

Der Token-Pass-Algorithmus, seiner Zeit voraus?

Proseminar - Verarbeitung Gesprochener Sprache

Nguyen Hoang Trung

30.08.2020

Abstract

1989 erschien eine wissenschaftliche Arbeit, die ein Modell und ein Verfahren zu musterbasierten Spracherkennung beschrieb. Dabei basiert die Erkennung auf die simulierte Weitergabe von abstrakten Informationsträgern innerhalb und zwischen Netzwerken. Doch obwohl diese Arbeit einen interessanten und unkonventionellen Ansatz zur Spracherkennung lieferte, dauerte es lange, bis weitere Forschungsarbeiten darauf aufbauten und das Prinzip weiterentwickelten. Diese Proseminararbeit untersucht die Arbeit zum Token-Pass-Algorithmus in dem größeren Zusammenhang der Forschung und möchte sowohl Erklärungsansätze für die Geschehnisse liefern als auch die Arbeit selbst kritisch unter verschiedenen Gesichtspunkten betrachten.

Inhaltsverzeichnis

1 Grundlagen des Token-Pass-Algorithmus	3
1.1 Einführung	3
1.2 Isolierte Worterkennung	3
1.3 Verknüpfte Worterkennung	5
2 Interessante Erkenntnisse	7
2.1 Zusammenhang zu anderen musterbasierten Erkennungsalgorithmen . . .	7
2.2 Zusammenhang zu HMM- und DTW- basierter Erkennung	7
2.3 Abstraktionsebenen	7
3 Zusammenhänge zu anderen Arbeiten ihrer Zeit	8
3.1 Ein Überblick der Arbeiten	8
3.2 Token-Pass im Vergleich	9
4 Die Bedeutung des Token-Pass-Prinzips	10
4.1 Kritische Betrachtung der Arbeit selbst	10
4.2 Kritische Betrachtung der äußeren Umstände	10
4.3 Der Einfluss der Token-Pass-Arbeit	11
5 Fazit	11

1 Grundlagen des Token-Pass-Algorithmus

1.1 Einführung

Um überhaupt die Wirkungsweise des Token-Pass-Algorithmus näher analysieren zu können, muss zuerst klar sein, mit welchem grundlegenden Prinzip die Spracherkennung erfolgt. In der Forschung haben sich dabei zwei Richtungen herauskristallisiert: der Mustervergleich-Ansatz und der statistische Ansatz.

Betrachten wir zuerst die Spracherkennung über Mustervergleich. Jedes Wort in einem Vorrat wird als ein Muster betrachtet. Bei einer gegebenen Eingabe wird diese nun mit allen verfügbaren Mustern in der Datenbank verglichen, dabei wird ermittelt, wie nah die Eingabe und das Muster aneinander liegen. Je kleiner diese „Distanz“ ist, desto ähnlicher sind sich die Eingabe und das Muster. Daher herrscht eine umgekehrte Proportionalität zwischen der Distanz und der Ähnlichkeit.

Beim statistischen Ansatz hingegen geht man davon aus, dass sich der spezifische Klang eines jeden Wortes in einem Vorrat statistisch beschreiben lässt. Die Erkennung einer Eingabe erfolgt also über das Finden des Wortes, dessen statistische Beschreibung am wahrscheinlichsten zur Eingabe passt.

Im Kontext der Zeit hat sich der Ansatz des Mustervergleichs bereits um die 1970er herausgebildet, während der 1980er begann die Forschungsgemeinschaft jedoch vermehrt zur Anwendung des statistischen Ansatzes überzugehen. [5]

Wir werden uns nun im Verlaufe dieses gesamten Abschnittes tiefer mit der Arbeit „Token Passing: a Simple Conceptual Model for Connected Speech Recognition Systems“ aus dem Jahr 1989 beschäftigen [8]. Sie wird der Dreh- und Angelpunkt dieser Proseminararbeit sein und Hauptgegenstand unserer Untersuchungen. Wir betrachten im nächsten Teilabschnitt ein Konzept zur Spracherkennung, welches die Arbeit einführt.

1.2 Isolierte Worterkennung

Die isolierte Worterkennung behandelt folgendes Problem: Wie ähnlich ist eine unbekannte Eingabe einem bestimmten Wort der Sprache? Dabei wird zuerst angenommen, dass die Eingabe in Form einer Sequenz von Vektoren ist. Dann wird ein sogenanntes Wortmodell etabliert, ein Übergangsnetzwerk von Zuständen (dargestellt in Abbildung 1). Zwischen je zwei Zuständen i und j existiert ein Übergang p_{ij} und für jeden Zustand k existiert eine lokale Funktion $d_k(x)$. Das Konzept der Distanzermittlung beim Mustervergleichsansatz wird über das Konzept von Kosten umgesetzt. Jeder Übergang besitzt einen fixen Kostenbetrag und jede lokale Funktion ist eine Kostenfunktion, die einen Betrag abhängig von einem Eingabevektor berechnet.

Einer Sequenz von k Eingabevektoren kann nun eine Sequenz von Zuständen des Wortmodells zugeordnet werden. Jede Sequenz von Zuständen, die mit dem Startzustand q_0 des Wortmodells anfängt und insgesamt eine Anzahl von $k + 1$ Zuständen hat, ist eine mögliche Anordnung des Wortmodells zur Eingabe. Man kann nun eine Distanz zwischen einer Anordnung $q = q_0q_1q_2\dots q_T$ zur Eingabe mit folgender Formel berechnen:

$$S(q) = \sum_{k=1}^T (p_{q_{k-1}q_k} + d_k(k)) \quad (1)$$

Man beachte hierbei, dass die lokale Kostenfunktion $d_k(x)$ als Input die Nummerierung des Eingabevektors erhält und der Wert der Ausgabe von diesem Eingabevektor abhängt.

Bei der Einzelworterkennung auf Basis des Mustervergleichs geht es um die Beantwortung folgender Fragestellung: Was ist die minimale Distanz, die eine mögliche Anordnung des Wortmodells zur Eingabe haben kann? Dazu kann mit folgender Formel rekursiv ermittelt werden, was für einen gegebenen Zustand j des Wortmodells die kleinstmögliche Distanz zur Eingabe ist, wenn die Anordnung des Wortmodells mit dem Startzustand des Wortmodells anfängt und mit dem Zustand j endet:

$$s_j(t) = \underbrace{\min}_i (s_i(t-1) + p_{ij}) + d_j(t) \quad (2)$$

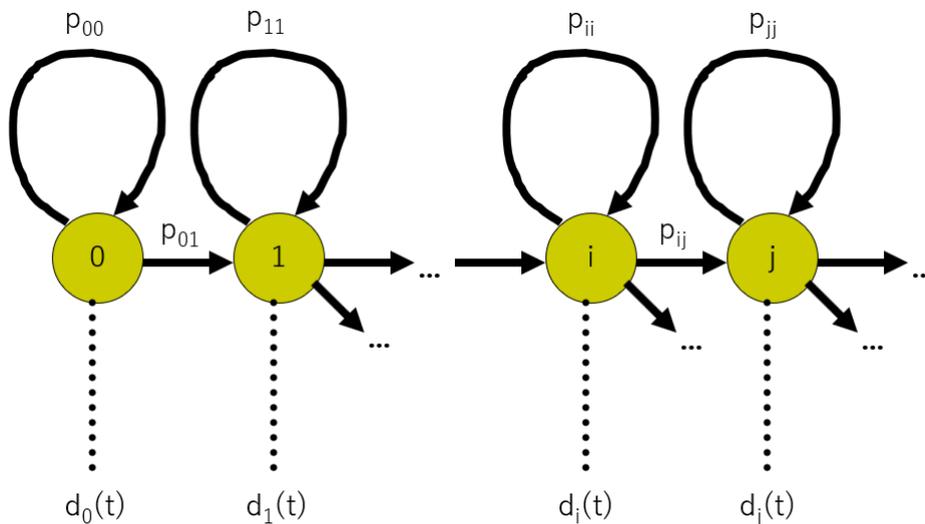


Abbildung 1: Grobe Skizze eines Wortmodells

Anschließend wird das Minimum von allen $s_j(q)$ für jeden möglichen Endzustand j ermittelt. Damit erhält man die absoluten minimalen Zuordnungskosten $S_{min}(q)$ für die gegebene Eingabe.

Allgemein wird ein umgekehrter proportionaler Zusammenhang zwischen den allgemeinen minimalen Zuordnungskosten und der Ähnlichkeit angenommen: Je kleiner $S_{min}(q)$ für eine Eingabe q ist, desto eher passt das Wortmodell (und im übertragenen Sinne das Wort) zur Eingabe. Die Erkennung erfolgt also letztendlich über die Ermittlung des kleinstmöglichen Kostenbetrages zur Zuordnung und damit die minimale Distanz zwischen der Eingabe und dem Wort.

In der Arbeit wird nun ein alternatives Verfahren zur rekursiven Berechnung vorgestellt, welches auf Tokens basiert. Ein Token (zu Deutsch: Marke) ist eine abstrakte Einheit,

die sowohl Informationen speichern kann als auch vervielfältigt und weitertransportiert werden kann. Es wird nun angenommen, dass sich in den Zuständen eines jeden Wortmodells beliebig viele Tokens befinden können und ein Tokenaustausch zwischen Zuständen möglich ist. Nun wird Algorithmus in Abbildung 2 ausgeführt.

- Die Eingabe sei eine Sequenz von T Eingabevektoren.
- Jeder Anfangszustand enthält einen Token mit dem Wert 0.
- Jeder andere Zustand erhält einen Token mit dem Wert ∞ .
- Für $t := 1$ bis $t := T$ tue Folgendes:
 - Für jeden Zustand i tue Folgendes:
 - Reiche eine Kopie des Tokens auf i allen weiteren Zuständen j weiter, die direkt von i erreichbar sind, wobei der Wert der weitergereichten Tokens um $p_{ij} + d_j(\mathbf{t})$ erhöht wird.
 - Lösche nun alle Tokens, die vorher schon auf den Zuständen lagen.
 - Für jeden Zustand i tue Folgendes:
 - Finde den Token mit dem niedrigsten Wert, welcher auf i liegt, und lösche alle übrigen Tokens auf i .
- Durchsuche alle Endzustände und nimm alle Tokens. Ermittle aus diesen den Token mit dem kleinsten Wert. Dieser Wert ist der minimale Übereinstimmungswert.

Abbildung 2: Der Token-Pass-Algorithmus

Am Ende erhält man als Ergebnis einen Wert, der Rückschlüsse auf die Ähnlichkeit zwischen Eingabe und Wort erlaubt.

1.3 Verknüpfte Worterkennung

Bei der verknüpften Worterkennung geht es um die Frage, welche Wortsequenz zu einer unbekanntem Eingabe am besten passt. Die Lösung lässt sich intuitiv bereits in Form von Netzwerken aus Wortmodellen erahnen (siehe Abbildung 3). Man erlaubt den Fluss von Tokens sowohl innerhalb als auch zwischen Wortmodellen und erhält bei Durchführung des Algorithmus' die minimalen Zuordnungskosten.

Das Problem ist nur noch, dass man ohne Modifizierung des Verfahrens nicht direkt erfährt, welche Wortsequenz denn nun diejenige ist, die am besten zur gegebenen Eingabe passt. Um diese zusätzliche Anforderung zu bewerkstelligen, wird der vorhin erläuterte Algorithmus erweitert, und zwar an der Stelle, wo Tokens potenziell zwischen Wortmodellen ausgetauscht werden. In Abbildung 4 ist die notwendige Modifizierung angegeben.

Durch die Weise, wie die Pfadidentifizierer jedes Tokens aktualisiert werden, kann man am Ende nach der Ermittlung des Tokens mit dem niedrigsten Wert durch die Interpretation des WLR dieses Tokens festgestellt werden, welche Sequenz von Wörtern am besten zur gegebenen Eingabe passt. Denn die WLR selbst enthält Referenzen auf weitere Pfadidentifizierer, und diese wiederum zeigen auf weitere WLR, so dass sich am

Ende eine verknüpfte Kette von Objekten ergibt, die Rückschlüsse über die Sequenz geben kann.

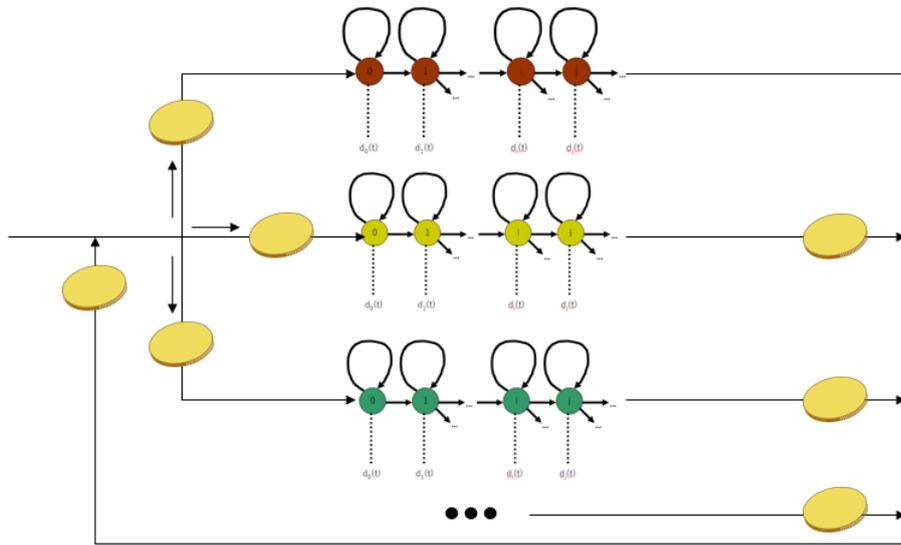


Abbildung 3: Tokenfluss zwischen Wortmodellen

- Für jeden Token, der von einem Wortmodell nach außen zum Zeitpunkt t weiterverbreitet wird, tue Folgendes:
 - Erstelle einen WLR (Word Link Record), der folgendes enthält:
 - den Inhalt des Tokens
 - das sendende Wortmodell
 - die Zeit t
 - Der Pfadidentifizierer des Tokens soll nun auf diesen erstellten WLR zeigen.

Abbildung 4: Die Modifizierung

2 Interessante Erkenntnisse

2.1 Zusammenhang zu anderen musterbasierten Erkennungsalgorithmen

In der Arbeit zum Token-Pass-Algorithmus wird in der Einführung eine weitere wissenschaftliche Arbeit erwähnt, die sich mit dem OP-Algorithmus (OP) und dem LB-Algorithmus (Level Building) auseinandersetzt. Genauer wird in dieser Arbeit gezeigt, dass die beiden Algorithmen, die auf Mustervergleich basieren, unter bestimmten Umständen identisch sind.

Der Zusammenhang wird in der Token-Pass-Algorithmus-Arbeit noch weiter ausgebaut. Es lässt sich zeigen, dass sich sowohl der OP-Algorithmus als auch der LB-Algorithmus mit demselben Algorithmus umsetzen lassen: dem Token-Pass-Algorithmus selbst. Die Unterschiede manifestieren sich in der Topologie der Wortmodelle, die notwendig ist, um die jeweiligen Algorithmen zu simulieren. Man kann also beispielsweise den LB-Algorithmus mit dem Token-Pass-Algorithmus umsetzen, wenn vorher die Verbindungen zwischen den Wortmodellen dementsprechend angepasst wurden.

2.2 Zusammenhang zu HMM- und DTW- basierter Erkennung

In gewisser Hinsicht lässt sich ein erstaunlicher Zusammenhang zu zwei scheinbar grundlegend verschiedenen Verfahren herstellen. Schon die Parallelen zwischen dem vorgestellten Wortmodell und der allgemeinen Struktur eines Hidden-Markov-Modells sind fast schon offensichtlich: So sind die Übergangskosten p_{ij} zwischen zwei Zuständen i und j nur von genau diesen abhängig und nicht von anderen Zuständen, die in einer Anordnung des Wortmodells früher zugeordnet werden. Außerdem ist die lokale Kostenfunktion $d_k(t)$ nur von dem jeweiligen Zustand k abhängig, nicht davon, von welchen Zuständen aus der Zustand k erreicht wurde. Insofern ähnelt die Struktur eines Wortmodells die eines HMMs, da ja dort sowohl die Übergangswahrscheinlichkeiten als auch die Beobachtungswahrscheinlichkeiten bekanntlich genauso wenig von vorherigen Zuständen und Beobachtungen abhängen.

Es besteht auch eine mathematische Verbindung zwischen der Ermittlung der minimalen Zuordnungskosten bei einem Wortmodell und der Erkennung bei einem HMM, dasselbe gilt auch für die Erkennung auf Basis von Dynamic Time Warping. Ohne näher auf die konkrete Theorie hinter der Verbindung eingehen zu wollen, erlauben diese neu gewonnenen Erkenntnisse uns, das Erkennungsverfahren mit DTW als Spezialfall des Erkennungsverfahrens für HMMs zu betrachten.

2.3 Abstraktionsebenen

Die beiden vorherigen Teilabschnitte lieferten zusammengefasst zwei wichtige Erkenntnisse: Zum einen kann ein individuelles Wortmodell sowohl das HMM-Erkennungsprinzip als auch das DTW-Erkennungsprinzip anwenden. Zum anderen können zwei unterschiedliche Algorithmen zur verknüpften Worterkennung, der LB- und der OP-Algorithmus, mithilfe

desselben Verfahrens umgesetzt werden, vorausgesetzt, die Topologie der Wortmodelle wird angemessen umgeformt.

Damit lässt sich nun sagen, dass sich in diesem Fall die verknüpfte Worterkennung in gewisser Weise von der Einzelworterkennung abstrahieren lässt. Letztendlich läuft der Erkennungsprozess von Wortsequenzen darauf hinaus, dass Tokens zwischen Wortmodellen fließen können und dass sich die Geschichte eines Tokens über geregelte Mechanismen vermerkt werden kann. Die Verarbeitung von Tokens innerhalb eines Wortmodells kann also klar von der Verarbeitung von Tokens aus der Makroperspektive getrennt werden. Damit lässt sich die Architektur des Erkennersystems als Ganzes unabhängig von dem zugrundeliegenden Erkennungsprinzip in den einzelnen Wortmodellen entwickeln. Die Wortmodelle selbst können dann als unabhängige Mustererkennungseinheiten betrachten, und somit ergibt sich automatisch eine Schnittstelle zwischen den einzelnen Einheiten und der Steuerung des Systems.

3 Zusammenhänge zu anderen Arbeiten ihrer Zeit

3.1 Ein Überblick der Arbeiten

Wir betrachten nun eine ausgewählte Sammlung von wissenschaftlichen Arbeiten im Bereich der Spracherkennung, die ungefähr im gleichen Zeitraum wie die Arbeit zum Token Passing erschienen sind. Insbesondere dient dies dazu, diese Arbeit in den Kontext der Zeit setzen zu können.

Betrachten wir als Erstes „A Comparative Study of Several Dynamic Time-Warping Algorithms for Connected-Word Recognition“, eine Arbeit aus dem Jahr 1981, untersucht mehrere ausgewählte Algorithmen im Bereich der verknüpften Worterkennung, die auf dem DTW-Prinzip basieren [3]. Der primäre Aspekt dieser Arbeit ist die Untersuchung von Unterschieden in der Leistung und das Treffen von Schlussfolgerungen über die praktische Nutzbarkeit von gewissen Algorithmen.

Nehmen wir nun den Artikel „A Maximum Likelihood Approach to Continuous Speech Recognition“ aus dem Jahr 1983 [1]. Zusammengefasst werden mehrere statistische Modelle zur verknüpften Worterkennung auf Effizienz und Fehlerrate geprüft. Außerdem werden zwei Decodier-Verfahren vorgestellt, die für bestimmte Sprachen und damit für verschiedene Anwendungen besser funktionieren. Allgemein ist die Arbeit sehr auf die Bestimmung optimaler Parameter für gegebene Trainingsdaten fokussiert.

„A Distributed, Developmental Model of Word Recognition and Naming“ ist eine wissenschaftliche Arbeit, erschienen in der Zeitschrift „Psychological Review“ im Jahr 1989 [6]. Es wird dabei ein Modell zur visuellen Erkennung von Wörtern und deren Aussprache präsentiert, wobei die Funktionsweise dieses Modells auf eine Interaktion von orthografischen, phonologischen als auch versteckten Einheiten basiert und das Training ein wichtiger Aspekt des Modells ist. Doch der Kern dieses Artikels beschäftigt sich eigentlich mit der Frage, inwiefern solche und ähnliche Modelle genutzt werden können, um das Verhalten von Menschen zu analysieren, wenn sie die Aussprache von Wörtern erkennen sollen. Insbesondere werden Zusammenhänge zwischen dem frühen Verhalten

des Modells beim Training und zu Kindern in ihrer Lernphase als auch zu Menschen mit kognitiven Einschränkungen untersucht.

„Modular Construction of Time-Delay Neural Networks“ aus dem Jahr 1989 greift Probleme in der Spracherkennung auf, die sich bei der Entwicklung von Netzwerkmodellen ergeben haben [7]. Konkret geht es um die Problematik der Zeit und der Skalierung: Bezogen auf die Zeit war es bei der Erkennung von Wortsequenzen mit dem klassischen Ansatz problematisch, wenn einzelne Phoneme bei uneindeutiger Akustik ermittelt werden müssen. Bezogen auf die Skalierung musste man sich bei der Ausweitung eines Modell (z.B. mehr Daten, größere Struktur) fragen, ob dies aus praktischer Sicht sinnvoll war und ob der Nutzen die Kosten rechtfertigt. Die Arbeit schlägt einen kombinierten Lösungsansatz vor: Zur Lösung des Zeitproblems werden Time Delay Neural Networks genutzt; zur Lösung des Skalierungsproblems wird das Erkennersystem modular aufgebaut und besteht aus mehreren Subnetzen, so dass Erweiterungen aus praktischer Sicht möglich sind.

„Hidden Markov Models for Speech Recognition“ ist ein wissenschaftlicher Artikel aus dem Jahr 1991 über Hidden-Markov-Modelle in der Spracherkennung und über Fragestellungen, die damals noch im Raum standen und gelöst werden müssten [2]. Im Kontext der Zeit griff dieser Artikel den Trend zu dieser Zeit auf, dass vermehrt HMMs zur Spracherkennung genutzt wurden und somit ein größeres Interesse bestand. Über den Inhalt des Artikels lässt sich sagen, dass es das grundlegende Erkennungsprinzip und dessen Stärken in der Spracherkennung erläutert: Das Hidden-Markov-Modell ist beständiges, vielseitig einsetzbares und flexibles Modell, welches verglichen zu anderen Modellen eine sehr hohe Leistungsfähigkeit zeigt.

1994 erschien „Shortlist: a connectionist model of continuous speech recognition“ in der Zeitschrift „Cognition“ [4]. Der Artikel baut auf eine andere wissenschaftliche Arbeit auf, die ein neuartiges Modell zur verknüpften Worterkennung vorgestellt hat. Dieses Modell hieß TRACE und hatte das Ziel, Aspekte in die Erkennung miteinzubringen, die sich bisher mit rekursiven Netzwerkmodellen nicht gut umsetzen ließen. Genauer war eine Korrektur der Ausgabe durch spätere Analyse des Kontexts für solche Modelle nicht möglich, mit TRACE jedoch gab es diese Möglichkeit. Doch die Architektur selbst war wenig elegant, kompliziert und hatte daher ungewünschte Konsequenzen auf die Leistung in Experimenten. Diese Arbeit selbst präsentiert wiederum ein neues Modell, welches zum Ziel hat, die Vorteile von TRACE und dem klassischen rekursiven Netzwerkmodell zu kombinieren, aber gleichzeitig einen praktischen, verständlicheren Aufbau aufwies.

3.2 Token-Pass im Vergleich

Wir können einige grundlegende Gemeinsamkeiten zwischen den präsentierten Arbeiten im vorherigen Abschnitt feststellen: Für gewöhnlich ist die Leistungsfähigkeit des jeweiligen Modells einer der wichtigsten Aspekte der Arbeit, was in einem Forschungsbereich mit hoher praktischer Relevanz natürlich einleuchtet. Das betrifft nicht nur die Analyse von Daten aus Experimenten, sondern auf Fähigkeiten, die dieses Modell von anderen Modellen abhebt. Die Nutzbarkeit ist ein wiederkehrendes Motiv in all diesen Arbeiten.

Die Arbeit zum Token-Pass-Algorithmus ist im Vergleich zu all den Arbeiten, die im vorherigen Abschnitt vorgestellt wurden, insofern speziell, als dass sie auf etwas

Anderes hinausläuft: Es wird ein neuartiges Verfahren vorgestellt, was nicht zwangsläufig von dem Verlangen nach Performanzsteigerung motiviert ist, sondern von der Suche nach einer neuen Perspektive, die einem das Finden von interessanten Zusammenhängen ermöglicht. Andererseits ist das Konzept von abstrakten Informationsträgern innerhalb eines Netzwerks

4 Die Bedeutung des Token-Pass-Prinzips

4.1 Kritische Betrachtung der Arbeit selbst

Wir setzen uns nun mit der Arbeit zum Token-Pass-Algorithmus kritisch auseinander, und zwar insofern, als dass wir mögliche Gründe für die anfängliche Bedeutungslosigkeit in der Forschung suchen.

Naheliegender wäre zunächst, dass Methodik der Tokenvervielfältigung und -verbreitung in einem Netzwerk vielleicht zu verkompliziert wirkt, um einen wahrlich praktischen Nutzen zu haben. Der Tokenverkehr selbst wird zur Laufzeit des Erkenners praktisch unübersichtlich und schwer nachvollziehbar, da von jedem Zustand aus Tokens kopiert und weitergereicht werden. Natürlich muss man hier anmerken, dass die Korrektheit des Verfahrens trotzdem intuitiv einleuchten kann. Dann könnte man einwerfen, dass durch die offensichtliche Ähnlichkeit des Wortmodells zum klassischen Hidden-Markov-Modell ersteres obsolet erscheinen lässt.

Doch möglicherweise sind die äußeren Umstände viel gravierender und hatten tatsächlich einen größeren Einfluss auf die Rezeption der Arbeit.

4.2 Kritische Betrachtung der äußeren Umstände

In der Geschichte der Spracherkennungswissenschaft befand sich die Forschung zur Zeit der Veröffentlichung in einem Übergang zur Entwicklung des statistischen Ansatzes als Alternative zum Mustervergleich. Die Erkenntnisse einer wissenschaftlichen Arbeit, die auf einen veralteten Ansatz fokussiert, könnten so schnell übersehen werden, auch wenn die Erkenntnisse für sich genommen eigentlich Bedeutung hatten.

Die präsentierten Arbeiten aus dem Zeitraum der Token-Pass-Arbeit, die in dieser Arbeit präsentiert wurden, lassen zumindestens erahnen, wie sich der Blick auf die Spracherkennung in der Forschungsgemeinde gewandelt hat. Hidden-Markov-Modelle gewannen in dieser Zeit zunehmend an Popularität und spielen bis heute noch eine relevante Rolle. Hauptsächlich waren zum einen die Paradigma, auf die die Theorie der HMMs basierte (Markov-Eigenschaft), vielversprechend und durchaus plausibel anwendbar auf das Erkennungsproblem. Zum anderen suchte die Leistungsfähigkeit und Genauigkeit, die mit solchen Modellen erreicht werden konnte, ihresgleichen und setzte einen Maßstab für kommende Ansätze in der Spracherkennung. Daher wurde es als geeignetes Framework in den verschiedensten Arbeiten beschrieben.

Man könnte vermuten, dass der Ansatz, der in der Token-Pass-Arbeit erarbeitet wurde, von den Forschern als der Verfolgung nicht würdig eingestuft wurde. Könnte es gar sein,

dass dies einfach Ausdruck der schwindenden Bedeutung des Mustervergleichsansatzes war? Auch wenn die Essenz der Token-Pass-Arbeit nicht zwangsläufig auf der Idee des Mustervergleichs basiert, sondern vielmehr auf dem neuartigen Werkzeug des Token-austausches, könnte es deswegen dazu gekommen sein, dass Wissenschaftler der Arbeit keine Aufmerksamkeit widmeten und lieber vielversprechendere Denkanstöße verfolgten, die in der Zeit populär geworden waren. Und auch wenn interessante Zusammenhänge und Chancen in dieser Arbeit aufgedeckt wurden, hätte man diese zur damaligen Zeit als irrelevant für den Werdegang der Forschung betrachten können, da zu der Zeit der OP-, der LB-Algorithmus, die DTW-Technologie und der Mustervergleichsansatz keinen Anklang mehr gefunden hatten.

Letztendlich erhebt diese Seminararbeit nicht den Anspruch auf das Festsetzen des plausibelsten Erklärungsansatzes und möchte Raum für Diskussionen schaffen.

4.3 Der Einfluss der Token-Pass-Arbeit

Sucht man „Token passing: a simple conceptual model for connected speech recognition systems“ in Google Scholar, so lässt sich schnell herausfinden, dass die Arbeit zum Zeitpunkt des 30.08.2020 von 492 wissenschaftlichen Artikeln zitiert wird. Im Zeitraum 1989 - 1999 wurde die Arbeit von lediglich 51 Arbeiten zitiert. Man kann also erahnen, dass die Bedeutung dieser Arbeit erst mit fortschreitender Zeit wahrgenommen wurde. Heutzutage spielt der Token-Pass-Algorithmus insofern eine Rolle, als dass damit den Forschern ein neues Werkzeug zur Verfügung steht, welches für eigene Zwecke verwendet und weiterentwickelt werden kann.

5 Fazit

Die Arbeit zum Token-Pass-Algorithmus hat ein neuartiges Konzept zur Erkennung von Wörtern und Wortsequenzen gefunden: Die Entwicklung eines Netzwerkmodells zur Spracherkennung war keine Neuheit, doch die Erkennung über einen systematisch geregelten Verkehr von abstrakten Informationsträgern war ein durchaus innovativer Ansatz, der weitaus früher hätte aufgegriffen und weiterentwickelt werden können. In dieser Arbeit wurde untersucht, inwiefern äußere Einflüsse als auch die Natur der Arbeit selbst dazu geführt haben könnten, dass diese Arbeit erst im Laufe der Zeit an Relevanz gewann. Heutzutage ist das Token-Pass-Prinzip ein abstraktes Werkzeug von vielen, die eventuell einen wichtigen Beitrag zur Lösung neuer Herausforderungen in der Forschung leisten könnten.

Literatur

- [1] Lalit R Bahl, Frederick Jelinek, and Robert L Mercer. A maximum likelihood approach to continuous speech recognition. *IEEE transactions on pattern analysis and machine intelligence*, (2):179–190, 1983.
- [2] Biing Hwang Juang and Laurence R Rabiner. Hidden markov models for speech recognition. *Technometrics*, 33(3):251–272, 1991.
- [3] Cory S Myers and Lawrence R Rabiner. A comparative study of several dynamic time-warping algorithms for connected-word recognition. *Bell System Technical Journal*, 60(7):1389–1409, 1981.
- [4] Dennis Norris. Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3):189–234, 1994.
- [5] Beat Pfister and Tobias Kaufmann. *Sprachverarbeitung*. Springer, 2008.
- [6] Mark S Seidenberg and James L McClelland. A distributed, developmental model of word recognition and naming. *Psychological review*, 96(4):523, 1989.
- [7] Alex Waibel. Modular construction of time-delay neural networks for speech recognition. *Neural computation*, 1(1):39–46, 1989.
- [8] Stephen John Young, NH Russell, and JHS Thornton. *Token passing: a simple conceptual model for connected speech recognition systems*. Cambridge University Engineering Department Cambridge, 1989.