# Übungsblatt 1 – Algorithmisches Lernen

Abgabe 30.04.2001 vor der Vorlesung

## *Aufgabe 1.1*

a) Gegeben ist eine Inselgruppe $I := \{I_1, \ldots, I_n\}$ mit jeweils $m_i \quad für\ i = 1 \ldots n$ pro Insel spezifischen Tierarten, die eine Insel eindeutig charakterisieren und dementsprechend auf keiner anderen Insel vorkommen. Daneben gibt es auch Tierarten, die auf allen Inseln gleichermaßen vorkommen, für die Unterscheidung der Inseln also keinen Informationsgehalt liefern. Eine Insel kann also anhand eines dort einzigartigen Tiers identifiziert werden. Aus der Aufgabenstellung ist die eine Insel beschreibende Definition eines Konzepts übernommen

$$C_x := \{t_1, \ldots, t_k\} \cup \{u_{x,1}, u_{x,2}, \ldots, u_{x,j_x}\} \tag{1.1}$$

Für das Erlernen eines Konzepts wird eine erste Hypothese $h_0$ gewählt, die alle Literale enthält und somit für alle Inseln zutrifft

$$\begin{aligned} h_0 = t_1 \wedge t_2 \wedge \ldots \wedge t_k \quad \wedge \quad & u_{1,1} \wedge u_{1,2} \wedge \ldots \wedge u_{1,m_1} \\ & \wedge u_{2,1} \wedge u_{2,2} \wedge \ldots u_{2,m_2} \\ & \ldots \\ & \wedge u_{n,1} \wedge u_{n,2} \wedge \ldots u_{n,m_n} \end{aligned} \tag{1.2}$$

Die Gottheit wird jetzt Gegenbeispiele liefern, die leicht von der aktuellen Hypothese abweichen. Dabei wird die Gottheit darauf achten, dass sich möglichst nur ein Literal unterscheidet, um möglichst viele Gegenbeispiele liefern zu können und damit viele Räucherstäbchen zu erhalten. Ein erstes Gegenbeispiel könnte so aussehen:

$$\begin{aligned} y_0 = t_1 \wedge t_2 \wedge \ldots \wedge t_k \quad \wedge \quad & \neg u_{1,1} \wedge u_{1,2} \wedge \ldots \wedge u_{1,m_1} \\ & \wedge u_{2,1} \wedge u_{2,2} \wedge \ldots u_{2,m_2} \\ & \ldots \\ & \wedge u_{n,1} \wedge u_{n,2} \wedge \ldots u_{n,m_n} \end{aligned} \tag{1.3}$$

In diesem Gegenbeispiel ist das Literal $u_{1,1}$, dass das erste Tier auf Insel 1 repräsentiert, negiert. Damit sagt die Gottheit dem Lerner, dass Insel 1 nicht in Frage kommt und alle Literale für Insel gestrichen werden können.

Die Gottheit könnte dem Lerner nach dem positiven Gegenbeispiel im Folgenden weitere Gegenbeispiele für Insel 1 geben, die aber negativ und redundant wären. Der Lerner wird nur positive Gegenbeispiele akzeptieren und so für jede Menge an Literalen pro Insel nur ein Gegenbeispiel benötigen. So gelangt er nach n Gegenbeispielen und damit n Räucherstäbchen zu einem Konzept und hat dieses erlernt:

$$Gegenbeispiel_C(C) = n \tag{1.4}$$

Für das Erlernen aller n Konzepte $C_x$ der Konzeptklasse C benötigt der Lerner dementsprechend $n^2$ Räucherstäbchen.

b) Die VC-Dimension ist definiert als die Kardinalität der größten Beispielmenge S, die von einer Konzeptklasse gerade noch zertrümmert werden kann. Formal mit $P(S) = \{S \cap c \mid c \in C\}$ hergeleitet. Für Monome und Monotonmonome der Länge n gilt, dass $VC(MONOM_N) = VC(MONOTON\text{-}MONOM_N) = n$ für $n \geq 2$.

Betrachtet man die Konzeptklasse C der Inseln, dann kann diese als ein Monotonmonom betrachtet werden, wobei k Literale durch die gemeinsamen Tierarten vorgegeben sind und

$$m_{all} = \sum_i^n m_i \tag{1.5}$$

Literale durch die spezifischen Tierarten. In der Summe erhält man für die VC-Dimension

$$VC(C) = k + m_{all} \tag{1.6}$$

c) Naja, das Folgende trifft es jetzt wohl nicht ganz, aber dennoch eine Überlegung: Wenn man bedenkt, dass eine auf der Insel spezifische Tierart eine Insel eindeutig identifiziert, könnten Konzepte und Hypothesen mit wesentlich weniger Literalen geschrieben werden. Zudem kann auf die ohnehin identischen Literale für die gemeinsamen Tiere verzichtet werden. Somit ergäben sich Konzepte mit

$$C_{X2} := \{u_{x,1}\} \tag{1.7}$$

### *Aufgabe 1.2*

Die Konzeptklasse C mit Konzepten

$$C_K := \{(x, y) \in \mathbb{R}^2 \mid (x, y) \text{ liegt in D}\} \tag{1.8}$$

Beschreibt Kreise in der Ebene. Die Frage ist, wie groß ihre VC-Dimension ist. Damit muß geklärt werden, wie groß die Kardinalität der größten Beispielmenge ist, die von der Konzeptklasse noch zertrümmert wird. Die Beispielmenge besteht dabei aus Punkten in der Ebene.

Für Beispielmengen mit keinem, einem, zwei oder drei Punkten sind Kreise aus C gegeben, die die Potenzmenge der Beispielmenge zertrümmern. Für Beispielmengen der Kardinalität sind keine Überdeckungen der Potenzmenge mehr möglich, wie im Folgenden illustriert wird. Die VC-Dimension für C beträgt daher

$$VC(C) = 3 \tag{1.9}$$

Sind die vier Punkte einer Beispielmenge $S = \{x_1, x_2, x_3, x_4\}$, so können diese entweder in einem Kreis angeordnet sein oder drei von ihnen ein Dreieck bilden mit dem vierten Punkt in ihrer Mitte (also konkav und konvexe Vierecke, wenn man die Punkte zu einem solchen verbindet).

Im zweiten Fall gibt es keinen Kreis, der die drei Punkte nicht ohne den inneren abgedeckt und somit kann damit diese Beispielmenge S nicht durch C zertrümmert werden. Das eingezeichnete Dreieck im linken Bild verdeutlicht dies.

Für den ersten Fall, bei dem die vier Punkte  im Kreis angeordnet sind, gibt es keinen Kreis aus C, der nur die beiden jeweils diagonalen Punkte überdeckt und die beiden verbliebenden Punkte ausschließt.

Die maximale Anzahl von mit dieser Konzeptklase zertrümmerten Punkten beträgt daher 3.

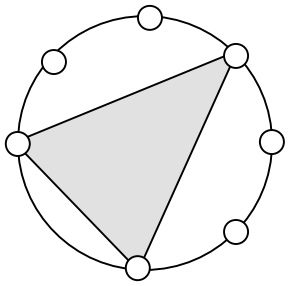
***Aufgabe 1.3***

Für die Konzeptklasse C der Dreiecke in der Ebene

$$C_D := \{(x, y) \in \mathbb{R}^2 \mid (x, y) \text{ liegt in D}\} \tag{1.10}$$

kann eine ähnliche Betrachtung wie für die Klasse der Kreise gemacht werden.

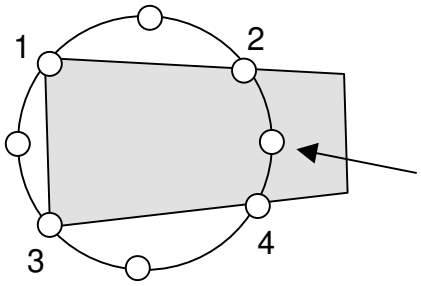
Zunächst kann gesagt werden, dass die Klasse der Dreiecke Beispielmengen größerer Kardinalität zertrümmern kann als die Klasse der Kreise. Dazu beschreibt man einen Kreis, der alle Punkte der Beispielmenge verbindet (sollte das nicht möglich sein, scheidet diese Beispielmenge zum Zertrümmern aus). In folgendem Schaubild wird dies illustriert



Für jede Kombination von Punkten lassen sich Dreiecke aus der Konzeptklasse finden, so dass Beispielmengen mit bis zu sieben Elementen von der Konzeptklasse zertrümmert werden können. Dies stellt gleichzeitig die größte Kardinalität dar und somit gilt

$$VC(C) = 7 \tag{1.11}$$

Mehr als sieben und damit beispielsweise acht Punkte können in keinem Fall zertrümmert werden, wie das folgende Bild verdeutlicht




Die Punkte 1, 2, 3 und 4 werden durch kein Dreieck abgedeckt werden, ohne diesen diesen Punkt einzuschließen!