

Die empirische Grammatik von Marktgesprächen

Eine neuro-symbolische Rekonstruktion

Paul Koop

1994–2026

Zusammenfassung

Dieser Beitrag rekonstruiert die empirische Grammatik von acht Verkaufsgesprächen, die 1994 auf dem Aachener Marktplatz aufgenommen wurden, mittels der Algorithmisch Rekursiven Sequenzanalyse (ARS). Die Analyse führt von den Roh-Transkripten über Terminalzeichenketten, Übergangszählung und Wahrscheinlichkeitsinduktion zu einer probabilistischen kontextfreien Grammatik (PCFG) mit empirisch optimierten Übergangswahrscheinlichkeiten. Ich biete eine interpretative Analyse der gelernten Wahrscheinlichkeiten und unterscheide zwischen *konstitutiven Regeln* (structurellen Beschränkungen, die Wohlgeformtheit definieren) und *statistischen Regularitäten* (empirischen Häufigkeiten, die kontingente Muster widerspiegeln). Die Grammatik offenbart sowohl ritualisierte Sequenzen (Begrüßungen, Verabschiedungen) als auch strategische Optionen (Upselling-Schleifen, Neustarts). Der Beitrag schließt mit einer Diskussion, wie die Grammatik in neuro-symbolischen Frameworks implementiert werden kann, um qualitative Interpretation und computergestützte Ausführung zu verbinden.

Inhaltsverzeichnis

1	Einleitung: Von Rohdaten zur formalen Grammatik	3
2	Daten und Interpretation	3
2.1	Die acht Transkripte	3
2.2	Zuordnung der Terminalzeichen	4
3	Die optimierte Grammatik	5
3.1	Induktionsmethode	5
3.2	Die optimierte probabilistische Grammatik	6
3.3	Grafische Darstellung	6
4	Interpretative Analyse der Wahrscheinlichkeiten	8
4.1	Konstitutive Regeln (strukturelle Notwendigkeiten)	8
4.2	Statistische Regularitäten (empirische Kontingenzen)	8
4.3	Die Upselling-Schleife	9
4.4	Die Neustart-Option	9
4.5	Die Ankunft eines neuen Kunden	9
5	Konstitutive Regeln vs. statistische Regularitäten	10
5.1	Die Unterscheidung	10
5.2	Von der Statistik zur Konstitutivität	10
5.3	Die methodologische Bedeutung	11
6	Hin zur neuro-symbolischen Implementierung	11
6.1	Die Dual-Dynamics-Architektur	11
6.2	Erklärbarkeit durch Beweisbäume	12
6.3	Implementierungsoptionen	12
7	Fazit	13

1 Einleitung: Von Rohdaten zur formalen Grammatik

Die Algorithmisch Rekursive Sequenzanalyse (ARS) beruht auf einer einfachen, aber wirkmächtigen Idee: Soziale Interaktionen hinterlassen physikalische Spuren (Tonaufnahmen), die transkribiert, interpretiert und in formale Grammatiken überführt werden können. Dieser Beitrag dokumentiert die gesamte Pipeline von den Roh-Transkripten bis zur empirisch optimierten Grammatik, basierend auf acht Verkaufsgesprächen, die im Juni/Juli 1994 auf dem Aachener Marktplatz aufgenommen wurden.

Der Beitrag ist dreifach:

1. **Empirisch:** Ich präsentiere die vollständige optimierte Grammatik, induziert aus den acht Transkripten, mit allen Übergangswahrscheinlichkeiten.
2. **Interpretativ:** Ich analysiere die gelernten Wahrscheinlichkeiten und unterscheide zwischen konstitutiven Regeln (strukturellen Notwendigkeiten) und statistischen Regularitäten (empirischen Kontingenzen).
3. **Methodologisch:** Ich diskutiere, wie die Grammatik in neuro-symbolischen Frameworks implementiert werden kann, um qualitative Interpretation und computergestützte Ausführung zu verbinden.

2 Daten und Interpretation

2.1 Die acht Transkripte

Das empirische Material umfasst acht Transkripte von Verkaufsgesprächen, die im Juni/Juli 1994 auf dem Aachener Marktplatz aufgenommen wurden. Die Transkripte variieren in Länge, Vollständigkeit und Gesprächstyp:

Tabelle 1: Korpus der Marktgespräche

Text	Datum	Ort	Beteiligte
T1	28.06.1994	Metzgerei	Verkäuferin (w), Kunde
T2	28.06.1994	Kirschenstand	Verkäufer (m), K1, K2
T3	28.06.1994	Fischstand	Verkäufer (m), Kunde
T4	28.06.1994	Gemüsestand	Verkäufer (m), Kunde
T5	26.06.1994	Gemüsestand	Verkäufer (m), K1, K2
T6	28.06.1994	Käseverkauf	Verkäufer (m), K1
T7	28.06.1994	Bonbonstand	Verkäufer (m), Kunde
T8	09.07.1994	Bäckerei	Verkäuferin (w), Kunde

2.2 Zuordnung der Terminalzeichen

Jede Äußerung wurde einem Terminalzeichen aus einem vordefinierten Kategoriensystem zugeordnet, das durch sequenzielle Mikroanalyse nach der Methode der objektiven Hermeneutik entwickelt wurde (Oevermann et al., 1979). Die zwölf Terminalzeichen sind:

Tabelle 2: Terminalzeichen und ihre Bedeutungen

Symbol	Bedeutung
KBG	Kunden-Gruß
VBG	Verkäufer-Gruß
KBBd	Kunden-Bedarf (konkret)
VBBd	Verkäufer-Nachfrage
KBA	Kunden-Antwort
VBA	Verkäufer-Reaktion
KAE	Kunden-Erkundigung
VAE	Verkäufer-Auskunft
KAA	Kunden-Abschluss
VAA	Verkäufer-Abschluss
KAV	Kunden-Verabschiedung
VAV	Verkäufer-Verabschiedung

Die resultierenden Terminalzeichenketten für die acht Transkripte sind:

T1: KBG, VBG, KBBd, VBBd, KBA, VBA, KBBd, VBBd, KBA, VAA, KAA, VAV, KAV

T2: VBG, KBBd, VBBd, VAA, KAA, VBG, KBBd, VAA, KAA
 T3: KBBd, VBBd, VAA, KAA
 T4: KBBd, VBBd, KBA, VBA, KBBd, VBA, KAE, VAE, KAA, VAV, KAV
 T5: KAV, KBBd, VBBd, KBBd, VAA, KAV
 T6: KBG, VBG, KBBd, VBBd, KAA
 T7: KBBd, VBBd, KBA, VAA, KAA
 T8: KBG, VBBd, KBBd, VBA, VAA, KAA, VAV, KAV

3 Die optimierte Grammatik

3.1 Induktionsmethode

Die Grammatik wurde durch Zählung der Übergänge über alle acht Transkripte induziert und zu Wahrscheinlichkeiten normalisiert:

$$P(\sigma_j \mid \sigma_i) = \frac{\text{Zahl}(\sigma_i \rightarrow \sigma_j)}{\sum_k \text{Zahl}(\sigma_i \rightarrow \sigma_k)}$$

Die resultierenden Wahrscheinlichkeiten wurden iterativ verfeinert, indem künstliche Ketten generiert und deren Häufigkeitsverteilungen mit den empirischen Daten verglichen wurden. Die finale Grammatik erreichte eine Korrelation von $r = 0,925$ zwischen empirischen und generierten Häufigkeiten.

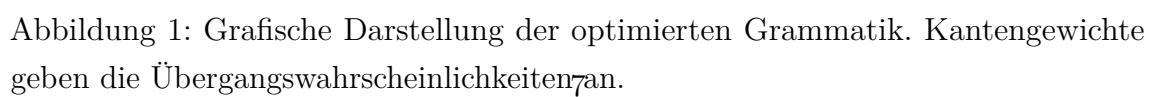
3.2 Die optimierte probabilistische Grammatik

Tabelle 3: Optimierte Übergangswahrscheinlichkeiten

Startsymbol	Folgesymbole mit Wahrscheinlichkeiten
KBG	VBG (0,667), VBBd (0,333)
VBG	KBBd (1,0)
KBBd	VBBd (0,667), VAA (0,167), VBA (0,167)
VBBd	KBA (0,444), VAA (0,222), KBBd (0,222), KAA (0,111)
KBA	VBA (0,5), VAA (0,5)
VBA	KBBd (0,5), KAE (0,25), VAA (0,25)
VAA	KAA (0,857), KAV (0,143)
KAA	VAV (0,75), VBG (0,25)
VAV	KAV (1,0)
KAE	VAE (1,0)
VAE	KAA (1,0)
KAV	VAV (0,5), KBBd (0,5)

3.3 Grafische Darstellung

Abbildung 1 visualisiert die Grammatik als gerichteten Graphen mit Wahrscheinlichkeiten.



4 Interpretative Analyse der Wahrscheinlichkeiten

4.1 Konstitutive Regeln (strukturelle Notwendigkeiten)

Einige Übergänge haben die Wahrscheinlichkeit 1,0, was darauf hinweist, dass sie nicht bloß statistische Regularitäten, sondern *konstitutive Regeln* des Interaktionsformats sind:

- **VBG $\rightarrow$ KBBd (1,0)**: Ein Verkäufer-Gruß muss von einer Kunden-Bedarfsäußerung gefolgt sein. Dies ist eine strukturelle Eigenschaft von Verkaufsgesprächen. Ohne diesen Übergang würde dem Gespräch sein transaktionaler Zweck fehlen.
- **KAE $\rightarrow$ VAE (1,0)**: Eine Kunden-Erkundigung muss von einer Verkäufer-Auskunft beantwortet werden. Dies spiegelt die normative Erwartung von Reziprozität und Kooperation wider.
- **VAE $\rightarrow$ KAA (1,0)**: Auf Verkäufer-Auskunft folgt immer ein Kunden-Abschluss. Dies deutet darauf hin, dass der Informationsaustausch in Marktgesprächen eng mit der Transaktionsbeendigung gekoppelt ist.
- **VAV $\rightarrow$ KAV (1,0)**: Verkäufer-Verabschiedungen werden immer durch Kunden-Verabschiedungen erwidert. Dies ist eine ritualisierte Abschlussequenz, die das Ende der Interaktion markiert.

4.2 Statistische Regularitäten (empirische Kontingenzen)

Andere Übergänge haben Wahrscheinlichkeiten zwischen 0 und 1, die empirische Häufigkeiten widerspiegeln, die je nach Kontext variieren können:

- **KBG $\rightarrow$ VBG (0,667) vs. KBG $\rightarrow$ VBBd (0,333)**: In zwei Dritteln der Fälle wird ein Kunden-Gruß erwidert; in einem Drittel antwortet der Verkäufer direkt mit einer Nachfrage (Überspringen des reziproken Grußes).
- **KBBd $\rightarrow$ VBBd (0,667) vs. KBBd $\rightarrow$ VAA (0,167) vs. KBBd $\rightarrow$ VBA (0,167)**: Die meisten Kunden-Bedarfe werden von Verkäufer-Nachfragen gefolgt, manchmal aber auch direkt von Abschluss (sofortiger Kauf) oder Reaktion (beratende Antwort).
- **VBBd $\rightarrow$ KBA (0,444) vs. VBBd $\rightarrow$ VAA (0,222) vs. VBBd $\rightarrow$ KBBd (0,222) vs. VBBd $\rightarrow$ KAA (0,111)**: Verkäufer- Nachfragen haben die vielfältigsten Folgen – Antworten, Abschlüsse, Bedarfsschleifen (Upselling) oder Kunden-Abschlüsse (vorzeitiger Ausstieg).

- **KBA $\rightarrow$ VBA (0,5) vs. KBA $\rightarrow$ VAA (0,5)**: Kunden-Antworten führen mit gleicher Wahrscheinlichkeit zu Verkäufer-Reaktionen oder direkten Abschlüssen. Dies deutet auf einen strategischen Wahlpunkt hin.
- **VAA $\rightarrow$ KAA (0,857) vs. VAA $\rightarrow$ KAV (0,143)**: Die meisten Verkäufer-Abschlüsse werden von Kunden-Abschlüssen gefolgt; selten direkt von Kunden-Verabschiedung (abgekürzter Abschluss).
- **KAA $\rightarrow$ VAV (0,75) vs. KAA $\rightarrow$ VBG (0,25)**: Drei Viertel der Kunden-Abschlüsse führen zu Verkäufer-Verabschiedungen; ein Viertel führt zu einem Neustart des Gesprächs (neue Begrüßung). Die Neustart-Option tritt auf, wenn ein neuer Kunde eintrifft oder derselbe Kunde einen zusätzlichen Kauf tätigt.

4.3 Die Upselling-Schleife

Der Übergang **VBA $\rightarrow$ KBBd (0,5)** verdient besondere Aufmerksamkeit. Diese Schleife – von Verkäufer-Reaktion zurück zu Kunden-Bedarf – ist die grammatische Repräsentation des *Upsellings*. Wenn ein Verkäufer "Sonst noch etwas?" (VBA) sagt, antwortet der Kunde oft mit einem zusätzlichen Bedarf (KBBd).

Entscheidend ist, dass diese Schleife nur in der Hälfte der Fälle auftritt (0,5). Die andere Hälfte führt zu Abschluss (VAA, 0,25) oder Erkundigung (KAE, 0,25). Dies deutet darauf hin, dass Upselling weder obligatorisch noch selten ist – es ist eine strategische Option, die Verkäufer einsetzen können, und Kunden können sie annehmen oder abweisen.

4.4 Die Neustart-Option

Der Übergang **KAA $\rightarrow$ VBG (0,25)** ist besonders interessant. Ein Kunden-Abschluss (KAA) wird normalerweise von einer Verabschiedung (VAV, 0,75) gefolgt, aber in einem Viertel der Fälle folgt eine Verkäufer-Begrüßung (VBG). Dies zeigt, dass Gespräche nach dem Abschluss neu gestartet werden können – zum Beispiel, wenn ein neuer Kunde eintrifft (wie in T2 und T5) oder wenn derselbe Kunde einen zusätzlichen Kauf tätigt.

4.5 Die Ankunft eines neuen Kunden

Der Übergang **KAV $\rightarrow$ KBBd (0,5)** zeigt, dass eine Kunden-Verabschiedung "zurückgenommen" werden kann, wenn ein neuer Kunde eintrifft. In T5 folgt auf eine

Verabschiedung (KAV) sofort ein neuer Kunden-Bedarf (KBBd). Dies ist der einzige Übergang, der die ansonsten strikte Phasenabfolge durchbricht.

5 Konstitutive Regeln vs. statistische Regularitäten

5.1 Die Unterscheidung

Das ARS-Framework hält eine strikte Unterscheidung zwischen zwei Beschreibungsebenen ein:

1. **Konstitutive Regeln:** Strukturelle Beschränkungen, die definieren, was als wohlgeformte Sequenz gilt. Diese sind binär (eine Sequenz entspricht entweder den Regeln oder nicht). Im Korpus sind Übergänge mit Wahrscheinlichkeit 1,0 Kandidaten für konstitutive Regeln.
2. **Statistische Regularitäten:** Empirische Häufigkeiten von Übergängen. Diese sind probabilistisch und können aktualisiert werden, wenn neue Daten eintreffen. Sie spiegeln wider, was *tatsächlich* vorkommt, nicht was *notwendigerweise* vorkommen muss.

Tabelle 4: Aus dem Korpus identifizierte konstitutive Regeln

Regel	Wahrscheinlichkeit	Interpretation
VBG $\rightarrow$ KBBd	1,0	Verkäufer-Gruß muss von Kunden-Bedarf gefolgt sein
KAE $\rightarrow$ VAE	1,0	Kunden-Erkundigung muss beantwortet werden
VAE $\rightarrow$ KAA	1,0	Information muss zum Abschluss führen
VAV $\rightarrow$ KAV	1,0	Verabschiedungen sind reziprok

5.2 Von der Statistik zur Konstitutivität

Eine statistische Regularität kann durch methodologische Entscheidung zu einer konstitutiven Regel werden. Der Übergang VBG $\rightarrow$ KBBd (1,0) variierte im Korpus nie. Wir könnten ihn als konstitutive Regel behandeln: "In Verkaufsgesprächen muss auf einen Verkäufer-Gruß eine Kunden-Bedarfsäußerung folgen." Dies verwandelt eine empirische Beobachtung in eine normative Beschränkung.

Diese Entscheidung ist jedoch nicht automatisch. Sie erfordert methodologische Reflexion: Ist dies wirklich ein notwendiges Merkmal des Interaktionsformats, oder

könnte es in anderen Kontexten variieren? Der ARS-Ansatz behandelt Übergänge zunächst als statistisch, bis Gegenbeispiele eine Revision der strukturellen Grammatik erzwingen.

5.3 Die methodologische Bedeutung

Diese Flexibilität ist entscheidend für XAI und neuro-symbolische Integration. Sie erlaubt dem Analysten:

1. Mit rein statistischem Lernen zu beginnen (keine vorherigen Beschränkungen).
2. Übergänge zu identifizieren, die in den Daten nie variieren.
3. Sie nach methodologischer Reflexion zu konstitutiven Regeln zu erheben.
4. Das resultierende hybride System für Vorhersage und Erklärung zu nutzen.

Das System bewegt sich so von rein empirischer Mustererkennung (System 1) zu regelbasierter Erklärung (System 2) – ein Wechsel, der Kahnemans Unterscheidung zwischen schnellem und langsamem Denken widerspiegelt (Kahneman, 2011).

6 Hin zur neuro-symbolischen Implementierung

6.1 Die Dual-Dynamics-Architektur

Die ARS-Grammatik legt nahe eine Dual-Dynamics-Architektur für die neuro-symbolische Implementierung:

1. **Symbolische Komponente** (System 2): Die Grammatikregeln, einschließlich sowohl konstitutiver Regeln als auch statistischer Wahrscheinlichkeiten. Diese Komponente ist inspizierbar, falsifizierbar und erklärbar.
2. **Neuronale Komponente** (System 1): Ein neuronales Netz, das Übergangswahrscheinlichkeiten aus Daten lernt und nächste Symbole vorhersagt. Diese Komponente ist schnell, musterbasiert und skalierbar.
3. **Hybride Integration**: Die neuronale Komponente liefert schnelle Vorhersagen; die symbolische Komponente validiert sie gegen konstitutive Regeln und liefert Erklärungen.

[caption=Pseudocode für neuro-symbolische ARS] Pseudocode: ARS 5.0 Dual-Dynamics-Architektur

```
class ARSNeuroSymbolicSystem:
```

```
def init(self): SymbolischeKomponente
    self.grammar = ARSGrammar(PCFGmitWahrscheinlichkeiten)
    self.counts = zero_matrix(12, 12)
    Übergangszähler

    NeuronaleKomponente
    self.neural_network = NeuralNetwork(input_dim = 12, hidden = 64, output_dim = 12)

    KonstitutiveRegeln(harteBeschränkungen)
    self.constitutive_rules = (KBG, VBG) : True, (VBG, KBG) : False

    def update_symbolic(self, from_sym, to_sym) :
        SSchnelles, zählbasiertesUpdate(System2)
        self.counts[from_sym][to_sym] += 1
        row_sum = sum(self.counts[from_sym])
        self.grammar.probs[from_sym] = self.counts[from_sym] / row_sum

    def update_neural(self, from_sym, to_sym) :
        Langsames, gradientenbasiertesUpdate(System1)
        loss = cross_entropy(self.neural_network(from_sym), to_sym)
        loss.backward()
        optimizer.step()

    def predict_next(self, from_sym) :
        HybrideVorhersage
        neural_probs = self.neural_network(from_sym)
        symbolic_probs = self.grammar.probs[from_sym]
        return 0.5 * neural_probs + 0.5 * symbolic_probs

    def is_valid(self, from_sym, to_sym) :
        PrüfungkonstitutiverRegeln
        return self.constitutive_rules.get((from_sym, to_sym), False)
```

6.2 Erklärbarkeit durch Beweisbäume

Die symbolische Komponente ermöglicht Erklärbarkeit durch Beweisbäume. Für jede wohlgeformte Sequenz kann die Grammatik eine Ableitung produzieren:

```
[caption=Beweisbaum für eine wohlgeformte Sequenz] well_formed([KBG, VBG, KBG], start(KBG))
0.667 1.0 0.0 = 0.0
```

Dieser Beweisbaum macht das Schließen transparent. Jeder Schritt wird durch eine Regel gerechtfertigt, und jede Regel hat eine explizite Wahrscheinlichkeit.

6.3 Implementierungsoptionen

Die ARS-Grammatik kann in verschiedenen neuro-symbolischen Frameworks implementiert werden:

Tabelle 5: Implementierungsoptionen für ARS 5.0

Framework	Sprache	Am besten für
DeepProbLog	Prolog/Python	Forschung, Erklärbarkeit
PyTorch	Python	Flexibilität, Prototyping
Flux.jl	Julia	Wissenschaftliches Rechnen, Performance
Candle	Rust	Produktion, Edge-Computing

Jedes Framework hat seine eigenen Stärken, aber alle können dieselbe Dual-Dynamics-Architektur implementieren.

7 Fazit

Dieser Beitrag hat die empirische Grammatik von acht Marktgesprächen rekonstruiert, von den Roh-Transkripten bis zu einer empirisch optimierten probabilistischen Grammatik. Die Analyse ergab drei Hauptkenntnisse:

1. **Empirisch:** Die optimierte Grammatik erreicht eine hohe Korrelation mit den Daten ($r = 0,925$) und offenbart sowohl konstitutive Regeln (Übergänge mit Wahrscheinlichkeit 1,0) als auch statistische Regularitäten.
2. **Interpretativ:** Die Upselling-Schleife ($VBA \rightarrow KBBd, 0,5$) und die Neustart-Option ($KAA \rightarrow VBG, 0,25$) heben strategische Wahlmöglichkeiten in Verkaufsinteraktionen hervor. Die Verabschiedungs-Zurücknahme ($KAV \rightarrow KBBd, 0,5$) zeigt, wie neue Kunden in laufende Interaktionen eintreten können.
3. **Methodologisch:** Die Unterscheidung zwischen konstitutiven Regeln und statistischen Regularitäten bietet eine Grundlage für erklärbare, falsifizierbare und anpassungsfähige neuro-symbolische Systeme.

Die hier präsentierte Grammatik ist kein statisches Artefakt. Sie kann aktualisiert werden, wenn neue Daten eintreffen (statistische Plastizität), und revidiert werden, wenn strukturelle Anomalien erkannt werden (strukturelle Stabilität). In diesem Sinne ist sie ein lebendiges neuro-symbolisches Modell – eine empirische Grammatik, die lernt, während sie erklärbar bleibt.

Literatur

- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Koop, P. (1994). *Grammatikinduktion empirisch gesicherter Verkaufsgespräche*. Scheme-Quellcode.
- Koop, P. (2026). *Von Scheme zu DeepProbLog: Die ARS als methodologischer Bauplan für moderne neuro-symbolische Programmierung*. the-last-freedom.org.
- Manhaeve, R., Dumancic, S., Kimmig, A., Demeester, T., & De Raedt, L. (2018). DeepProbLog: Neural probabilistic logic programming. *Advances in Neural Information Processing Systems*, 31.
- Oevermann, U., Allert, T., Konau, E., & Krambeck, J. (1979). Die Methodologie einer objektiven Hermeneutik. In H.-G. Soeffner (Hrsg.), *Interpretative Verfahren in den Sozial- und Textwissenschaften* (S. 352-434). Metzler.