

Proseminar “Generierung”

2. Sitzung: Einführung

Alexander Koller
29. Oktober 2010

```
Unuz Ono, "The Attack of Hatchets"
A: 361, mostly goblins, 131 losses
D: 447 elves, 120 losses
Defender was victorious.

In 192, the human Ozud Oldstolen's left eye was torn
out by the elf Avafi Matchedglossed the Assaulted
Slaughters.
In 192, the elf Avafi Matchedglossed the Assaulted
Slaughters's left foot was ripped by the human Ozud
Oldstolen.
In 192, the elf Masami Glossrams's left upper arm was
stabbed by the elf Arstruk Rabblevile.
In 192, the goblin Stâsost Flysoots was shot and
killed by the elf Arane Pageford the Mighty Romance in
Beancyclones.
In 192, the goblin Strodno Feverdoomed was shot and
killed by the elf Gica Ownmusic the Enjoyable Sun of
Venerating in Beancyclones.
In 192, the elf Ula Snakesperplexes was struck down by

ESC: done.      Shift + ESC: back to age.
```

Übersicht

- Wie man gute Vorträge hält
- Generierung

Wie hält man gute Vorträge?

- Inhalt
- Geschichte
- Motivation
- Gestaltung
- Vom Artikel zu den Folien

Inhalt

- Grundlegendes Ziel eines Vortrags:
Zuhörer sollen verstehen, wie es geht.
- Vortrag muss also:
 - ▶ inhaltlich korrekt und vollständig sein
 - ▶ verständlich
 - ▶ unterhaltsam
 - ▶ zur Diskussion einladen

Struktur

- Ein Vortrag hat
 - ▶ eine Einleitung: Was ist das Problem? Warum ist es interessant?
 - ▶ einen Hauptteil: Wie löst man das Problem?
 - ▶ einen Schluss: Zusammenfassung; was für Probleme bleiben noch offen?
- Dabei darf man durchaus einen Punkt in verschiedenen Teilen wiederholen.

Geschichte

- Beim Folienmalen immer Geschichte vor Augen haben.
- Zentrale Funktion: Fokussieren.
 - ▶ Man stößt oft auf die Frage “Soll ich über Thema X reden oder nicht?”
 - ▶ Antwort: Rede über X, wenn man es braucht, um die Geschichte zu erzählen. Sonst nicht.
- Ein Artikel kann u.U. mehrere mögliche Geschichten haben => die richtige auswählen!

Geschichte

- Jeder gute Vortrag erzählt eine Geschichte.
 - ▶ Was sind die wirklich wichtigen Punkte?
 - ▶ Was ging vorher nicht, das jetzt geht?
 - ▶ Warum ist es schwierig?
 - ▶ Wie geht es?
- Der neue Beitrag des Artikels muss unbedingt klar werden.

Der Zuhörer

- Es ist wichtig, dass die Zuhörer, die wirklich im Publikum sitzen, die Geschichte verstehen.
- Was wissen diese Zuhörer, was nicht?
Unbekanntes einführen, an Bekanntes erinnern.
- Zuhörer nicht überschätzen!
- Bezug zu früheren Themen im Seminar? Gerne alte Punkte aufgreifen und eigenen Artikel damit vergleichen.

Motivation = Drama

- Zuhörer muss verstehen, warum er Euch 45 Minuten lang zuhören soll.
- Daher deutlich machen:
 - ▶ Was war vorher schlecht an der Welt?
 - ▶ Warum ist es schwierig, es besser zu machen?
 - ▶ Wir haben es besser gemacht! Juhu!
- Andreas Zeller: Vorträge sind Märchen.

Vom Artikel zum Vortrag

- Ihr müsst selbst verstehen, was im Papier steht.
 - ▶ ggf. mehrfach lesen
 - ▶ Hintergrundliteratur suchen, lesen, ggf. einbringen
 - ▶ wichtige Unklarheiten mit mir diskutieren
- Verständnis = mit eigenen Worten erklären können
 - ▶ genau das macht man dann im Vortrag

Vom Artikel zum Vortrag

- Beispiele sind nützlich.
 - ▶ linguistische Phänomene: Textbeispiele
 - ▶ Algorithmen: Ablauf am Beispiel zeigen
 - ▶ Bilder!
 - ▶ NB: Papier und Vortrag sind sehr verschiedene Medien, Darstellung muss daran angepasst werden.
- Einem Freiwilligen probewhalten.
- Bei Problemen: Holt Euch bei mir Hilfe.

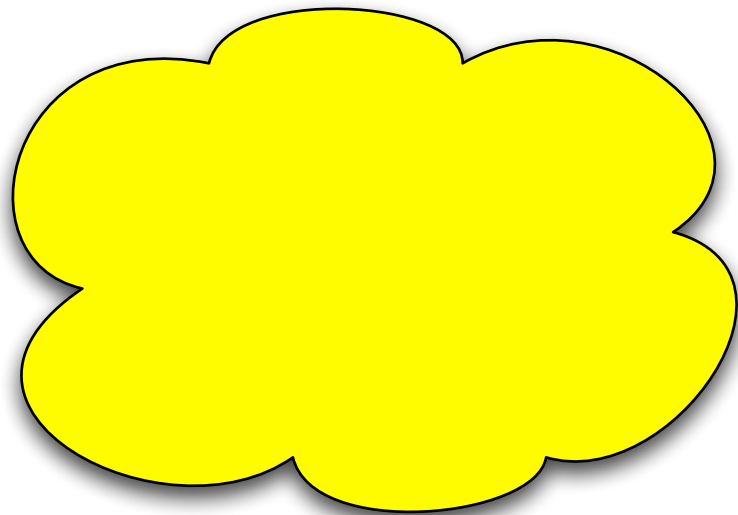
Gestaltung

- Zweck einer Folie:
 - ▶ Verständnis des Vortrags leichter machen.
 - ▶ Sekundär: Spickzettel für Vortragenden.
- Achtet auf ordentliche und einheitliche Gestaltung.
- Folien nicht zu voll machen.
- Schriftgröße: *mindestens* 24 Punkte.

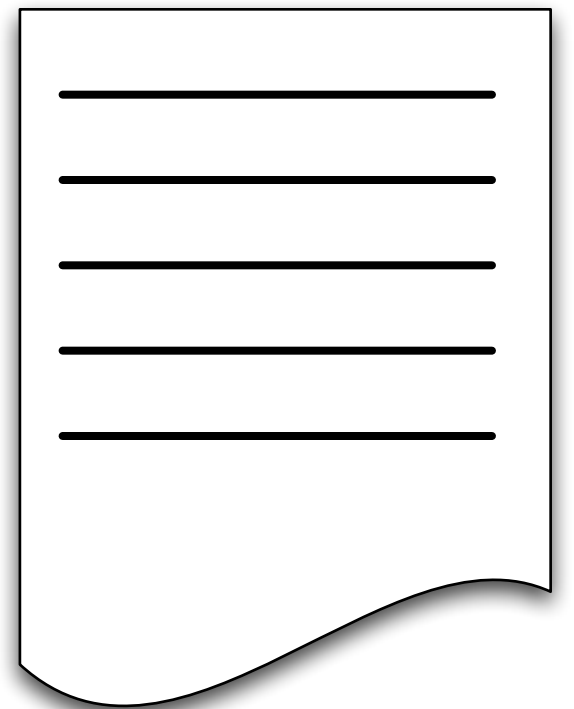
Körpersprache

- Kontakt zum Publikum!
- Raum einnehmen.
- Leute anschauen.

Was ist Generierung?



irgendwas



Text

Beispiel: Wetterberichte

Input Data			Output Text
Hour of the day	Wind Direction	Wind Speed	
0	S	10.0	
3	S	12.0	
6	N	15.0	
9	S	20.0	
12	S	15.0	
15	S	13.0	
18	S	12.0	
21	S	10.0	
24	S	10.0	

Generate >>>

S 8-13 backing N 13-18 by early morning, then veering S 18-23 by morning, and easing 8-13 by midnight.

<http://www.csd.abdn.ac.uk/research/sumtime/>

Reiter et al. (2005)

Beispiel: Museum



“This exhibit is a lekythos, created during the archaic period. It dates from circa 500 BC. It was painted by Amasis with the red figure technique and it originates from Attica.”

“... Unlike the previous exhibit, it originates from Attica.”

“Questo reperto è una lekythos.”

“Αυτο' το ε'κθεμα ει'ναι μια λη'κυθος.”

Beispiel: Zusammenfassen

Home » Business » iPad

Apple profit surges 70%, again blowing away expectations

Sales of iPhones leap 91% from a year earlier, accounting for almost half of Apple's record overall sales in the period and pushing the company past BlackBerry's maker in the rankings of smart-phone manufacturers.



Apple sold 14.1 million iPhones and 4.19 million iPads in the quarter.

By David Chidley
October 18, 2010 | 1:59 pm

Apple Inc. trounced analyst expectations yet again Monday while continuing its string of record quarters.

Riding booming sales of its blockbuster iPhone 4 -- the company said it sold 14.1 million smart phone during the new model's debut quarter -- Apple reported an all-time high in sales. That handily beat Wall Street analysts' expectations of \$19.3 billion.

Sales of the new iPhone rose more than 90% over the previous quarter, an increase that mute concerns that the company might have lost momentum after the phone stumbled with its antenna shortly after its release.

"We are blown away to report over \$20 billion in revenue and over \$4 billion in after-tax profit, both all-time records for Apple," said Apple CEO Steve Jobs, adding that, "We still have a lot of surprises left for the remainder of this calendar year."

Apple also sold 4.19 million iPads, a nearly 30% increase over the 3.3 million the tablet device's debut quarter.

Shares of Apple hit an all-time high of \$319 a share, representing a 50% gain since January.

Related: Chart: iPhone sales

Topics: New Products, Apple iPhone, Earnings

See more topics »

ADS BY GOOGLE

Hot Stock Alert - EHSI
Profit from healthcare explosion.
Millionaires Created Today
EmergingTech.com/stockalerts.com

Free Stock Advice
Become an informed investor with our
Free Monthly Shares Newsletter

Balsillie fires back at Apple's Jobs

Last Updated: Tuesday, October 19, 2010 | 6:27 PM ET | Comments 99 | Recommend 105
CBC News

Waterloo, Ont.-based BlackBerry maker Research In Motion fired back Tuesday in a war of words with Apple Inc. Co-CEO Jim Balsillie sliced apart Apple's criticisms of his company, issuing a lengthy statement aimed at Apple CEO Steve Jobs.

October 19, 2010 6:50 PM

RIM: Jobs Telling Half the Truth

2 comments

Posted by Erica Ogg

Steve Jobs certainly knows how to rile up his competitors.

His rant yesterday about the fragmentation of Android, Google's obfuscation and the shortcomings of 7-inch touch-screen tablets has earned him responses. Android posted a passive-aggressive tweet on the subject, and Android developer TweetDeck made a more direct correction of Chief Executive Steve Jobs' claims about its development process.

Research In Motion's co-CEO also has a few things to say. He's clearly decided to show how the iPhone is leaving the BlackBerry in its dust go unchallenged.

7 Key Turning Points That Made Apple No. 1

By Brian X. Chen | May 26, 2010 | 7:54 pm | Categories: Accessories and Peripherals, Miscellaneous



Steve Jobs and Jim Balsillie

Apple climbed to an all-time high in the Cupertino, California, market, at least.

computer market.

Apple profit: Apple fourth-quarter profit surges 70%, again exceeding expectations

Propelled by huge demand for the latest iPhone model, Apple Inc. on Monday posted tremendous increases in earnings and sales as it overtook BlackBerry's maker in the rankings of smart-phone manufacturers. ([article 1](#)) The tech juggernaut once again trounced analysts' expectations for earnings as well as sales, but its stock slumped in after-hours trading because of an unexpected drop in a key measure of profitability. ([article 1](#)) Waterloo, Ont.-based BlackBerry maker Research in Motion fired back Tuesday in a war of words with Apple Inc. Co-CEO Jim Balsillie sliced apart Apple's criticisms of his company, issuing a lengthy statement aimed at Apple CEO Steve Jobs. ([article 3](#)) (Dave Chidley/Canadian Press) A day earlier, Jobs bragged that Apple sold more iPhones than RIM sold BlackBerrys in the most recent financial quarter. ([article 3](#)) The head of Android posted a passive-aggressive tweet on the subject, and Android developer TweetDeck made a more direct correction of Chief Executive Steve Jobs' claims about its development process. ([article 4](#)) Earlier in the day, the stock hit an all-time high of \$319 a share, representing a 50% gain since January. ([article 2](#)) Apple surpassed longtime rival Microsoft in market capitalization, making the Cupertino, California, company the most valuable technology company in the world, for the moment, at least. ([article 5](#))

Ein einführendes Beispiel

Umuz Ono, "The Attack of Hatchets"

A: 361, mostly goblins, 131 losses

D: 447 elves, 120 losses

Defender was victorious.

In 192, the human Ozud Oldstolen's left eye was torn out by the elf Avafi Matchedglossed the Assaulted Slaughters.

In 192, the elf Avafi Matchedglossed the Assaulted Slaughters's left foot was ripped by the human Ozud Oldstolen.

In 192, the elf Masami Glossrams's left upper arm was stabbed by the elf Arstruk Rabblevile.

In 192, the goblin Stâsost Flysoots was shot and killed by the elf Arane Pageford the Mighty Romance in Beancyclones.

In 192, the goblin Strodno Feverdoomed was shot and killed by the elf Gica Ownmusic the Enjoyable Sun of Venerating in Beancyclones.

In 192, the elf Ula Snakesperplexes was struck down by

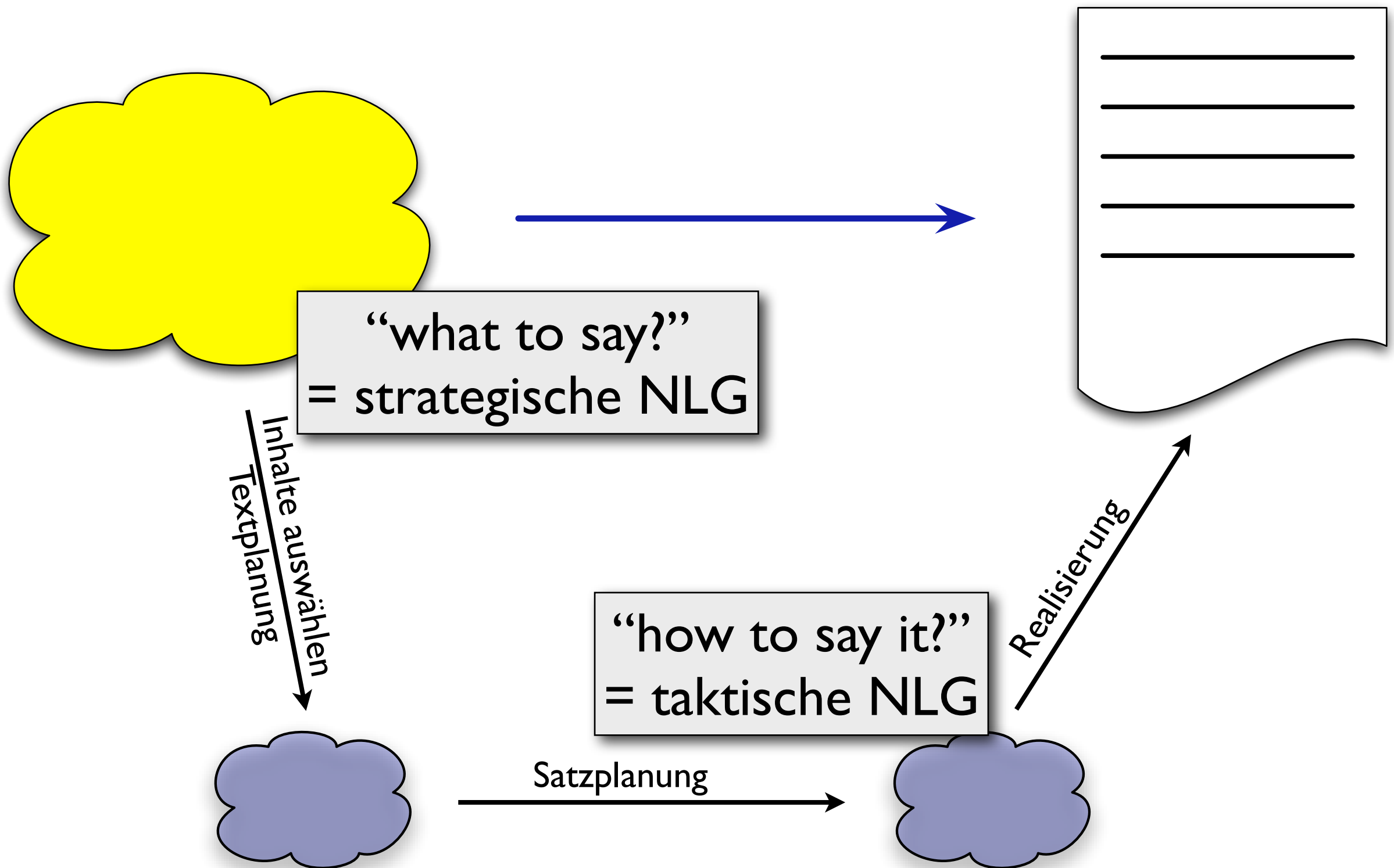


ESC: done.

Shift + ESC: back to age.

("Dwarf Fortress", generated in-game text, see <http://www.bay12games.com/dwarves/>)

Struktur eines NLG-Systems



Content determination

- Content selection:
 - ▶ what objects do we want to talk about?
 - ▶ what do we want to say about them?
 - ▶ if necessary, interpret domain data
- Discourse planning:
 - ▶ how to arrange information into a discourse?
 - ▶ how to break information up into sentence-sized chunks?
- Both are typically domain-specific.

Beispiel: Dialogsysteme

S: *I found 28 round trips from Edinburgh to Brussels on October 31st and returning Edinburgh on November 6th. For option 1, I have an outbound flight with KLM departing at 7 am and arriving at 8:49 am non-stop. A return flight is with KLM departing at 10:30 am and arriving at 11:28 am with 1 stopover. Ticket price is £165. Please say "next option", or, "flight details", or, "I'll take it".*

U: Next option

S: *For option 2, I have an outbound flight with KLM departing at 3:30 pm and arriving at 5:19 pm nonstop. A return flight is with KLM departing at 8:25 pm and arriving at 12:28 pm with 1 stopover. Ticket price is £165 please say "next option", or, "flight details", or, "I'll take it."*

U: Next option

S: ...

Naives System

U: Hello. I'd like to book a flight from Edinburgh to Brussels for tomorrow arriving at around 3pm.

S: *There are four flights with availability in business class. None are on KLM. The only direct flight arrives at 5:30 pm, which is later than you requested. To arrive earlier, you'll have to make a connection.*

If you're willing to travel economy, there is a direct flight on BMI, arriving at 1:10 pm. There's also a KLM flight arriving at 1:50 pm, but it requires a connection in Amsterdam. All other economy flights take longer or arrive after 4 pm.

U: Can you tell me more about the business class flights that arrive earlier?

S: *The two flights with shortest travel time are both on British Airways and cost 308 pounds. The first takes 4 hours and 20 minutes and arrives at 1:20 pm. The other takes 4 hours and 30 minutes and arrives at 3:30pm.*

There's a cheaper flight on Lufthansa. It arrives at 3:10pm and costs just 270 pounds, but it takes 5 hours and 35 minutes. All other business class flights arrive very early or too late.

Demberg & Moore (2006)

Content selection

	depart	arrive	airline	business	direct	price
1	9:00	13:20	BA	+	-	308
2	9:27	13:50	KLM	-	-	200
3	10:00	15:10	Lufthansa	+	-	270
4	11:00	15:30	BA	+	-	308
5	11:10	13:10	BMI	-	+	150
6	15:30	17:19	KLM	-	+	165
7	15:40	17:30	BMI	+	+	350

Demberg & Moore (2006)

User model: + direct, + business, + KLM

Input: arrive around 3pm

Discourse planning

1. Flights in business class: 4

a) KLM: none

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

- Take specification of each sentence and translate it into the actual sentence in the output language.
- Input:
 - ▶ specification of sentence (e.g., semantic representation)
 - ▶ a grammar or something similar
- Output:
 - ▶ one sentence

Surface realization

1. Flights in business class: 4

a) KLM: none

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, *but* -direct

Template-based realization

Hard-coded rule:

“flights in business class: X”

→ “There are X flights in business class”.

Flights in business
class: 4



“There are four flights
in business class.”

Surface realization

1. There are four flights with availability in business class.

a) KLM: none

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.

a) None are on KLM.

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.

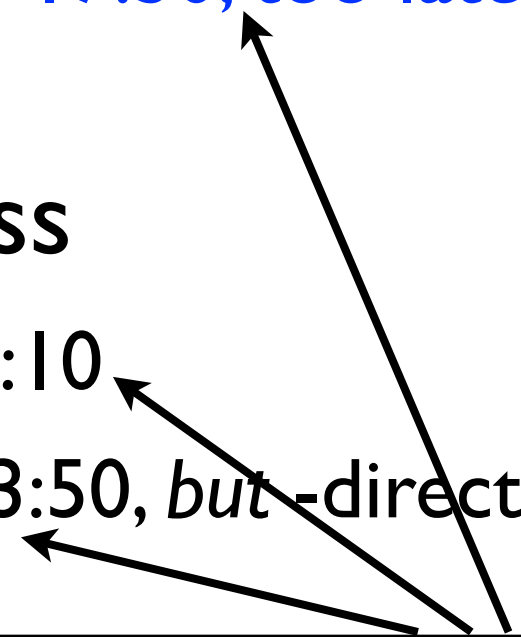
a) None are on KLM.

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, but -direct



don't want separate template
for each combination

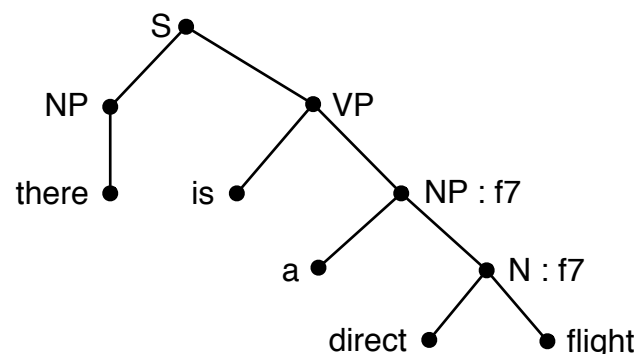
The diagram consists of three arrows originating from a yellow rectangular box at the bottom right. One arrow points to the phrase 'too late' in item 1b. Another arrow points to the phrase 'arrives 13:10' in item 2a. The third arrow points to the phrase 'arrives 13:50' in item 2b.

Surface realization with TAG

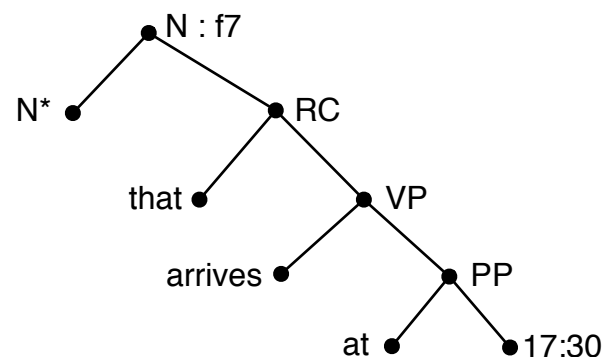
semantics

syntax

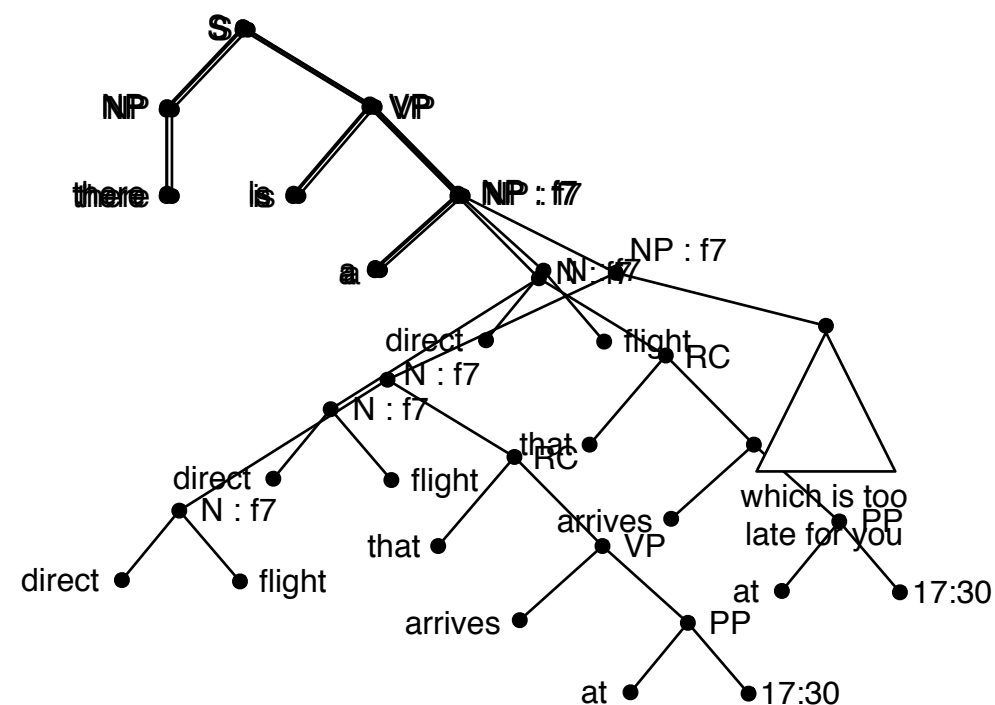
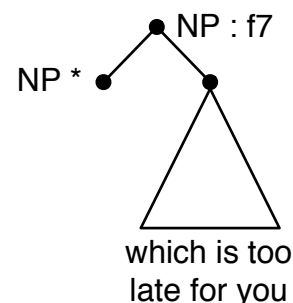
direct(f7)



arrives(f7, 17:30)



too_late(f7)



“There is a direct flight that arrives at 17:30, which is too late for you.”

Surface realization

1. There are four flights with availability in business class.

a) None are on KLM.

b) describe f7: +direct, arrives 17:30, too late

2. Flights in economy class

a) describe f5: BMI, arrives 13:10

b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.
2. Flights in economy class
 - a) describe f5: BMI, arrives 13:10
 - b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.
2. If you're willing to travel economy,
 - a) describe f5: BMI, arrives 13:10
 - b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.
2. If you're willing to travel economy,
 - a) There is a direct flight on BMI, arriving at 13:10.
 - b) describe f2: KLM, arrives 13:50, *but* -direct

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.
2. If you're willing to travel economy,
 - a) There is a direct flight on BMI, arriving at 13:10.
 - b) There's also a KLM flight arriving at 13:50, but it requires a connection in Amsterdam.

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.

2. If you're willing to travel economy,
 - a) There is a direct flight on BMI, arriving at 13:10.
 - b) There's also a KLM flight arriving at 13:50, but it requires a connect

**Merge these two sentences:
Aggregation.**

Surface realization

1. There are four flights with availability in business class.
 - a) None are on KLM.
 - b) There is a direct flight that arrives at 17:30, which is too late for you.
2. If you're willing to travel economy,
 - a) there is a direct flight on BMI, arriving at 13:10.
 - b) There's also a KLM flight arriving at 13:50, but it requires a connection in Amsterdam.

Ein einführendes Beispiel

Umuz Ono, "The Attack of Hatchets"

A: 361, mostly goblins, 131 losses

D: 447 elves, 120 losses

Defender was victorious.

In 192, the human Ozud Oldstolen's left eye was torn out by the elf Avafi Matchedglossed the Assaulted Slaughters.

In 192, the elf Avafi Matchedglossed the Assaulted Slaughters's left foot was ripped by the human Ozud Oldstolen.

In 192, the elf Masami Glossrams's left upper arm was stabbed by the elf Arstruk Rabblevile.

In 192, the goblin Stâsost Flysoots was shot and killed by the elf Arane Pageford the Mighty Romance in Beancyclones.

In 192, the goblin Strodno Feverdoomed was shot and killed by the elf Gica Ownmusic the Enjoyable Sun of Venerating in Beancyclones.

In 192, the elf Ula Snakesperplexes was struck down by

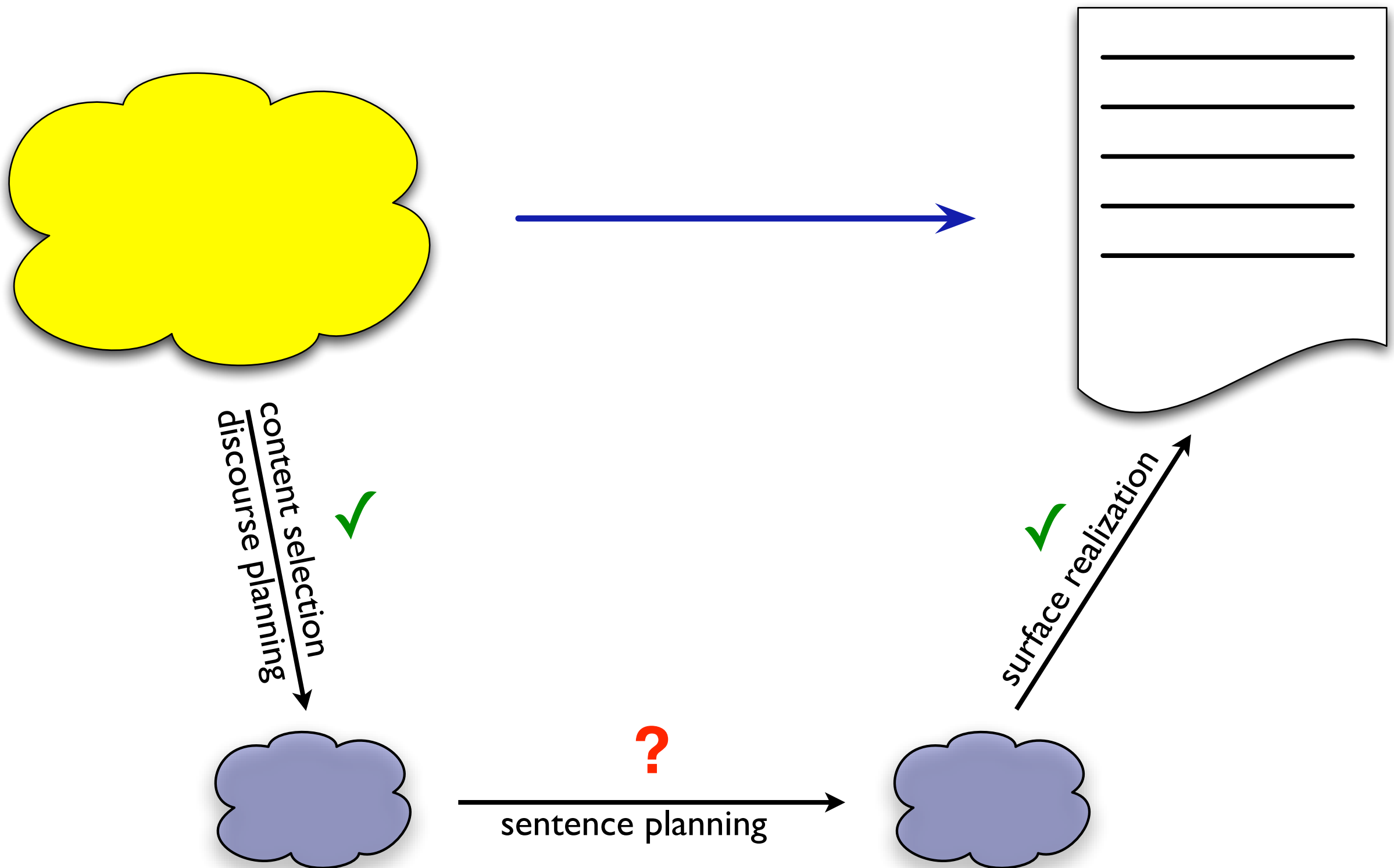


ESC: done.

Shift + ESC: back to age.

("Dwarf Fortress", generated in-game text, see <http://www.bay12games.com/dwarves/>)

Where are we now?



Sentence planning

- Output of content determination may be not quite suitable as input of surface realizer.
- Sentence planning: Everything that needs to happen to map CD output to SR input.
- Typically:
 - ▶ referring expression generation
 - ▶ lexical choice
 - ▶ etc.

Referring expressions

Knowledge base:



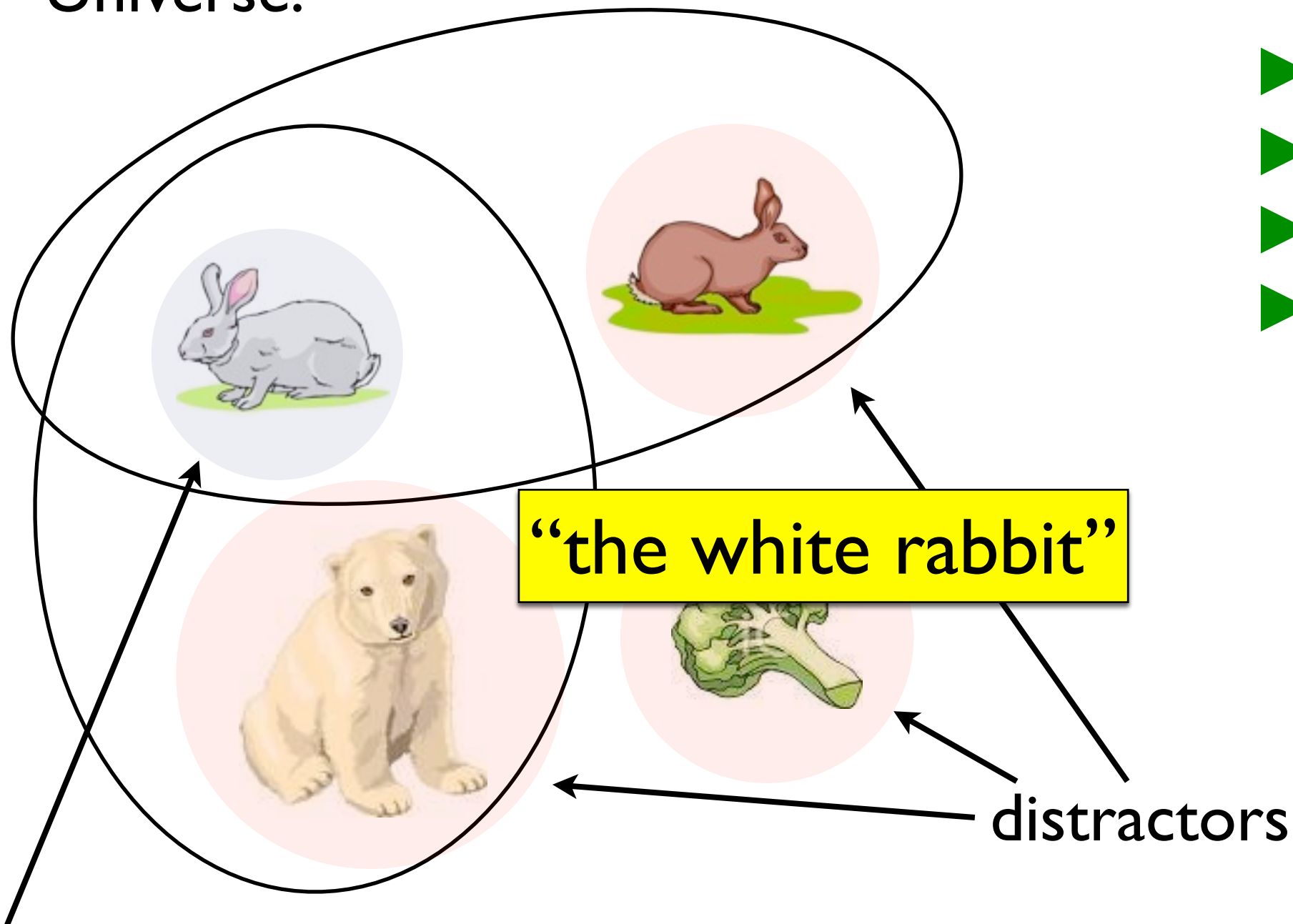
We want to say that
this guy sleeps.

“The white rabbit sleeps.”

Is this content determination? Surface realization?

RE generation

Universe:



List of properties:

- ▶ - rabbit ✓
- ▶ - polar bear
- ▶ - broccoli
- ▶ - white ✓
-

target referent

Incremental algorithm: Dale & Reiter (1995)

RE generation

- Until recently, very active research area.
- 1990s, early 2000s: algorithms for more expressive REs, dominated by logicians
- Recently: focus shift towards cognitive models of “good” REs

Lexical choice

- Lexical choice: Mapping semantic concepts to content words.
- Not as trivial as it sounds at first glance.
- How to realize generic concepts based on what they apply to?
 - ▶ The temperature **rose**.
 - ▶ The rain got **heavier**.
 - ▶ The revenue **increased**.

Lexical choice

- How to map real-world values to words?
 - ▶ what RGB values accepted as “red car”?
 - ▶ as “red wine”?
 - ▶ as “red hair”?
- How to distribute bits of meaning over different words?
 - ▶ swim across the lake
 - ▶ traverser le lac à la nage

Zusammenfassung

- Gute Vorträge:
 - ▶ Geschichten erzählen
 - ▶ An Zuhörer anpassen
- Überblick Generierung:
 - ▶ “what to say” = content determination
 - ▶ “how to say it” = Textplanung + Satzplanung + Realisierung
 - ▶ Spezielle Probleme: Referierende Ausdrücke, lexical choice, Aggregation