

# Event Segmentation mit Hilfe eines Rekurrenten Neuronalen Netzes mit LSTM

Exposé zur Bachelorarbeit von Steffen Schnürer  
bei Prof. Dr. Martin V. Butz, betreut von Sebastian Otte

Event Segmentation ist die (automatisierte) Erkennung diskreter Ereignisse zur Unterteilung einer kontinuierlichen Aktivität. Ein sehr einfaches Beispiel einer solchen unterteilbaren kontinuierlichen Aktivität ist das Bouncing Ball Szenario. Ein Ball gleitet mit steter Geschwindigkeit in einer quadratischen Ebene. Erreicht er eine Kante, prallt er von ihr ab und gleitet in die entsprechende Richtung weiter. Mit Hilfe eines Rekurrenten Neuronalen Netzes (RNN) mit LSTM (Long short-term memory) soll zunächst die Flugbahn eines solchen Balles erlernt werden.

Die Motivation hinter Event Segmentation ist vielfältig. Ein möglicher, aber auch sehr ambitionierter Anwendungsbereich ist Event Segmentation von menschlichen Aktionen. Aktionen können so unterteilt, klassifiziert und eventuell sogar vorhergesagt werden. So ist zum Beispiel die Aktion des Kaffee Trinkens aus einer Tasse unterteilbar in "Hand zu Tasse führen", "Hand umgreift Henkel", "Tasse wird zum Mund geführt" und "Kaffee wird getrunken", jeweils unterteilt durch Events wie z.B. "Hand erreicht Tasse", vergleichbar mit "Ball erreicht Wand" aus dem Bouncing Ball Szenario.

Ein Neuronales Netz ist ein Netz aus Zellen die intern Signale der vorigen Ebene auswerten und selbst wiederum von Zellen der nächsten Ebene unterschiedlich gewichtet beachtet werden. Die Gewichte werden mit Hilfe von Trainingsdaten, bestehend aus Inputdaten und den zugehörigen Outputdaten, über mehrere Epochen justiert.

In Rekurrenten Neuronalen Netzen verfügt das Netz außerdem noch über Verbindungen zum vorherigen Zeitschritt. So können Probleme gelöst werden, bei denen das Ergebnis von den vorigen Iterationsschritten abhängt, z.B. wie im Bouncing Ball Szenario. Um zu erraten wo der Ball als nächstes sein wird genügt nicht die derzeitige Position, man muss auch einbeziehen wo er zum vorigen Zeitpunkt war. Beim Training eines RNN mit BPTT(Backpropagation through time) werden die Fehler im Outputlayer mittels Gradienten an die vorigen Layer weiter propagiert, damit diese ihre Gewichte entsprechend anpassen können. Muss so ein Gradient durch mehrer Layer, kann sein Wert schnell verschwindet gering werden(Vanishing gradient problem). Eine Lösung hierzu ist LSTM.

Ein LSTM fügt einem RNN Speicherzellen mit Gates für In und Output zu. Diese bewirken, dass ein Wert über mehrere Iterationen hinweg gespeichert werden kann und nur bei bestimmten Umständen geändert und/oder ausgegeben wird. Konkret im Bouncing Ball Szenario könnte man z.B. eine Assoziation von Gates und Events derart finden, dass die Outputgates der Speicherzelle vom RNN geöffnet werden, wenn der Ball sich in Kantennähe befindet, um so das besondere Verhalten dessen und das Ereignis des Abprallens zu simulieren.

Ziel der Arbeit wird es sein zunächst diese Funktionalität zu erreichen und optimieren. Außerdem durch Experimentieren mit verschiedenen Parametern Zusammenhänge zwischen den Speicherzellen im LSTM und den Ereignissen der simulierten Aktivität aufzustellen.

Es besteht die Hoffnung durch Analyse der Speicherzellen Erkenntnisse über die Ereignisse zu bekommen, diese klar unterteilen und zuordnen zu können und eventuell sogar vorherzusagen zu können, da sich ja auch vorraus berechnen lässt unter welchen Umständen im Netz die Speicherzellen ein bestimmtes Verhalten aufweisen.

Zu diesem Thema wurde im vergangenen Semester hier bereits eine Masterarbeit geschrieben, bei der noch vieles offen blieb. Diese Arbeit soll nun hoffentlich für weitere Erkenntnisse sorgen.

#### Zeitplan:

Für die Bearbeitung der Bachelorarbeit sind 4 Monate vorgesehen.

1. Monat: Einarbeitung in die Erstellung von LSTMs, erst 1D-BB-Szenario, dann 2D-BB-Szenario erlernt mit LSTM.
2. Monat: Vorhersagen über Gates und Events anhand Aktivierungen im LSTM. Kontextuierung von Gates und Events im BB-Szenario.
3. Monat: Betrachten weiterer Szenarien: z.B. 3D-BB-Szenario, Bouncing Ball mit Schwerkraft, Discworld Szenario /oder: Dem LSTM-Agent die Möglichkeit zu Handeln zu geben, z.B: immer wenn Event/Kollision vorhergesagt wird, wird die Richtung geändert.
4. Monat: Schreiben der Bachelorarbeit.

#### Vorraussichtliche Quellen:

Skript zur Vorlesung Advanced Neural Networks sowie im Ilias beigefügte Paper.

Masterarbeit von Po-Hsuan Huang sowie dessen angegebenen Quellen.