**Neuroinformatik – Übungen**

**WS 2019/20**

Beispiel 4, Abgabe: 16.12.2019
(entwickelt von Manfred Klöbl, adaptiert durch GD)

Für diese Übung ist, wie in Übung 2, das Programm Simbrain (<http://www.simbrain.net/>) notwendig. Installieren Sie die Version 3-beta1 von

<http://www.meduniwien.ac.at/user/georg.dorffner/lv/folien>

Dazu ist es nur erforderlich, das entsprechende zip-File zu entpacken. Die Software wird durch Aufrufen der Dateien Simbrain.jar gestartet. Eine Dokumentation ist durch Klicken auf den Knopf „Help“ zugänglich – beachten Sie auch den Online Artikel unter <http://www.brains-minds-media.org/archive/1411>)

Dieses Beispiel baut vollinhaltlich auf Übung 2 auf. Siehe Grundbemerkungen dort.

# Beispiel 4 (30 Punkte)

***4a) Training***

Führen Sie ein vollständiges Training des Netzwerkes wie in Beispiel 2c) durch oder übernehmen Sie das trainierte Netz von dort. Wählen Sie des weiteren 10 zufällige Wörter wie in 2a) aus bzw, übernehmen diese von dort.

***4b) Verteilte Repräsentationen***

Analysieren Sie die verteilten Repräsentationen in den 10 Hidden Units (Muster an Aktivierungen) beispielhaft anhand ihrer Ähnlichkeiten zueinander. Ermitteln Sie dazu die internen Repräsentationen von drei Mal mindestens 3 möglichst unterschiedlichen Wörtern (wie „AUTO“ oder „MAMA“) mit der gleichen Pluralform („Autos“, „Mamas“), sowie von drei Mal mindestens 3 einander phonologisch ähnlichen Wörtern (wie „HAUS“ und „MAUS“) mit verschiedenen Pluralformen („Häuser“ vs. „Mäuse“). Die Werte können dabei von der graphischen Darstellung des Netzes abgelesen werden. Betrachten Sie die entsprechenden Repräsentationen als Vektoren und berechnen in jeder Gruppe von Wörtern ihre durchschnittlichen euklidischen Distanzen zueinander. Welche Ähnlichkeiten sind größer – die von den phonologisch ähnlichen oder den hinsichtlich der Pluralendungen ähnlichen Wörtern?

***4c) Häufigkeiten***

Ermitteln Sie durch Erstellung von Häufigkeitstabellen die relativen Häufigkeiten jeder der sieben deutschen Plural-Klassen. Nehmen Sie nun die zwei häufigsten Klassen und eliminieren zufällig 50% aller Wörter in diesen beiden Klassen und wiederholen das Training nach 4a). Dokumentieren – und, wenn möglich, interpretieren – Sie etwaige Änderungen zur Performanz an den 10 zufällig gewählten Wörtern, im Vergleich zu 2b).

***4d) Default***

Ermitteln Sie den „wahren“ Default des Netzes (lauter 0en als Input). Entspricht dieser einem der häufigsten in 4c) ermittelten Klassen??

# III. Umgang mit Simbrain

|  |  |
| --- | --- |
| Trainingsdaten ansehen | Rechtsklick auf das Label „Backprop“ 🡪 <Edit combined data> 🡪 linkes Icon über der jeweiligen Tabelle („Open .csv data ...“) (links Input, rechts Target) |
| Netz trainierenInputmuster testen | Rechtsklick auf Label „Backprop“ 🡪 <Train Backprop net ...> 🡪 <Run> oder <Batch> 🡪 Start-Button. Vor dem ersten Training zunächst auf <Randomize> klicken.In der Data World eine Zelle des gewünschten Inputmusters anklicken, dann „global step network update algorithm“-Taste (linker Pfeil) im Hauptmenü von Simbrain klicken. |
| TrainingseinstellungenLernkurve anzeigen | im Trainer Fenster: [Properties]Taste „show error plot“ im Trainer Fenster |
| Löschen von Hidden UnitsData World hinzufügen | Unit durch Anklicken aus wählen, Rechts-Klick 🡪 <Delete><Insert> 🡪 <New World> 🡪 <Data World> |
| Daten in Data World laden | in der Data World: <File> 🡪 <Import / Export .csv> 🡪 <Import> |
| Data World in Iteration Mode setzen | in der Data World: <Edit> 🡪 <Iteration mode> |
| Data World an Input koppeln | in der Data World: <Couple> 🡪 <Network X> |