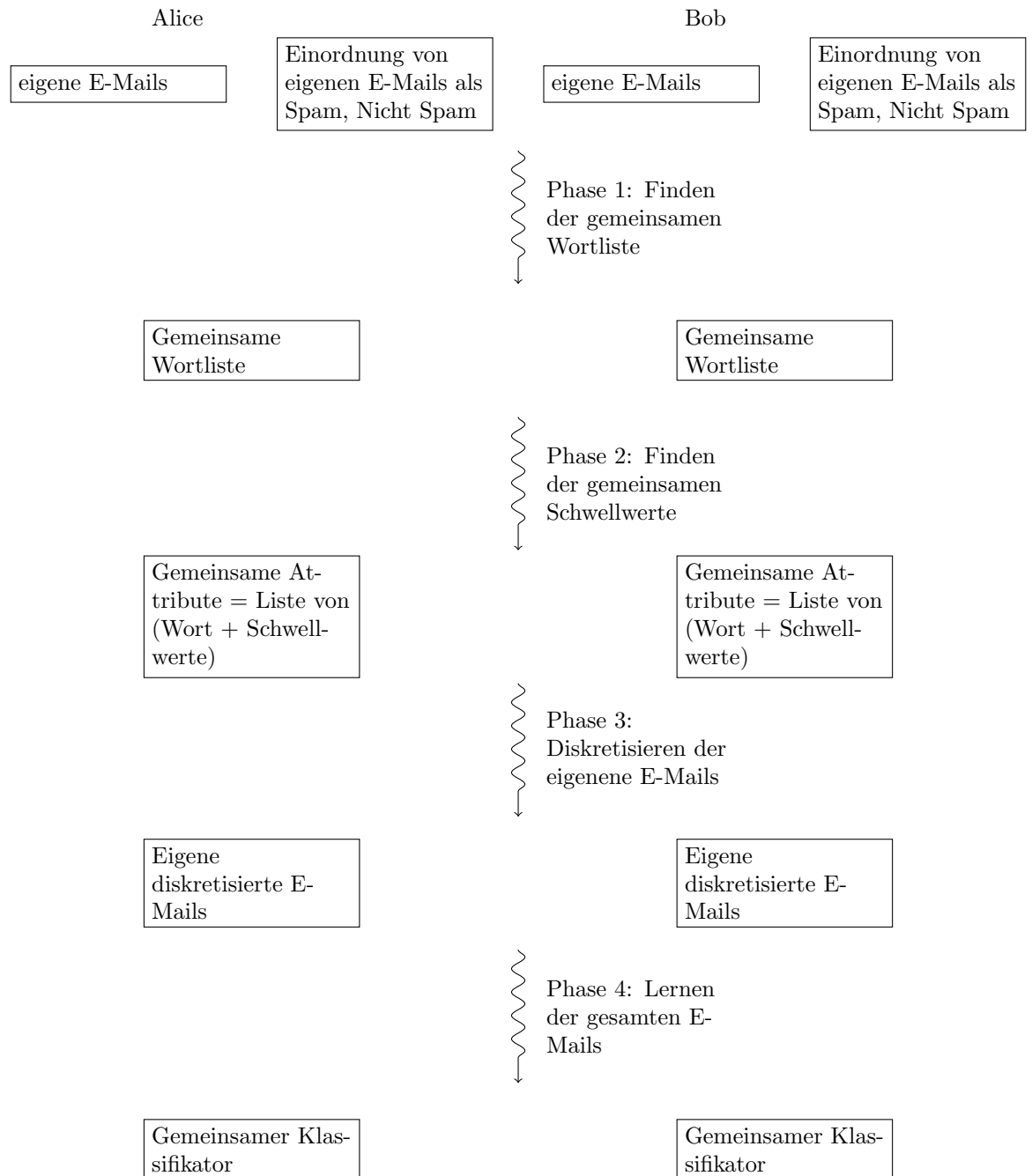


Figure 1: Phasen der Anwendung

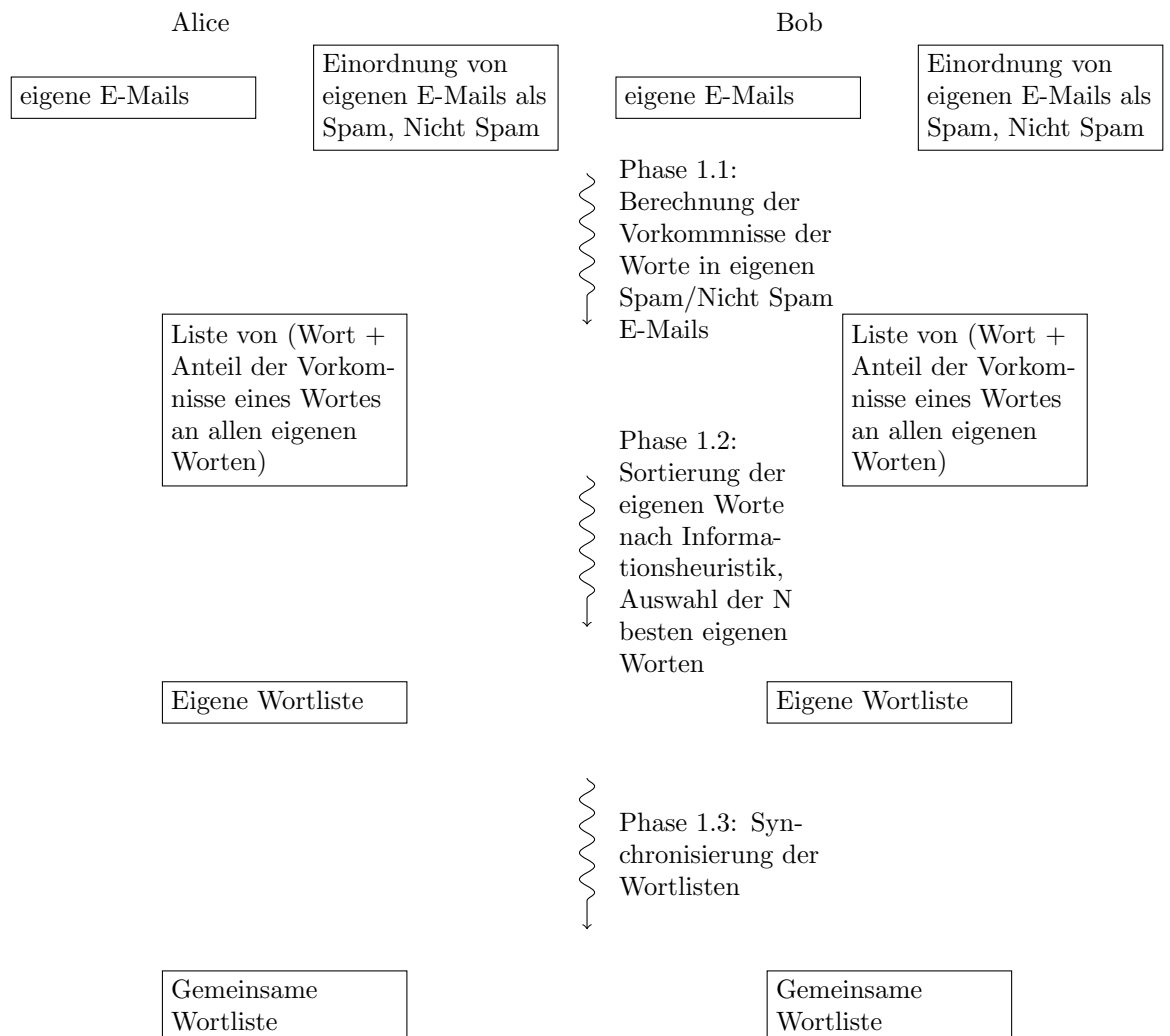
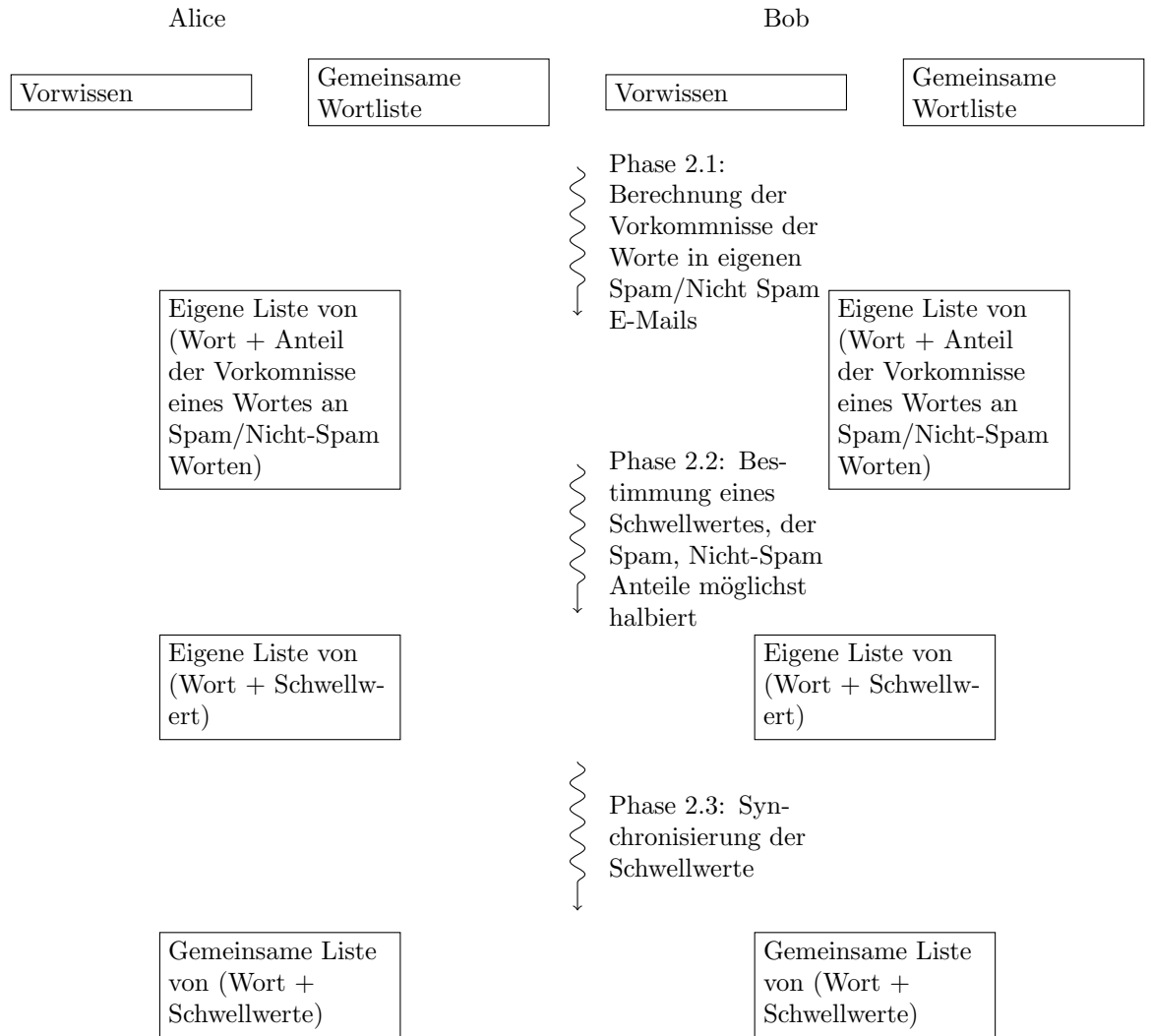


Figure 2: Schritte zum Berechnen der gemeinsamen Wortliste



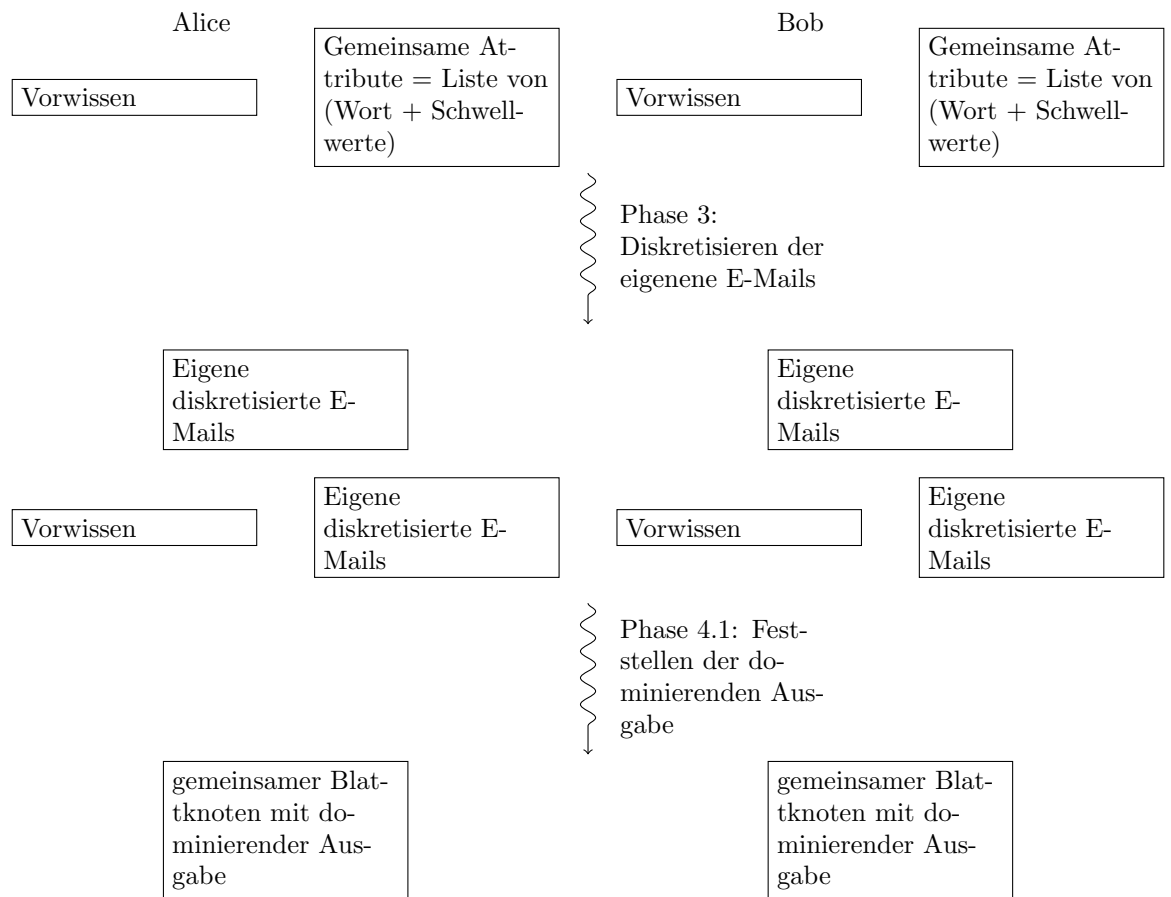


Figure 3: ID3-Algorithmus, Fall 1: Keine Attribute mehr vorhanden

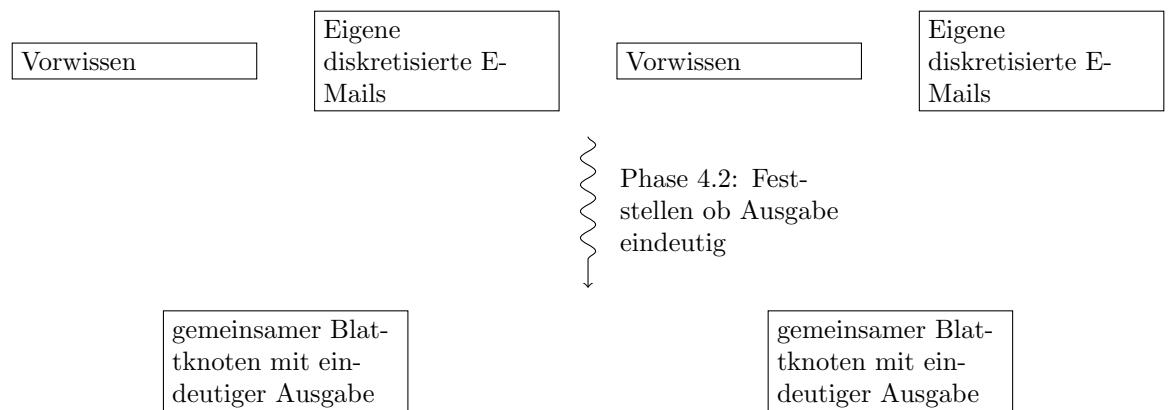


Figure 4: ID3-Algorithmus, Fall 2: Ausgabe eindeutig

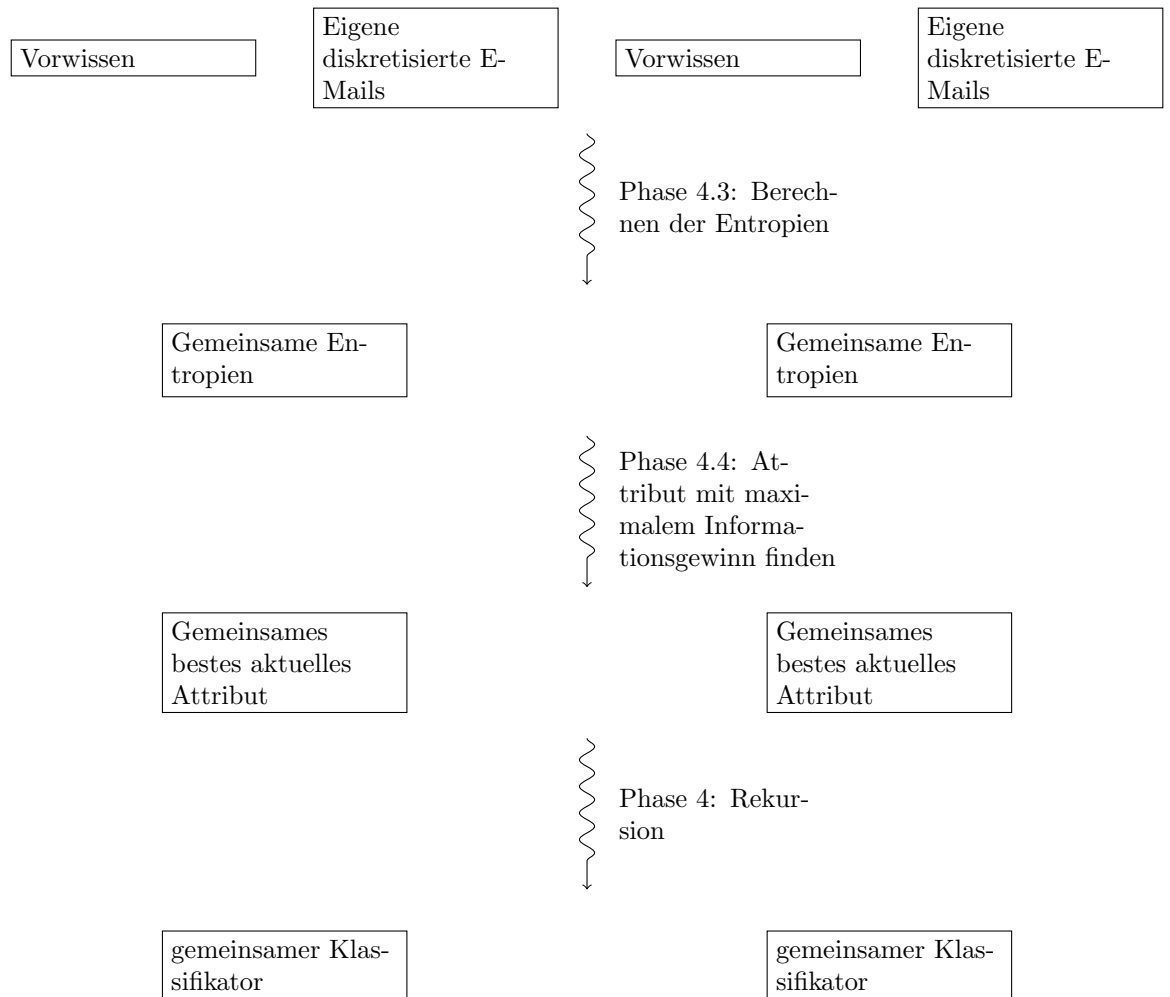


Figure 5: ID3-Algorithmus, Fall 3: Erzeugung eines Astes