

Dialogsysteme

Vera Demberg

Einführung in die Computerlinguistik

3. Februar 2015

Was sind Dialogsysteme?



hotel
email
news music
weatherbook selection
recommendationcalendar sights
restaurant stocks sports
booking



Und wozu sind Dialogsysteme gut?

Physische Motivation:

- Gerätebedienung für Behinderte: Aufzug, Rollstuhl, Fahrkartenautomaten, ... (barrierefreie Umgebung)



Physische Motivation:

- Gerätebedienung für Behinderte: Aufzug, Rollstuhl, Fahrkartenautomaten, ... (barrierefreie Umgebung)

Physische Motivation:

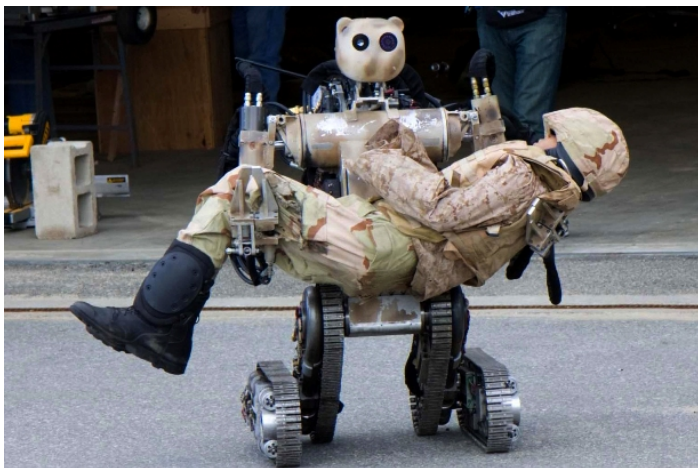
- Gerätebedienung für Behinderte: Aufzug, Rollstuhl, Fahrkartenautomaten, ... (barrierefreie Umgebung)
- Gerätebedienung: Tätigkeiten, bei denen Hände und Augen für andere Aufgaben benötigt werden.



ärztliche Diagnose, Steuerung von Instrumenten in der Chirurgie



Autofahren (weiterhin auf Strasse schauen, Lenkrad anfassen)



Generell: hands-free/eyes-on Anwendungen

Physische Motivation:

- Gerätebedienung für Behinderte: Aufzug, Rollstuhl, Fahrkartenautomaten, ... (barrierefreie Umgebung)
- Gerätebedienung: Tätigkeiten, bei denen Hände und Augen für andere Aufgaben benötigt werden.
- Telefonie: Anwendung, in denen nur der akustische Kanal zur Verfügung steht (Fahrplan-Auskunft, Kundendienst, Telebanking, ...)

Semantische Motivation

- Natürliche Sprache ist attraktiv, weil sie das natürliche Kommunikationsmedium des Menschen ist.
- Wir können auf dasselbe Objekt/ dieselbe semantische Information mit unterschiedlichsten sprachliche Mitteln zugreifen.
- Dialogschnittstellen ermöglichen dem Benutzer, seine Absichten inhaltsorientiert und intuitiv, ohne Spezialwissen und ohne Auswendiglernen von Sprachkommandos zu realisieren.

Semantische Motivation

- Natürliche Sprache ist attraktiv, weil sie das natürliche Kommunikationsmedium des Menschen ist.
- Wir können auf dasselbe Objekt/ dieselbe semantische Information mit unterschiedlichsten sprachliche Mitteln zugreifen.
- Dialogschnittstellen ermöglichen dem Benutzer, seine Absichten inhaltsorientiert und intuitiv, ohne Spezialwissen und ohne Auswendiglernen von Sprachkommandos zu realisieren.

Der Benutzer sagt, was er will, so, wie er es will.

Adaptive Dialogsysteme



Intuitive Gerätebedienung:
Sprachdialog für Nicht-Experten; spielerischer Roboterfreund

Adaptive Dialogsysteme



Adaptive Dialogsysteme



“Companion-Technologie”:
Service-Roboter im privaten Haushalt und in Pflege-Einrichtungen

Adaptive Dialogsysteme



Haus:

Unterstützt ältere Menschen, in einer immer technischer werdenden
Alltagsumgebung lange selbständig zu bleiben.

Ein einfaches Beispiel

Fahstuhldialog

User: Fahrstuhl

System: In welches Stockwerk möchten Sie bitte?

User: Vier / Vierter Stock

System: 〈Kommando wird ausgeführt〉

Ein einfaches Beispiel

Fahstuhldialog

User: Fahrstuhl

System: In welches Stockwerk möchten Sie bitte?

User: Vier / Vierter Stock

System: Ich fahre Sie in den vierten Stock.

System: 〈Kommando wird ausgeführt〉

Ein einfaches Beispiel

Fahstuhldialog

User: Fahrstuhl

System: In welches Stockwerk möchten Sie bitte?

User: Vier / Vierter Stock

System: Ich fahre Sie in den dritten Stock.

User: Nein, in den vierten Stock!

System: Ich fahre Sie in den vierten Stock.

System: 〈Kommando wird ausgeführt〉

Ein einfaches Beispiel

Fahstuhldialog

User: Fahrstuhl

System: In welches Stockwerk möchten Sie bitte?

User: Vier / Vierter Stock

System: Ich habe Sie nicht so genau verstanden.
Wollen Sie in den fünften Stock?

User: Nein, in den vierten Stock!

System: Ich fahre Sie in den vierten Stock.

System: 〈Kommando wird ausgeführt〉

Grounding

“Grounding” = Verständigungssicherung

- Klärungsfragen des Systems (explizites Grounding)
- Implizite Verständigungssicherung, die Gelegenheit zur Korrektur durch den Benutzer gibt (implizites grounding)

Optimistische Strategie Verzicht auf Grounding im Erfolgsfall effizient, bei Missverständnis Effektivitäts- und/oder Effizienzproblem

Vorsichtige Strategie Durchgängig (explizites) Grounding maximiert Effektivität, aber auf Kosten von Effizienz und Benutzerzufriedenheit

Flexible Strategie Je nach Konfidenz (Wahrscheinlichkeit der besten Erkennerrhypothese) kein Grounding/ implizites Grounding/ explizite Rückfrage

Qualitätskriterien für Dialogsysteme

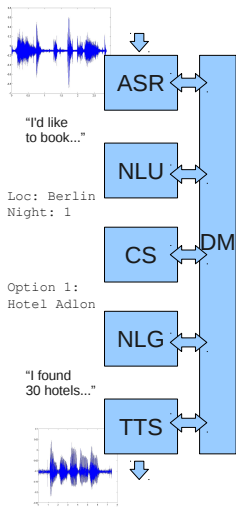
Qualitätskriterien

- **Effektivität:** Der Dialog führt zum (gewünschten) Ergebnis
- **Effizienz:** Der Dialog führt in angemessener Zeit/ in einer angemessenen Anzahl von Dialogschritten zum Ergebnis
- **Benutzerzufriedenheit**

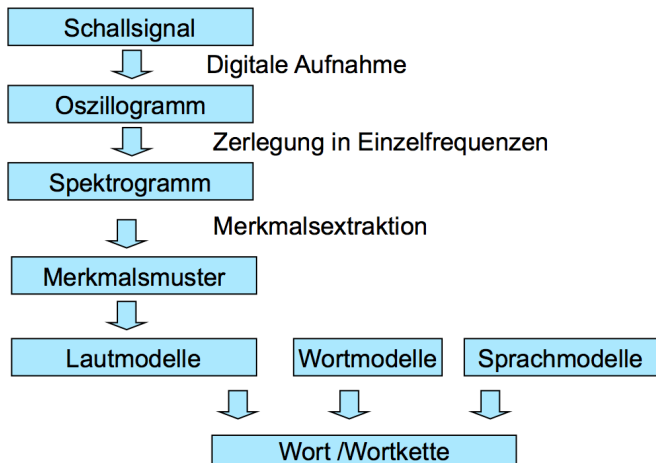
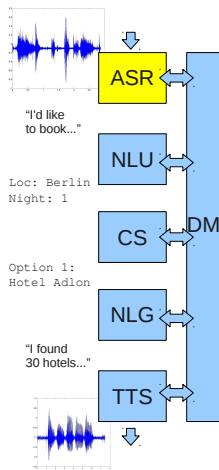
Teil 1: Aufbau von Dialogsystemen

Aufbau von Dialogsystemen

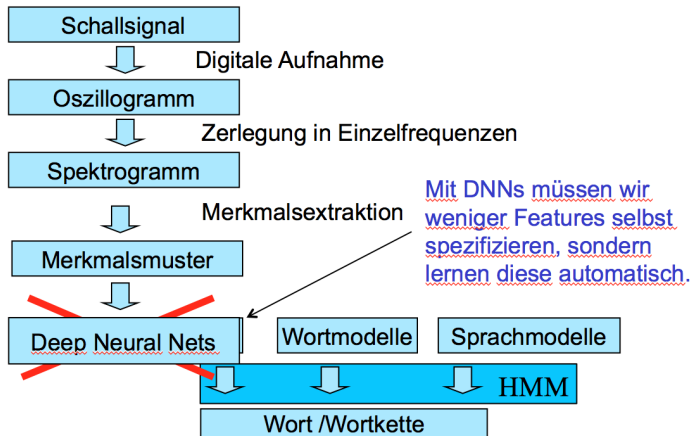
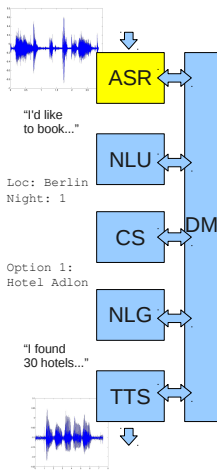
System Components



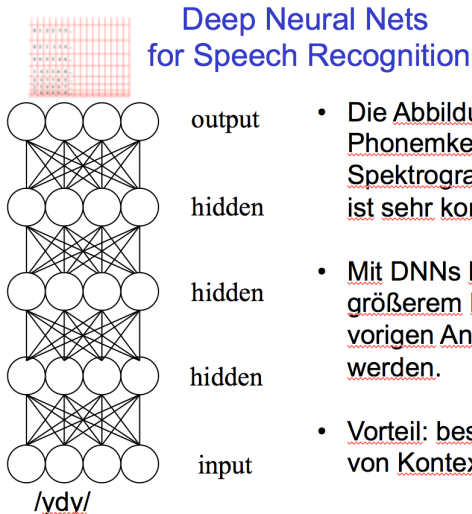
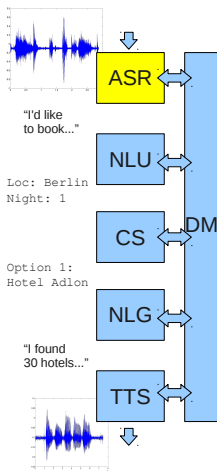
Automatic Speech Recognition (= Spracherkennung)



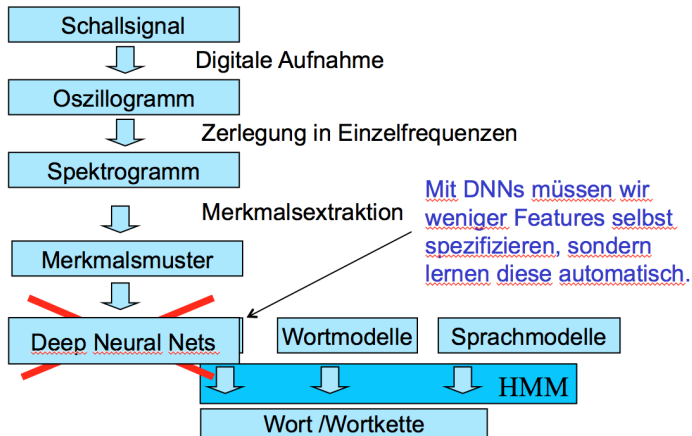
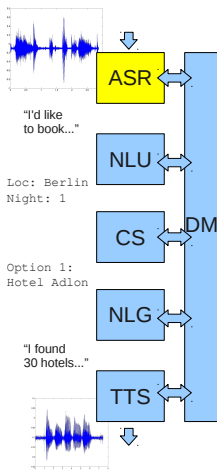
Automatic Speech Recognition (= Spracherkennung)



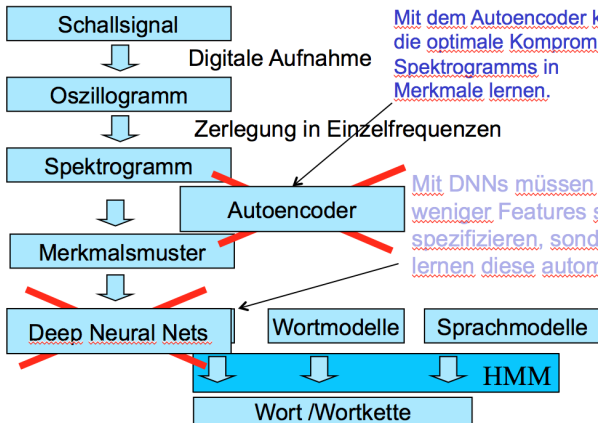
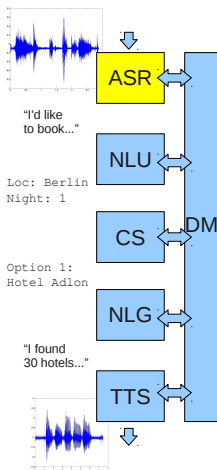
Automatic Speech Recognition (= Spracherkennung)



Automatic Speech Recognition (= Spracherkennung)



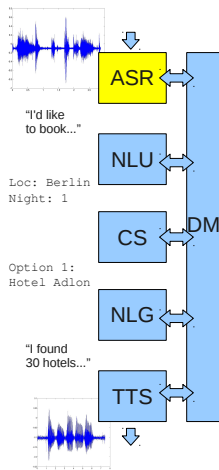
Automatic Speech Recognition (= Spracherkennung)



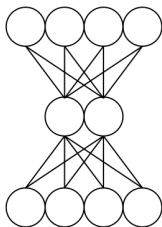
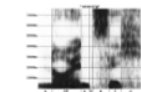
Mit dem Autoencoder können wir die optimale Kompromierung des Spektrogramms in Merkmale lernen.

Mit DNNs müssen wir weniger Features selbst spezifizieren, sondern lernen diese automatisch.

Automatic Speech Recognition (= Spracherkennung)



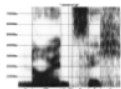
Autoencoder für Merkmalsmusterextraktion



output

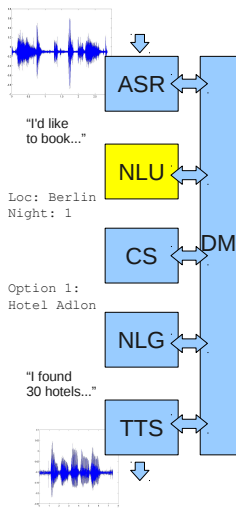
hidden

input



- Der Autoencoder soll als Output den Input möglichst akkurat reproduzieren.
- Durch die kleineren hidden Layers (hier nur schematisch als ein hidden Layer dargestellt, können aber mehr sein) wird die relevante Information mit nur minimalem Informationsverlust komprimiert.
- Vorteil: geringerer Informationsverlust als bei traditioneller Merkmalsextraktion mit Zeit-Frequenz-fenster.

Natural Language Understanding



Example for a hotel booking system

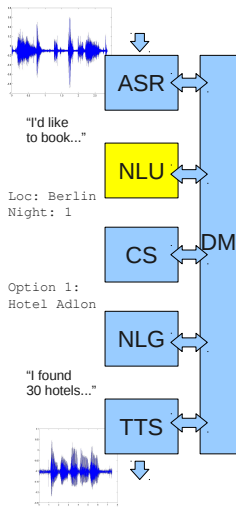
ASR output

I'd like to book a double room for 2 adults in Berlin for one night.

Frame

city:	Berlin
arrival:	
departure:	
duration:	1
# of people:	2
room type:	double
area:	Grunewald

Natural Language Understanding



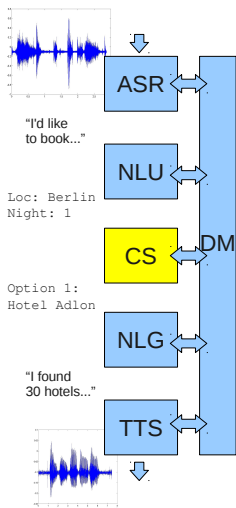
Strategien:

- key-word spotting
- parsing

Aufgabe: was bedeuten die Wörter in diesem Kontext? Wie verhalten sie sich zu einander? Was bedeutet es für die Dialogaufgabe?

Wir brauchen sowohl Linguistik als auch Maschinelles Lernen.

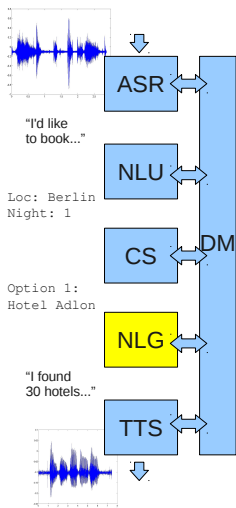
Content Selection



Inhalte aussuchen und strukturieren:

- 1 Relevante Informationen aus Datenbank abrufen
- 2 Eventuell Selektion der Information, die dem Nutzer präsentiert werden soll
- 3 Strukturierung der Information

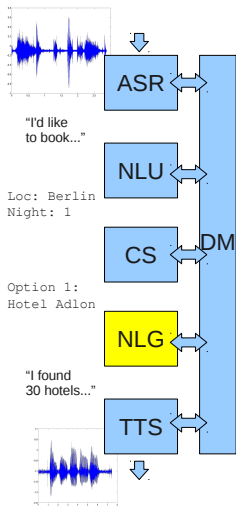
Sprachgenerierung: Natural Language Generation



Textgenerierung:

- ① Was soll in einem Satz gesagt werden?
- ② aus den Daten werden natürlichsprachliche Sätze gebaut
 - Templates
 - Grammatik-basierte Generierung
 - datengetriebene Generierung

Sprachgenerierung: Natural Language Generation



Sets of sentence templates, e.g.,

For the MonthlyTemperatureMsg:

TempString = case (TEMP - AVERAGETEMP)

[2.0 ... 2.9]: 'very much warmer than average.'

[1.0 ... 1.9]: 'much warmer than average.'

[0.1 ... 0.9]: 'slightly warmer than average.'

[-0.1 ... -0.9]: 'slightly cooler than average.'

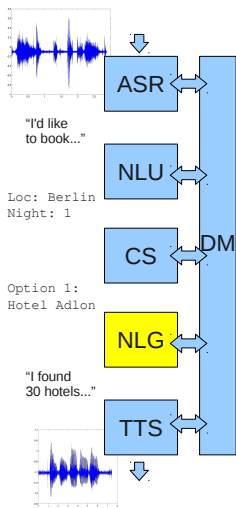
[-1.0 ... -1.9]: 'much cooler than average.'

[-2.0 ... -2.9]: 'very much cooler than average.'

endcase

Sentence = 'The month was' + TempString

Sprachgenerierung: Natural Language Generation



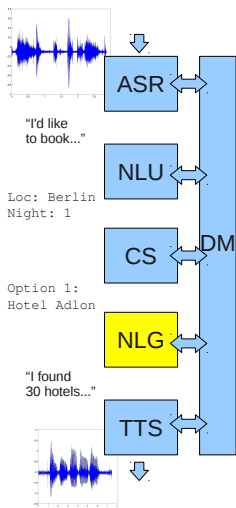
The Result:

The month was cooler than average.
 The month was drier than average.
 There was the average number of rainy days.
 The total rain for the year so far is well below average.
 There was rain on every day for 8 days from 11th to 18th.
 Rainfall amounts were mostly small.

What we'd really like:

The month was cooler and drier than average. The total rain for the year so far is well below average, even though there was an average number of rainy days this month. There was rain on every day for 8 days from 11th to 18th, but rainfall amounts were mostly small.

Sprachgenerierung: Natural Language Generation



Natürlicher klingende Sprache generieren:

*Heavy rain fell on the 27th.
Heavy rain fell on the 28th.*

↓Aggregation mit Konjunktion↓

Heavy rain fell on the 27th and heavy rain fell on the 28th.

↓Aggregation mit Ellipse ↓

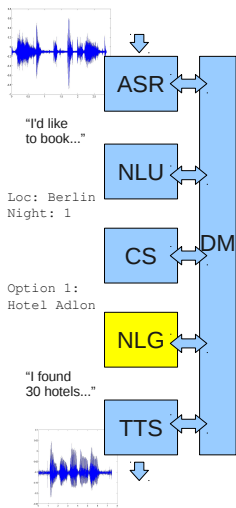
Heavy rain fell on the 27th and [] on the 28th.

↓Aggregation mit Menge↓

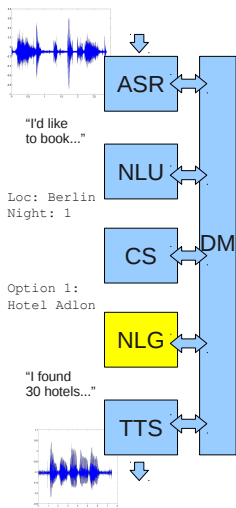
Heavy rain fell on the 27th and 28th.

Sprachgenerierung: Natural Language Generation

Es gibt fast immer Formulierungsalternativen.



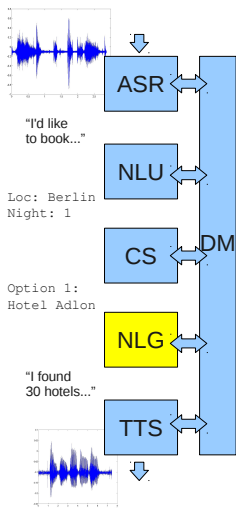
Sprachgenerierung: Natural Language Generation



Es gibt fast immer Formulierungsalternativen.

- S:** The flight is cheap / inexpensive / not expensive
- NP:** a flight with a low fare
- AP:** cheaper than average

Sprachgenerierung: Natural Language Generation



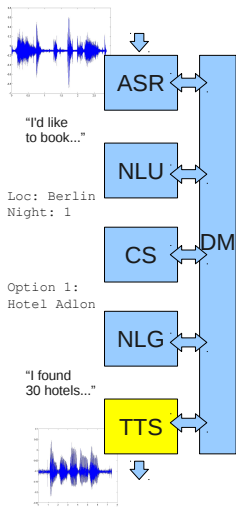
Es gibt fast immer Formulierungsalternativen.

- S:** The flight is cheap / inexpensive / not expensive
- NP:** a flight with a low fare
- AP:** cheaper than average

Beste Wahl hängt ab von...

- Nutzervorlieben
- Konsistenz mit vorherigen Formulierungsentscheidungen
- Fokus und Stil
- weitere Informationen die übermittelt werden sollen.

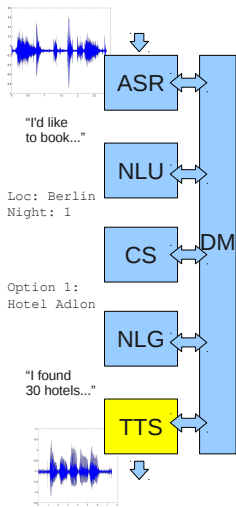
Text-to-speech Synthesis



Sprachsynthese:

- Wie muss ein Wort ausgesprochen werden?
 - Lesart
 - modern
 - Aktentaschen
 - 1/80 → ein Achtzigstel

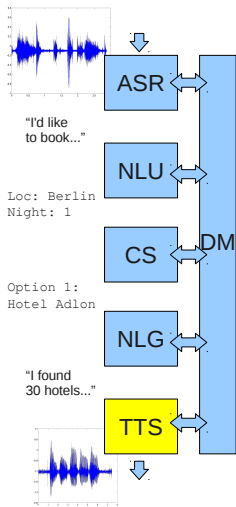
Text-to-speech Synthesis



Sprachsynthese:

- Wie muss ein Wort ausgesprochen werden?
 - Lesart
 - modern
 - Aktentaschen
 - 1/80 → ein Achtzigstel
- Methoden:

Text-to-speech Synthesis

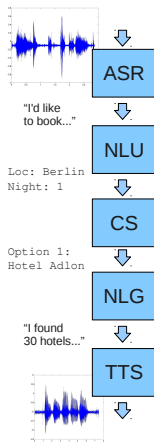


Sprachsynthese:

- Wie muss ein Wort ausgesprochen werden?
 - Lesart
 - modern
 - Aktentaschen
 - 1/80 → ein Achtzigstel
- Methoden:
 - Einzellautsynthese (unmöglich)
 - voraufgenommene Sprache (unpraktisch)
 - Diphonsynthese
 - Wortkonkatenation
 - Unit Selection
 - artikulatorische Synthese

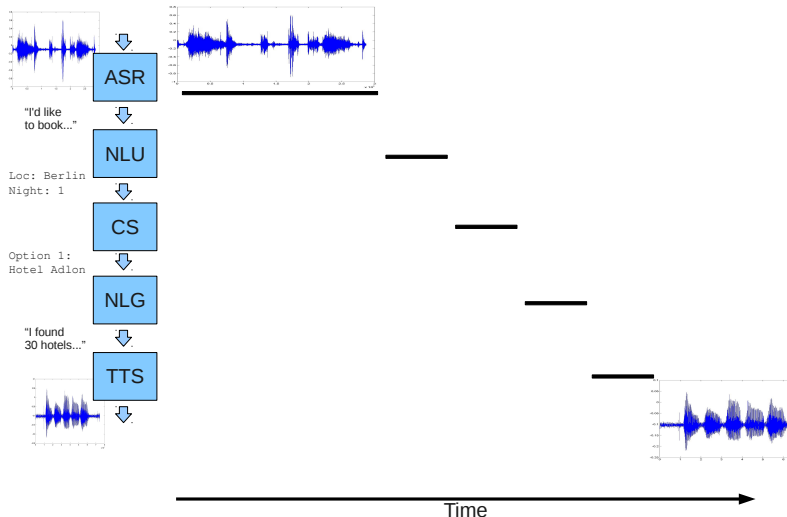
Inkrementelle Dialogsysteme

Ein traditionelles Dialogsystem:



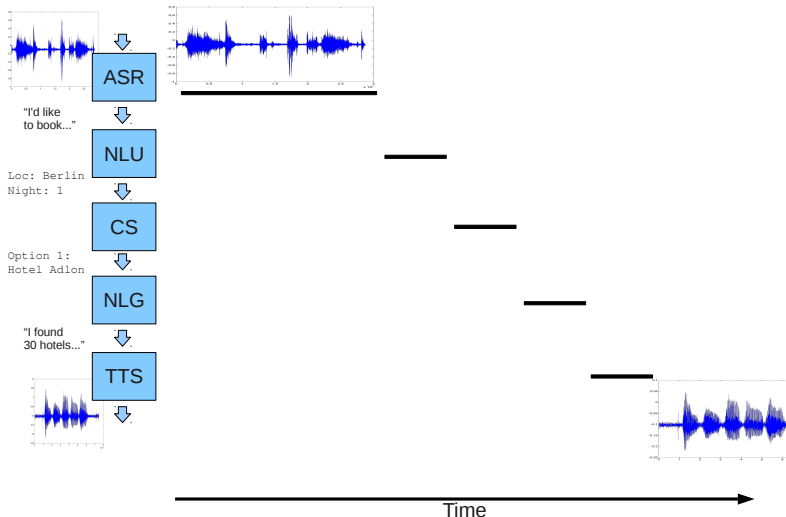
Inkrementelle Dialogsysteme

Ein traditionelles Dialogsystem:



Inkrementelle Dialogsysteme

Ein traditionelles Dialogsystem: Beispiel



Teil 2: Inkrementelle Dialogsysteme

Inkrementelle Dialogsysteme

Menschlicher Dialog

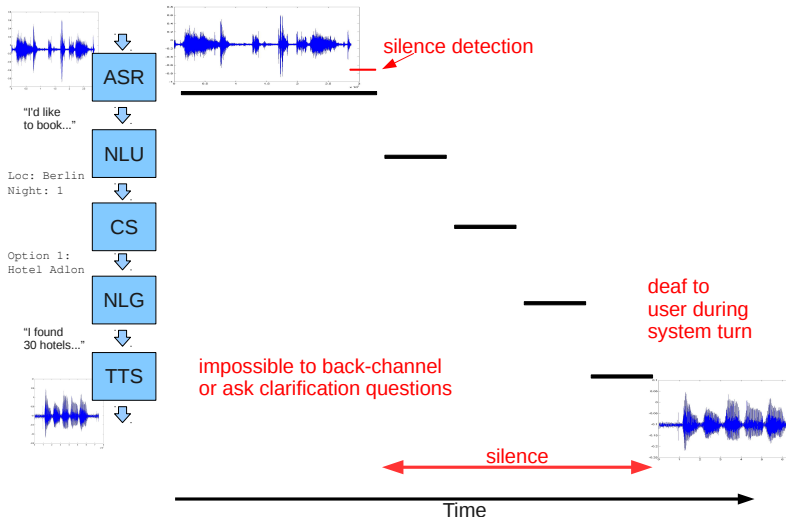


Wie unterscheiden sich typische menschliche Dialoge davon?

- keine langen Pausen
- Bestätigung (back-channeling: mhm, ok)
- Hesitationen (uh, uhm)
- Füller (ok let me see, yeah...)
- Unterbrechung mit Klärungsfragen (“barge in”)

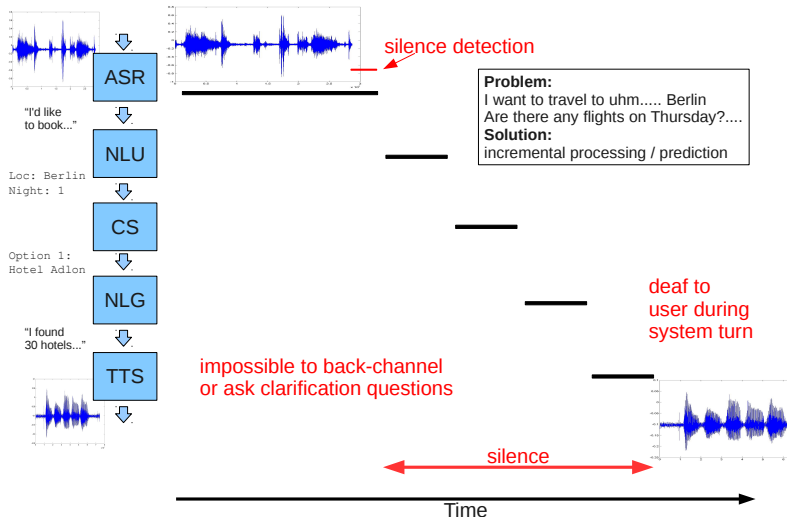
Incremental Processing in SDS

Ein traditionelles Dialogsystem



Incremental Processing in SDS

Ein traditionelles Dialogsystem



Sprachverarbeitung beim Menschen

Sprachverstehen: [e.g., Marslen-Wilson 1973, Kamide et al. 2003]

- Menschen verarbeiten Sprache inkrementell, und machen sogar Vorhersagen.

Sprachverarbeitung beim Menschen

Sprachverstehen: [e.g., Marslen-Wilson 1973, Kamide et al. 2003]

- Menschen verarbeiten Sprache inkrementell, und machen sogar Vorhersagen.

Wann sprechen im Dialog? [e.g., de Ruiter et al., 2006]

- extrem kurze Pausen zwischen Äußerungen
- nur möglich wenn vorhergesagt wird, wann der andere mit Sprechen fertig ist.

Sprachverarbeitung beim Menschen

Sprachverstehen: [e.g., Marslen-Wilson 1973, Kamide et al. 2003]

- Menschen verarbeiten Sprache inkrementell, und machen sogar Vorhersagen.

Wann sprechen im Dialog? [e.g., de Ruiter et al., 2006]

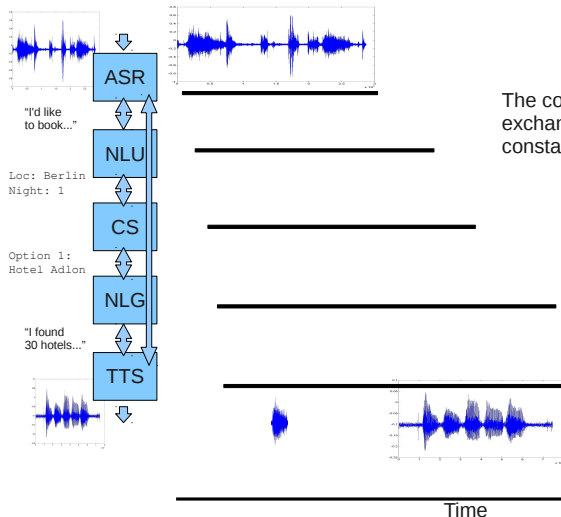
- extrem kurze Pausen zwischen Äußerungen
- nur möglich wenn vorhergesagt wird, wann der andere mit Sprechen fertig ist.

Sprachproduktion: [e.g., Ferreira, 1999; Gambi 2013]

- Sprecher fangen oft an zu sprechen, bevor sie die gesamte Äußerung zuendegeplant haben.
- Sprachgeschwindigkeit hängt von Planung der Äußerung ab
- Selbstverbesserung und Zögern sind häufig

Inkrementelle Verarbeitung in SDS

Was heisst inkrementelle Verarbeitung beim Dialogsystem?



Warum inkrementelle Systeme?

Vorteile eines inkrementellen Systems:

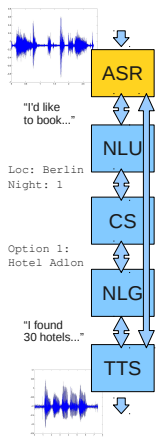
- schnellere Antworten
- kann Pause von Äußerungsende besser unterscheiden
- kann Bestätigungssignale geben
- kann auf Zwischenfragen reagieren

Wie inkrementell sind heutige Dialogsysteme?

Siri (nicht inkrementell) vs. Google (inkrementell)



Automatic Speech Recognition



Pausenerkennung:

traditionell: erkenne die Pause, initiiere dann nächsten Verarbeitungsschritt

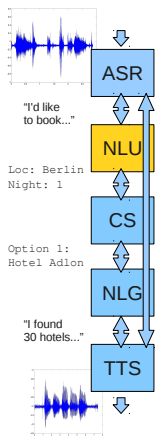
inkrementell: Pausen sind relevant für Sprechinitiative des Systems.

- NLU (tiefe Verarbeitung)
- TTS (flache Verarbeitung)

Spracherkennung:

inkrementell: beste Hypothesen werden ständig an NLU Komponente weitergegeben.

Natural Language Understanding



Example for a hotel booking system

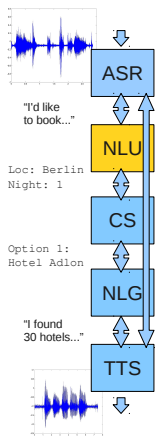
ASR output (traditional)

I'd like to book a double room for 2 adults in Berlin for one night.

Frame (traditional)

city:	Berlin
arrival:	
departure:	
duration:	1
# of people:	2
room type:	double

Natural Language Understanding

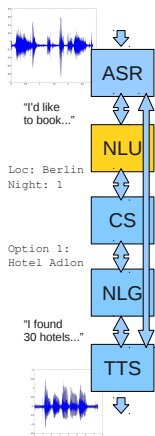


Example for a hotel booking system

ASR output (incremental)
I'd like

→ Frame (incremental)

Natural Language Understanding



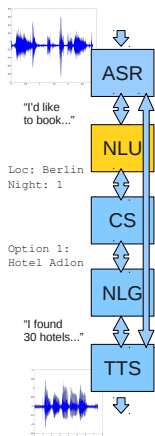
Example for a hotel booking system

ASR output (incremental)
I'd like to book

Frame (incremental)

city:
arrival:
departure:
duration:
of people:
room type:

Natural Language Understanding



Example for a hotel booking system

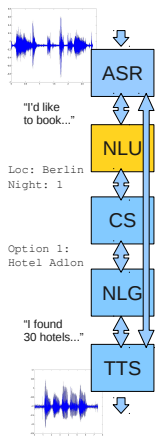
ASR output (incremental)
I'd like to book a double

Frame (incremental)

```

city:
arrival:
departure:
duration:
# of people:
room type: double
  
```

Natural Language Understanding



Example for a hotel booking system

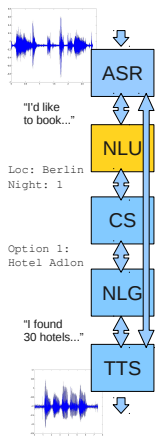
ASR output (incremental)

I'd like to book a double
room for 2 adults

Frame (incremental)

city:
arrival:
departure:
duration:
of people: 2
room type: double

Natural Language Understanding



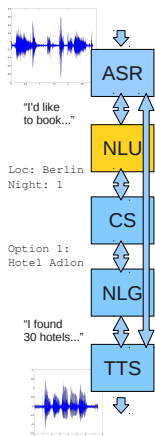
Example for a hotel booking system

ASR output (incremental)
I'd like to book a double
room for 2 adults in Berlin

Frame (incremental)

city: Berlin
arrival:
departure:
duration:
of people: 2
room type: double

Natural Language Understanding



Example for a hotel booking system

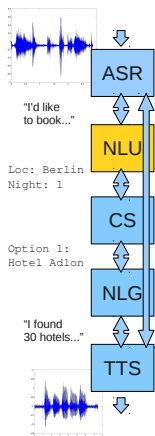
ASR output (incremental)

I'd like to book a double
room for 2 adults in Berlin
for one night.

Frame (incremental)

city:	Berlin
arrival:	
departure:	
duration:	1
# of people:	2
room type:	double

Natural Language Understanding



Example for a hotel booking system

ASR output (incremental)

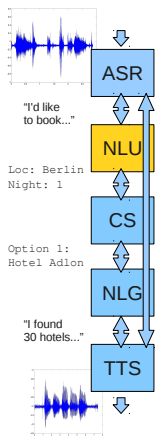
I'd like to book a double
room for 2 adults in Berlin
oh, I meant Potsdam.

Frame (incremental)

city: BerlinPotsdam
arrival:
departure:
duration: 1
of people: 2
room type: double

- Need to cope with self-corrections.

Natural Language Understanding



Example for a hotel booking system

ASR output (incremental)

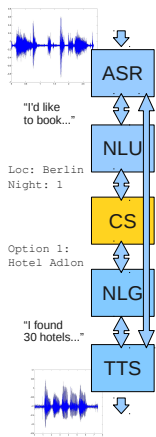
I'd like to book a double room for 2 adults in Berlin near Potsdam.

Frame (incremental)

city:	Berlin
arrival:	
departure:	
duration:	1
# of people:	2
room type:	double

- Keyword-spotting vs. full parsing / semantic interpretation

Content Planning, Selection and Structuring



Entscheiden, was das System sagen soll.

- Datenbankabfrage schon starten sobald ein Teil der Information verarbeitet ist.
- sobald wie möglich Reaktion auswählen
- eventuell muss Entscheidung revidiert werden

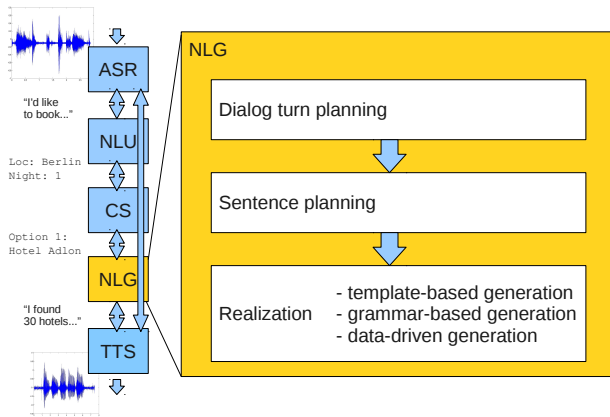
Example

USER: *I want to go from London in the morning.*

SYS: *From London in the morning, there are 10 flights.*

[e.g., Skantze and Hjalmarsen, 2013]

Natural language generation



[Buschmeier et al., 2012]
[Hough, 2012]

Generate output in chunks:

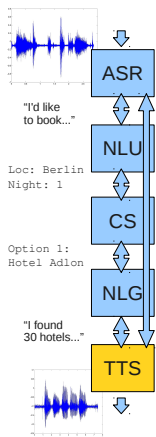
Your appointment| is today| at 4:25| in building C7.4| room 3.05.
sorry when?

at 4:25| in building C7.4...

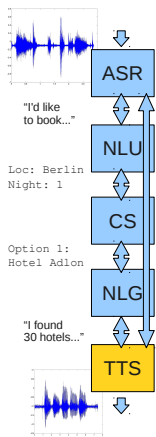
Sprachsynthese

Sprachsynthese

- **traditionell:** Synthese beginnt wenn vollständiger Satz generiert wurde
- **inkrementell:** muss mit Teilsätzen umgehen und trotzdem gute Satzmelodie generieren.



Sprachsynthese



Sprachsynthese

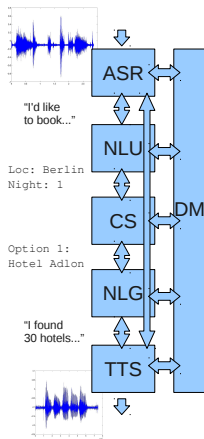
- **traditionell:** Synthese beginnt wenn vollständiger Satz generiert wurde
- **inkrementell:** muss mit Teilsätzen umgehen und trotzdem gute Satzmelodie generieren.

weitere Schwierigkeiten:

- Input kann sich ändern!
- Das Synthesesytem muss beobachten, wie weit es schon generiert hat um sich selbst verbessern zu können und auf Zwischenfragen reagieren zu können.
- Pausen füllen, falls noch kein weiterer Input vorhanden

[see also Buschmeier et al., 2012]

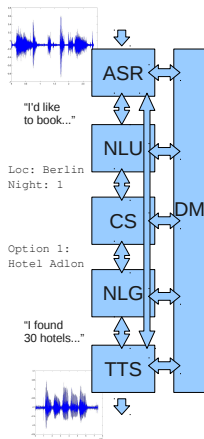
Dialog Management



Informationsfluss:

- Wann sollte ich sprechen / bestätigen?
- Was wurde schon gesagt?
- Updates und ggf. Selbstverbesserungen koordinieren.
- angemessen auf Zwischenfragen und Bestätigungen durch den Nutzer reagieren.

Dialog Management



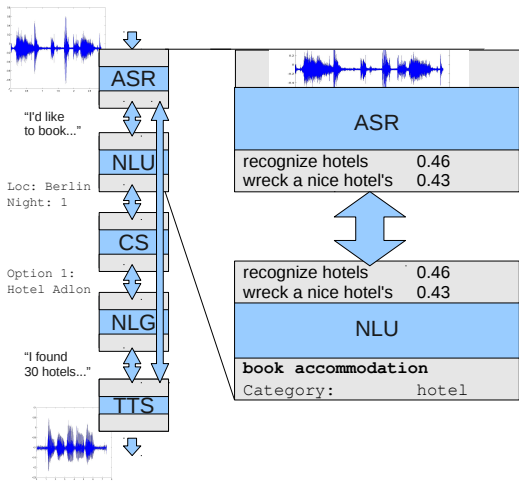
Informationsfluss:

- Wann sollte ich sprechen / bestätigen?
- Was wurde schon gesagt?
- Updates und ggf. Selbstverbesserungen koordinieren.
- angemessen auf Zwischenfragen und Bestätigungen durch den Nutzer reagieren.

Architektur: Wie kann die Information geteilt werden?

- Dialogmanager vs. geteilte Buffer

Dialog Management



Wann welche Information teilen?

- Beispiel ASR: fast sofortige Hypothesen, aber nicht stabil
- Kompromiss zwischen Inkrementalität und Hypothesenstabilität

Teil 3:

Nutzeranpassung

Beispiel für Nutzeranpassung

USER: I want to travel from Edinburgh to Brussels, arriving by 5 pm.

For a starving student:

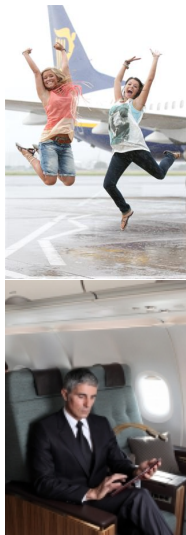
SYS: *There's a direct flight on BMI with a good price. It arrives at four ten p.m. and costs one hundred and twelve pounds. The cheapest flight is on Ryanair. It arrives at twelve forty five p.m. and costs just fifty pounds, but it requires a connection in Dublin.*

For a business traveller:

SYS: *You can fly business class on British Airways, arriving at four twenty p.m., but youd need to connect in Manchester. There's a direct flight on BMI, arriving at four ten p.m., but theres no availability in business class.*

Figure: Beispiel FLIGHTS system; White et al., 2010

Nutzermodellierung in SDS



Nutzermodellierung bei Informationspräsentation:

- System kann Alternativen aufzeigen
- Evaluierung zeigt:
 - Nutzer zufriedener
 - effizientere Interaktion
 - beste Option wird mit höherer Wahrscheinlichkeit gefunden

[Demberg et al., 2011]

Optimale Präsentation ist situationsabhängig



Bei gleichzeitigem Autofahren führte komplexere Strategie zu:

- mehr Fahrfehlern
- schlechterer Performance
- Nutzer fühlten sich überfordert

[Winterboer et al., 2007; Hu et al., 2007]

Sprechen und Autofahren



- Handy im Auto
 - gefährlich
 - Handy wird bei 5% der Fahrten genutzt, aber in 18% der Unfälle lag Handynutzung vor.
 - “inattention blindness”

Sprechen und Autofahren



- Handy im Auto
 - gefährlich
 - Handy wird bei 5% der Fahrten genutzt, aber in 18% der Unfälle lag Handynutzung vor.
 - "inattention blindness"
- Freisprechanlage
 - gleicher Effekt wie Handy
 - vergleichbar mit Trunkenheit am Steuer

Sprechen und Autofahren



- Handy im Auto
 - gefährlich
 - Handy wird bei 5% der Fahrten genutzt, aber in 18% der Unfälle lag Handynutzung vor.
 - "inattention blindness"
- Freisprechanlage
 - gleicher Effekt wie Handy
 - vergleichbar mit Trunkenheit am Steuer
- Beifahrer
 - ok

Sprechen und Autofahren



- Handy im Auto
 - gefährlich
 - Handy wird bei 5% der Fahrten genutzt, aber in 18% der Unfälle lag Handynutzung vor.
 - "inattention blindness"
- Freisprechanlage
 - gleicher Effekt wie Handy
 - vergleichbar mit Trunkenheit am Steuer
- Beifahrer
 - ok

Warum?

Was ist der Unterschied zwischen Beifahrer und Freisprechanlage?

Beifahrer passt sich an an Fahrsituation und Stress des Fahrers.

- in schwierigen Fahrsituationen ist Sprachkomplexität geringer bei Fahrer und Beifahrer (Drews et al., 2008)
- Thema wird auf Fahrsituation gelenkt (Villing 2009a,b)
- Weniger Äußerungen bei schwierigerer Stadtfahrt als auf Landstraße. (Crundall et al., 2005)



Was ist der Unterschied zwischen Beifahrer und Freisprechanlage?

Beifahrer passt sich an an Fahrsituation und Stress des Fahrers.

- in schwierigen Fahrsituationen ist Sprachkomplexität geringer bei Fahrer und Beifahrer (Drews et al., 2008)
- Thema wird auf Fahrsituation gelenkt (Villing 2009a,b)
- Weniger Äußerungen bei schwierigerer Stadtfahrt als auf Landstraße. (Crundall et al., 2005)

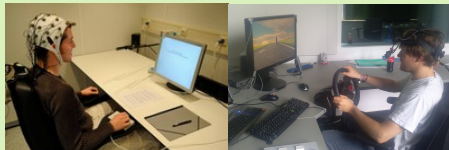


Dialogsysteme sollen sich verhalten wie der Beifahrer, nicht wie der entfernte Gesprächspartner via Freisprechanlage.

Wie kann Sprachkomplexität kontrolliert werden?

Experimentelle Psycholinguistik

- Welche Sprachkonstruktionen sind schwer zu verstehen?
- Wie interferiert Sprachkomplexität mit weiteren Aufgaben?
- Individuelle Unterschiede



Computerbasierte Modellierung

- Computerbasierte kognitive Modelle von menschlichem Sprachverständnis

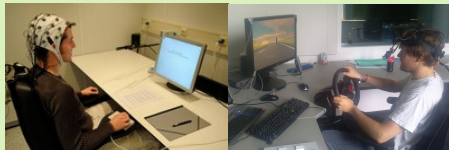
Sprachanwendungen

- Integration von psycholinguistischen Modellen in adaptive Generierungssysteme

Wie kann Sprachkomplexität kontrolliert werden?

Experimentelle Psycholinguistik

- Welche Sprachkonstruktionen sind schwer zu verstehen?
- Wie interferiert Sprachkomplexität mit weiteren Aufgaben?
- Individuelle Unterschiede



Computerbasierte Modellierung

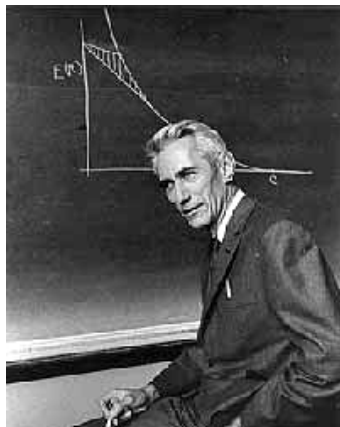
- Computerbasierte kognitive Modelle von menschlichem Sprachverständnis

Sprachanwendungen

- Integration von psycholinguistischen Modellen in adaptive Generierungssysteme

Modellierung von Verarbeitungsschwierigkeit: Surprisal

Informations-theoretisches Maß: Wieviel Information überträgt ein Wort?



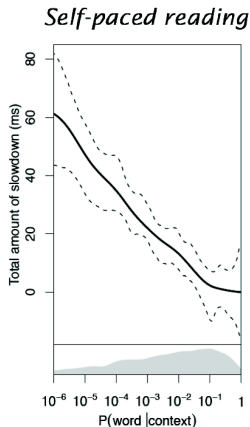
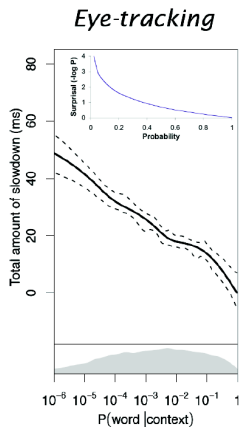
Claude Shannon

Modellierung von Verarbeitungsschwierigkeit: Surprisal

Verarbeitungsschwierigkeit beim Menschen proportional zum Informationsgehalt.

Modellierung von Verarbeitungsschwierigkeit: Surprisal

Verarbeitungsschwierigkeit beim Menschen proportional zum Informationsgehalt.



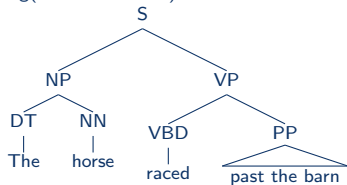
[from Smith
and Levy, 2013]

Modellierung von Verarbeitungsschwierigkeit: Surprisal

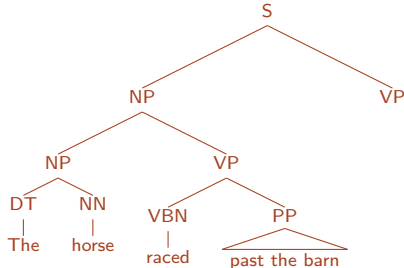
- Annahme: Verarbeitungsschwierigkeit beim Menschen proportional zum Informationsgehalt.
- $Surprisal(w_{k+1}) = -\log P(w_{k+1}|w_1, w_2, \dots, w_k)$
- Beispiel 1: Peter hit the nail on the **head**.
Ein sehr vorhersagbares Wort trägt wenig neue Information
→ deshalb ist es einfach zu verarbeiten.
- Beispiel 2: The boy kicked the ball **kicked** the ball.
Ein sehr überraschendes Wort liefert viel Information (hier vorallem dazu, wie das erste kicked the ball zu interpretieren ist)
→ daher ist es schwer zu verarbeiten.

Syntactic Surprisal

$$-\log(1.7766 \times 10^{-11}) = 35.712$$



$$-\log(1.06596 \times 10^{-15}) = 49.736$$



sum of both: $pp_{w_n} = 35.712$

How to calculate surprisal:

- Calculate prefix probabilities:

$$pp_{w_n} = -\log \sum_{T \in \text{Trees}} p(T | w_1 \dots w_n)$$

- Surprisal s of word w_n :

$$s_{w_n} = pp_{w_n} - pp_{w_{n-1}}$$

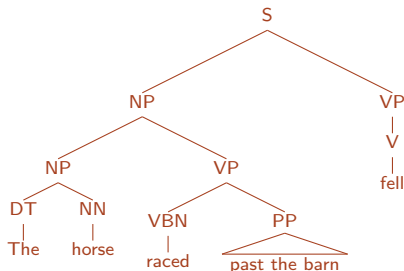
Example PCFG:

Rule	Probability of rule
$S \rightarrow NP VP$	$p = 0.6$
$VBD \rightarrow \text{raced}$	$p = 0.0005$
$VBN \rightarrow \text{raced}$	$p = 0.000001$
$DT \rightarrow \text{the}$	$p = 0.7$

Syntactic Surprisal

$$pp_{w_{n+1}} = -\log(1.06596 \times 10^{-15} \times 0.003)$$

$$= 58.12$$



How to calculate surprisal:

- Calculate prefix probabilities:

$$pp_{w_n} = -\log \sum_{T \in \text{Trees}} p(T | w_1 \dots w_n)$$

- Surprisal s of word w_n :

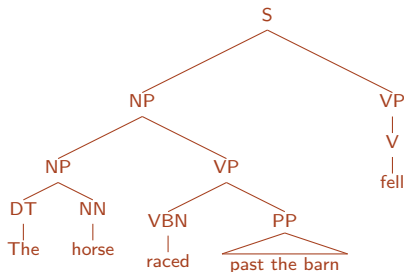
$$s_{w_n} = pp_{w_n} - pp_{w_{n-1}}$$

Example PCFG:

Rule	Probability of rule
$S \rightarrow NP VP$	$p = 0.6$
$VBD \rightarrow \text{raced}$	$p = 0.0005$
$VBN \rightarrow \text{raced}$	$p = 0.000001$
$DT \rightarrow \text{the}$	$p = 0.7$

Syntactic Surprisal

$$pp_{w_{n+1}} = -\log(1.06596 \times 10^{-15} \times 0.003) \\ = 58.12$$



$$pp_{w_{n-1}} = 35.712$$

$$pp_{w_n} = 58.12$$

$$\text{surprisal}(w_n) = 22.41$$

How to calculate surprisal:

- Calculate prefix probabilities:

$$pp_{w_n} = -\log \sum_{T \in \text{Trees}} p(T|w_1 \dots w_n)$$

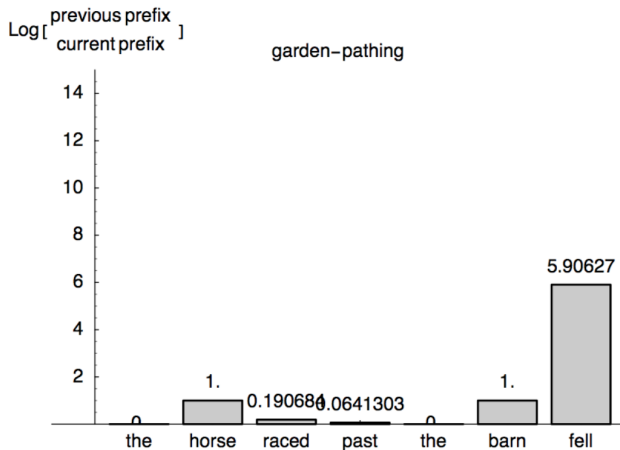
- Surprisal s of word w_n :

$$s_{w_n} = pp_{w_n} - pp_{w_{n-1}}$$

Example PCFG:

Rule	Probability of rule
$S \rightarrow NP VP$	$p = 0.6$
$VBD \rightarrow \text{raced}$	$p = 0.0005$
$VBN \rightarrow \text{raced}$	$p = 0.000001$
$DT \rightarrow \text{the}$	$p = 0.7$

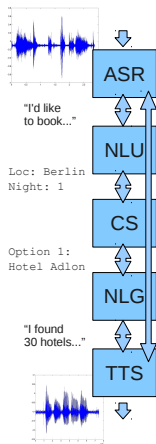
Surprisal – Modelling Results (1)



[Hale, 2001]

Wie kann sich ein Dialogsystem an den Nutzer anpassen?

Adaptive Dialogsysteme:



ASR Sprechererkennung und individuelle Spracherkennung

NLU richtige Momente für Bestätigung und Verständnisfragen, Frequenz an Nutzer anpassen

CS Anpassung des Inhalts

NLG Komplexität und Formulierungen and Situation und Nutzer anpassen

TTS Störungen insbesondere an Stellen von hohem Interesse für den Nutzer vermeiden