**Eidgenössische Technische Hochschule Zürich**
*Ecole polytechnique fédérale de Zurich*
*Politecnico federale di Zurigo*
*Swiss Federal Institute of Technology Zurich*

Dr. Jakob Bernasconi

## Informationsverarbeitung in neuronalen Netzwerken

Frühjahrssemester 2009

# Übung 5

Ausgabe am 24. April 2009
Abgabe am 8. Mai 2009

## 1. Das "Salon-Problem"

Das "Salon-Problem" ist eine Zwei-Personen-Variante des berühmten "El Farol Bar Problems" (W.B. Arthur, 1994). Es wurde 1998 von Reiner Franke (damals Uni Bremen) erfunden und wie folgt formuliert.

Die Marquise M hat zwei konkurrierende Verehrer (Spieler A und B), die sich jeden Abend (unabhängig voneinander!) entscheiden, ob sie in den Salon der Marquise gehen wollen oder doch lieber in die Oper, d.h. sie können wählen zwischen S (= Salon) und O (= Oper). Ein Abend im Salon der Marquise ist den beiden Kontrahenten natürlich viel mehr wert (z.B. $r = 1$) als ein Besuch der Oper (z.B. $r = 0.2$), aber nur, wenn der andere nicht auch dort ist. Sonst macht ein Salon-Besuch gar keinen Spass ($r = 0$).

Die Payoff-Matrix ($r_A$, $r_B$) = (Reward von A, Reward von B) sieht also folgendermassen aus:

| | | Spieler B | |
|---|---|---|---|
| | | S | O |
| Spieler A | S | (0, 0) | (1, 0.2) |
| | O | (0.2, 1) | (0.2, 0.2) |

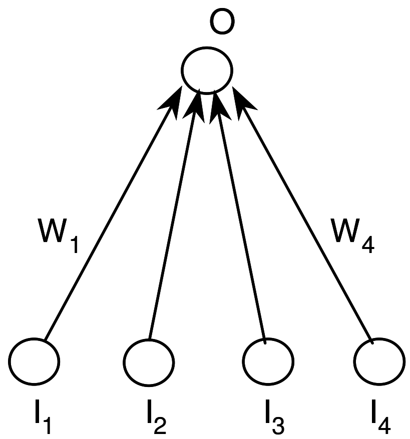
Das Spiel wird viele Male hintereinander gespielt, und jeder Spieler hat das Ziel, einen möglichst hohen Genuss (= Summe aller Rewards) zu erzielen.

Es stellt sich also die Frage, wie sich die zwei Spieler verhalten werden, d.h. zum Beispiel, ob sich im Verlauf der Zeit stabile Verhaltensmuster ergeben.

Um solche Fragestellungen zu untersuchen, wollen wir die beiden Spieler (oder eventuell nur einen davon) durch stochastische Neuronen simulieren.

## 2. Das Salon-Problem mit stochastischen Neuronen

Modellieren Sie die beiden Spieler durch je ein stochastisches Neuron mit 4 Eingängen und einem Ausgang:




$$O = \begin{cases} 1 & \text{mit Wahrsch.keit } p \\ 0 & \text{mit Wahrsch.keit } 1-p \end{cases}$$

$$p = f(W_1 I_1 + W_2 I_2 + W_3 I_3 + W_4 I_4)$$

$$f(x) = \frac{1}{1+\exp(-x)}$$

O = 1 bzw. 0 bedeutet, dass das Neuron die Wahl S (Salon) bzw. O (Oper) trifft. Die Eingänge $I_1$ bis $I_4$ nehmen ebenfalls nur die Werte 0 oder 1 an. Sie beziehen sich auf das Verhalten der beiden Spieler am vorhergehenden Abend:

| Spieler A | Spieler B |
|---|---|
| O | O |
| S | O |
| O | S |
| S | S |

——>

| $I_1$ | $I_2$ | $I_3$ | $I_4$ |
|---|---|---|---|
| 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 1 |

Es ist also immer nur genau ein $I_k = 1$, so dass $p = p_k = 1 / (1 + \exp(-W_k))$. Am nächsten Abend wählt das Neuron dann also mit Wahrscheinlichkeit $p_k$ das Verhalten S und mit Wahrscheinlichkeit $1 - p_k$ das Verhalten O.

Nach jedem Spiel verändert das Neuron sein Gewicht $W_k$ (d.h. dasjenige Gewicht, das zum Input mit $I_k = 1$ gehört) gemäss dem $A_{R-P}$ - Verfahren um $\Delta W_k$:

$$\Delta W_k = \eta\, [r\,(O - p_k) - \alpha\,(1 - r)\,(O - 1 + p_k)]\ .$$

(Die übrigen Gewichte $W_i$ bleiben unverändert!)

Auf diese Weise adaptiert das Neuron seine Taktik an die Strategie des Gegners.

<u>Hinweis</u>: Zu Beginn sind alle Gewichte gleich Null, d.h. alle $p_k$ sind am Anfang gleich 0.5.

## 3. Aufgaben

3a) Implementieren Sie ein entsprechendes Programm, und analysieren Sie das Verhalten der beiden Neuronen (zeitliche Entwicklung der Wahrscheinlichkeiten $p_A(k)$, $p_B(k)$ und der kumulierten Payoffs $G_A$ und $G_B$) für verschiedene Werte von $\eta$ und $\alpha$ (z.B. $\eta = 0.5, 1.0, 2.0$ und $\alpha = 0, 0.05$).

Beobachten und beschreiben Sie die Strategien, die die beiden Neuronen im Verlauf der Zeit entwickeln.

Hinweise:

- Spielen Sie pro Run etwa 2000 bis 5000 Spiele, und beobachten Sie die Werte von $p_A(k)$, $p_B(k)$, $G_A$ und $G_B$ nach jeweils 50 oder 100 Spielen.
- $G_A = \sum r_A \quad , \quad G_B = \sum r_B$ .

3b) Modifizieren Sie das Programm so, dass Sie die Rolle des Spielers B übernehmen können (interaktives Programm "Mensch gegen stochastisches Neuron"), und spielen Sie mit verschiedenen Strategien jeweils 50 oder 100 Spiele gegen das Neuron (z.B. mit $\eta = 2.0$ und $\alpha = 0$).

- Versuchen Sie, dem Neuron eine bestimmte Strategie aufzuzwingen (z.B. abwechslungsweise Salon und Oper zu wählen).
- Welchen Payoff erzielen Sie, wenn Sie "intuitiv" gegen das Neuron spielen?

## 4. Pseudo-Code für ein "Salon"- Programm

```
CHOOSE η AND α;

SET             WA(k) := 0  [k := 1 ... 4];
                WB(k) := 0  [k := 1 ... 4];

INITIALIZE      OA(-1), OB(-1)
                GA := 0;             [* Gewinn Neuron A *]
                GB := 0;             [* Gewinn Neuron B *]

LOOP

        k := 1 + OA(-1) + 2*OB(-1);

        PA := 1 / (1 + exp(-WA(k)));
        PB := 1 / (1 + exp(-WB(k)));

        RNDA, RNDB := Zufallszahlen zwischen 0 und 1;

        IF RNDA <= PA    THEN  OA := 1;
                         ELSE  OA := 0;
```

```
        IF RNDB <= PB    THEN  OB := 1;
                         ELSE  OB := 0;

        IF (OA, OB) = (1, 1)   THEN   (RA, RB) := (0, 0);
        IF (OA, OB) = (1, 0)   THEN   (RA, RB) := (1, 0.2);
        IF (OA, OB) = (0, 1)   THEN   (RA, RB) := (0.2, 1);
        IF (OA, OB) = (0, 0)   THEN   (RA, RB) := (0.2, 0.2);

        GA := GA + RA;
        GB := GB + RB;

        DWA := RA*(OA – PA) – α* (1 – RA)*(OA – (1 – PA));
        WA(k) := WA(k) + η*DWA;

        DWB := RB*(OB – PB) – α* (1 – RB)*(OB – (1 – PB));
        WB(k) := WB(k) + η*DWB;

        OA(-1) := OA;
        OB(-1) := OB;

        PRINT   GA, GB
        PRINT   PA(i) := 1 / (1 + exp(-WA(i)));, i := 1 ... 4
        PRINT   PB(i) := 1 / (1 + exp(-WB(i)));, i := 1 ... 4

        [* PRINT nur nach jeweils 100 Spielen! *]

    LOOPEND
```

Bemerkung:

Beim interaktiven Programm “Mensch gegen stochastisches Neuron“ wird nur ein Neuron simuliert (Neuron A), und die Wahl von OB (=1 oder 0) wird bei jedem Spiel von Hand eingegeben.

Die Ausgabe (PRINT bzw. DISPLAY) beschränkt sich dann auf

RA, RB     und     GA, GB,

erfolgt aber nach jedem der 50 oder 100 Spiele.