

## Aufgabe zum Programmwurf zur Vorlesung „Grundlegende Algorithmen und Verfahren der KI“

### Aufgabe 1

Es gilt folgender Satz: Jede aussagenlogische Formel in disjunktiver Normalform kann durch ein zweischichtiges, vorwärtsgetriebenes neuronales Netz realisiert werden.

Dies wollen wir uns klarmachen und dann das entsprechende Netz implementieren. Dazu folgende Eckpunkte / Anforderungen an die Architektur des Netzes:

- Wir gehen davon aus, dass die zugehörige DNF aus  $m$  Monomen mit jeweils  $n$  Aussagenvariablen  $x_1, \dots, x_n$  besteht. Dann sind die Eingangsneuronen des Netzes auch  $x_1, \dots, x_n$ .
- Die Eingangsneuronen können die Werte  $-1$  und  $1$  annehmen, also  $x_i \in \{-1, 1\}$  für  $i = 1, \dots, n$ . Dabei entspricht der Wert  $-1$  dem logischen falsch und der Wert  $1$  dem logischen wahr.
- Die Zwischenschicht besitzt  $m$  Neuronen  $z_1, \dots, z_m$  und die Ausgabeschicht ein Neuron  $y$ .
- Die verwendete Aktivierungsfunktion ist die Vorzeichenfunktion

$$\text{sgn}(x) = \begin{cases} 1 & \text{falls } x \geq 0 \\ -1 & \text{falls } x < 0 \end{cases}.$$

- Die Propagation im Netz lautet in Vektorform:

$$\begin{aligned} \underline{z} &= \text{sgn}(w \cdot \underline{x} - \underline{v}) \\ y &= \text{sgn}(W \cdot \underline{z} - V) \end{aligned}$$

mit den Gewichtsmatrizen  $w \in \mathbb{R}^{(m,n)}$  und  $W \in \mathbb{R}^{(1,m)}$  sowie den Schwellwerten  $\underline{v} \in \mathbb{R}^m$  und  $V \in \mathbb{R}$ . Die Aktivierungsfunktion ist dabei komponentenweise zu nehmen.

- Der Lernalgorithmus für die Gewichte und die Schwellwerte ist die Fehlerrückübertragung mit

$$\begin{aligned} \Delta W_{1j} &= \eta \cdot (p - y) \cdot z_j & \Delta V &= -\eta \cdot (p - y) \\ \Delta w_{jk} &= \eta \cdot W_{1j}(p - y) \cdot x_k & \Delta v_j &= -\eta \cdot (p - y) \cdot W_{1j} \end{aligned}$$

wobei  $p$  das Zielmuster ist und  $\eta > 0$  die Lernrate.

Wie müssen die Gewichte und Schwellwerte gewählt werden, damit das Netz eine gegebene DNF realisiert? Dazu folgende Erläuterungen:

- Um einen Transfer zwischen DNF und Netz zu erhalten muss die aussagenlogische Formel arithmetisiert werden.
- Sei  $F = M_1 \vee M_2 \vee \dots \vee M_m$  die DNF in den  $n$  Aussagenvariablen  $x_1, \dots, x_n$  sowie den  $m$  Monomen  $M_1, \dots, M_m$ .
- Jedes Monom  $M_i$  habe die Form  $M_i = l_{i1} \wedge l_{i2} \wedge \dots \wedge l_{in}$  mit den Literalen  $l_{ij}$ . Dabei kann  $l_{ij}$  entweder  $x_j$ ,  $\neg x_j$  oder  $w$  sein für  $i = 1, \dots, m$ .
- Dann gilt  $M_i = l_{i1} \wedge l_{i2} \wedge \dots \wedge l_{in}$  ist wahr genau dann, wenn  $a_{i1} \cdot x_1 + a_{i2} \cdot x_2 + \dots + a_{in} \cdot x_n = b_i$  gilt mit

$$a_{ij} = \begin{cases} 1 & \text{falls } l_{ij} = x_j \\ 0 & \text{falls } l_{ij} = w \\ -1 & \text{falls } l_{ij} = \neg x_j \end{cases}, \quad b_i = \sum_{k=1}^n |a_{ik}|$$

für  $i = 1, \dots, m$  und  $j = 1, \dots, n$ .

- Die Monome werden durch die Neuronen  $z_j$  der Zwischenschicht realisiert. Dabei gilt  $F = M_1 \vee M_2 \vee \dots \vee M_m$  ist wahr genau dann, wenn  $(z_1 + 1) + (z_2 + 1) + \dots + (z_m + 1) \geq 1$  gilt, oder  $z_1 + z_2 + \dots + z_m \geq 1 - m$ .

Wählen Sie eine zufällige DNF  $F$  mit mindestens 5 Monomen, mindestens 3 Literalen pro Monom und insgesamt mindestens 10 Aussagenvariablen  $x_i$  in  $F$  aus.

- Erstellen Sie eine Implementierung des Netzes mit Fehlerrückübertragung, welche  $F$  realisiert, unter Beachtung der Anforderungen.
- Geben Sie eine Beschreibung (Dokumentation) der implementierten Funktionalität an.
- Testen Sie Ihre Implementierung. Dazu gehen Sie folgendermaßen vor: Setzen Sie die Gewichte und Schwellwerte initial so, wie in den Erläuterungen beschrieben, dass die DNF direkt implementiert ist. Dann dürfen sich die Gewichte und Schwellwerte bei Eingabe eines Musters (Belegung) nicht mehr ändern.  
Dokumentieren und bewerten Sie die Testergebnisse.
- Nun werden die Gewichte und Schwellwerte leicht verändert. Starten Sie eine Lernsequenz für das Netzes und beobachten Sie, ob das Netz die korrekte Funktion lernt. Schließlich setzen Sie die Gewichte und Schwellwerte initial auf zufällige Werte und wiederholen Sie das Experiment.

Dokumentieren und analysieren Sie die erhaltenen Ergebnisse.