

Proseminar „Generierung“

Textplanen mit Schemata

Susanne Fertmann

5. November 2010

Textgenerierung

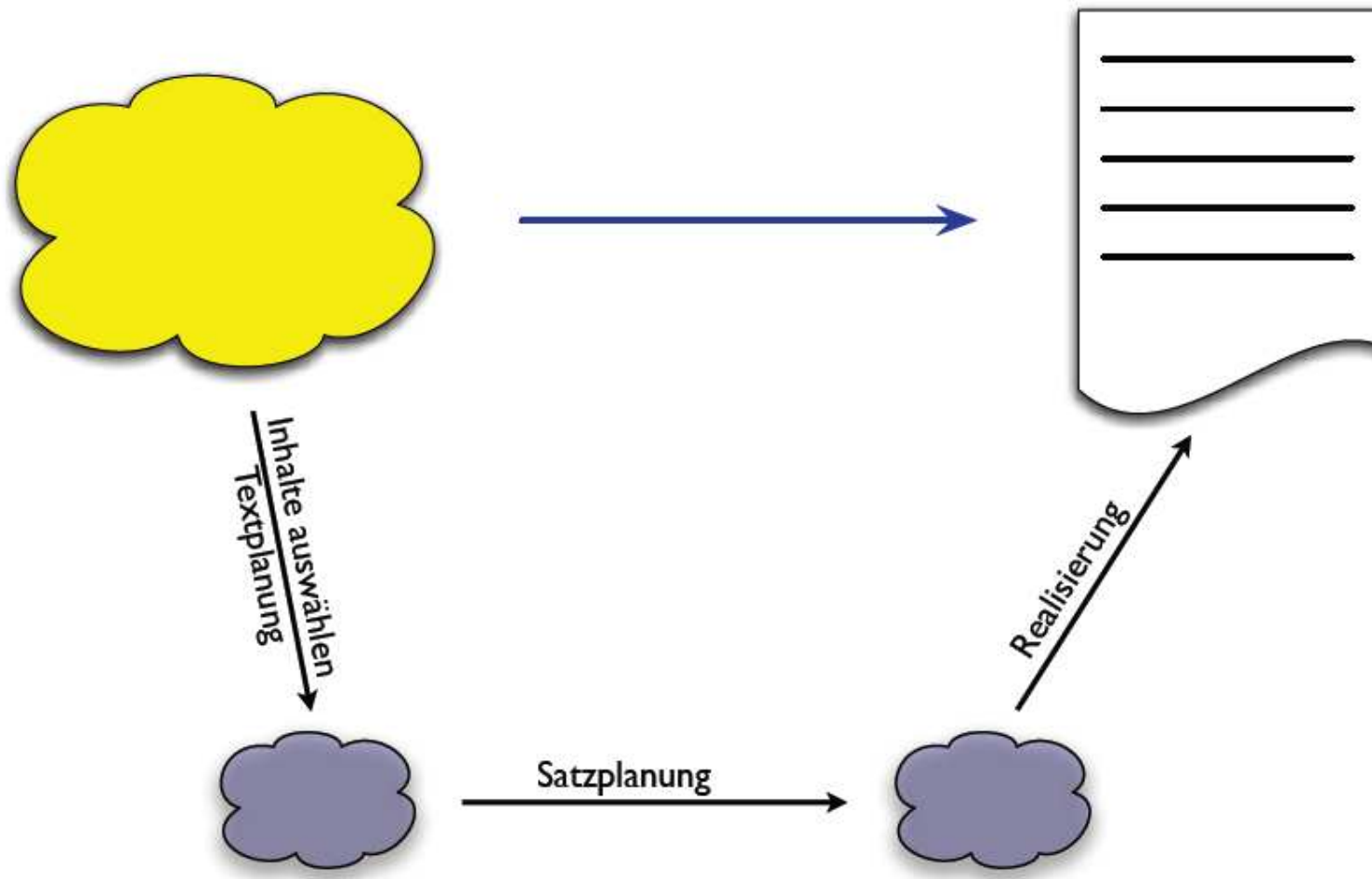
- Wir müssen entscheiden:
 1. Was wollen wir sagen? → Inhalt
 2. Wie wollen wir es sagen? → Struktur

Überblick

- Probleme beim Textplanen
- Diskursstruktur
- Schemata (McKeown, TEXT)
- Fokus
- Beispiel

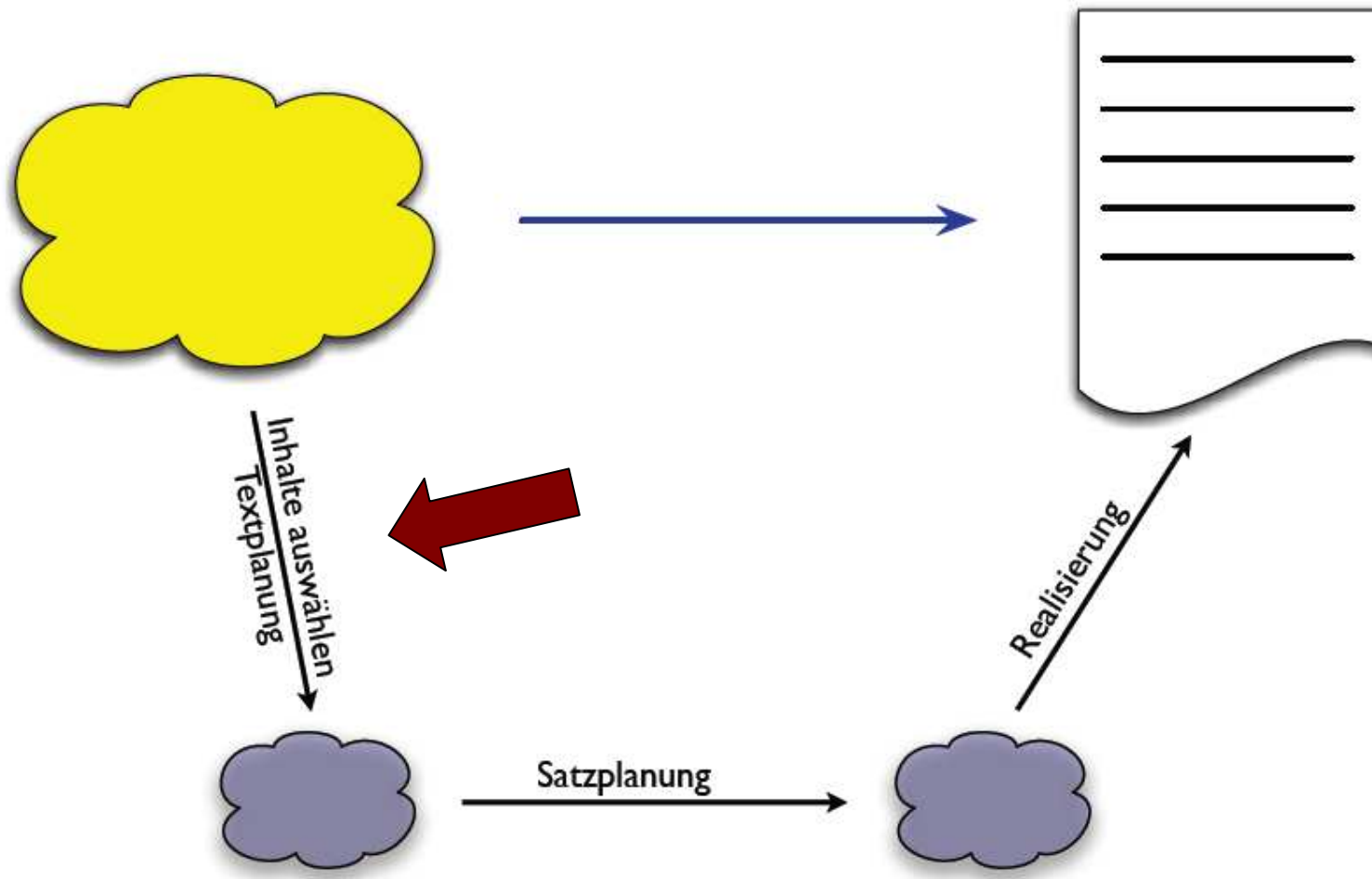
Zur Erinnerung...

Struktur eines NLG-Systems



Zur Erinnerung...

Struktur eines NLG-Systems



Probleme bei Textgenerierung

Ich habe mir ein Buch gekauft. Vorgestern war ich im Kino. Liest du auch gerne Krimis?

Probleme bei Textgenerierung

Ich habe mir ein Buch gekauft. Vorgestern war ich im Kino. Liest du auch gerne Krimis?

→ **Kohärenz** (die Sätze sollen sich aufeinander beziehen)

→ **Diskursstruktur**

Diskursstruktur

„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) The vineyards make wines that are emphatically of the Rheingau style, (3) with a considerable weight for a white wine. (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

- Textstruktur
 - die zur Erreichung eines Diskurszieles verwendet wird
 - durch die der Inhalt und dessen Anordnung im Text festgelegt ist (www.enzyklo.de)

Rhetorische Prädikate

- Charakterisieren eine Äußerung und beschreiben ihre Beziehung zu anderen Aussagen
- Beispiele:
 - Amplification
 - This book is about violence, to the spirit as well as to the body.
 - Analogy
 - You make it exactly the same way as red-wine sangria, except that you use any of your inexpensive white wines instead of one of your inexpensive reds.

Rhetorische Prädikate

– Attributive

- Mary has a pink coat.

– Identification

- Eltville, an important wine village of the Rheingau region.

– Particular-illustration

- Tautenberg and Sonnenberg are among vineyards of note.

Analyse von Prädikaten

- McKeown untersucht kurze Informationstexte von 10 Autoren
 - Satzteile werden mit Prädikaten beschrieben
 - Ergebnis: ähnliche Prädikate
- 4 Prädikatmuster (=Schema)
- Identification Schema
 - Constituency Schema
 - Attributive Schema
 - Compare & Contrast Schema

Schemata

- Identification Schema (v.a. für Definitionen in Wörterbüchern und Enzyklopädien)

Identification (class&attribute/function)

{Analogy/Constituency/Attributive/Renaming/Amplification}*
Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

{Particular-illustration/Evidence}

Identification Schema

Identification (class&attribute/function)

{Analogy/Constituency/Attributive/Renaming/Amplification}*
Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

{Particular-illustration/Evidence}

„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) The vineyards make wines that are emphatically of the Rheingau style, (3) with a considerable weight for a white wine. (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

Identification Schema

Identification (class&attribute/function)

{Analogy/Constituency/Attributive/Renaming/Amplification}*
Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

{Particular-illustration/Evidence}

„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) The vineyards make wines that are emphatically of the Rheingau style, (3) with a considerable weight for a white wine. (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

1. Identification (class&attribute)

Identification Schema

Identification (class&attribute/function)

{Analogy/Constituency/**Attributive**/Renaming/Amplification}*
Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

{Particular-illustration/Evidence}

„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) **The vineyards make wines that are emphatically of the Rheingau style,** (3) with a considerable weight for a white wine. (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

1. Identification (class&attribute)

2. **Attributive**

Identification Schema

Identification (class&attribute/function)

{Analogy/Constituency/Attributive/Renaming/**Amplification**}*

Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

{Particular-illustration/Evidence}

„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) The vineyards make wines that are emphatically of the Rheingau style, (3) **with a considerable weight for a white wine.** (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

1. Identification (class&attribute)
2. Attributive
3. **Amplification**

Identification Schema

Identification (class&attribute/function)

{Analogy/Constituency/Attributive/Renaming/Amplification}*
Particular-illustration/Evidence+

{Amplification/Analogy/Attributive}

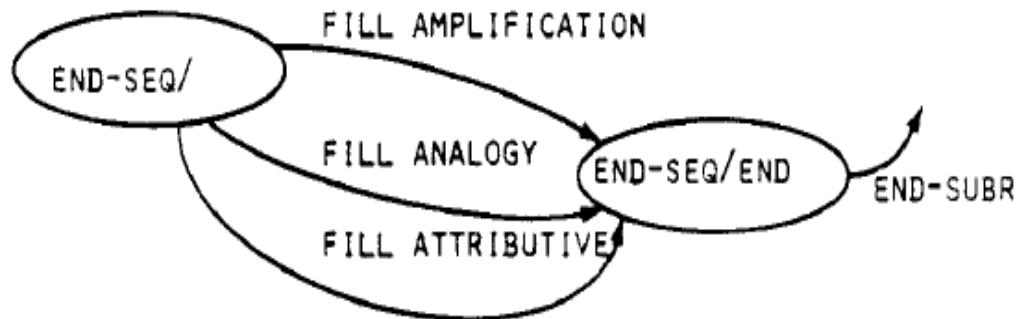
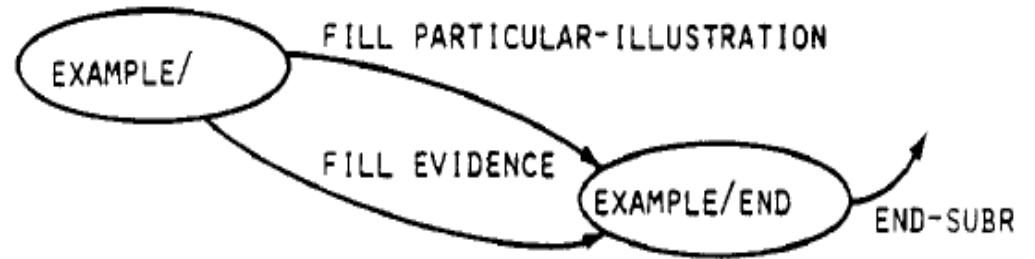
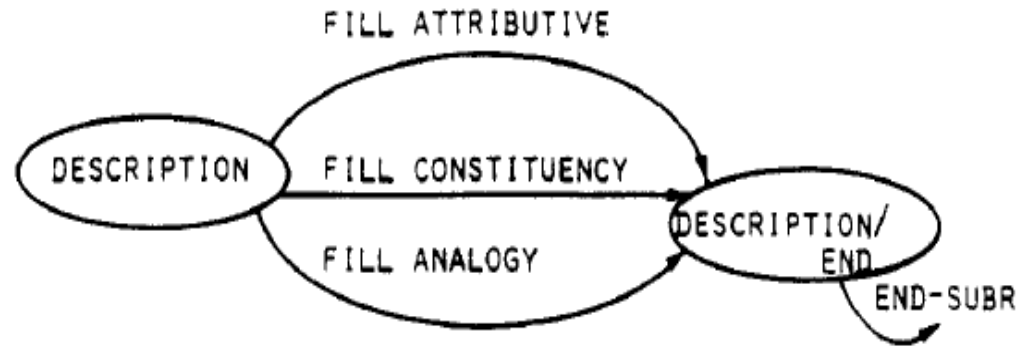
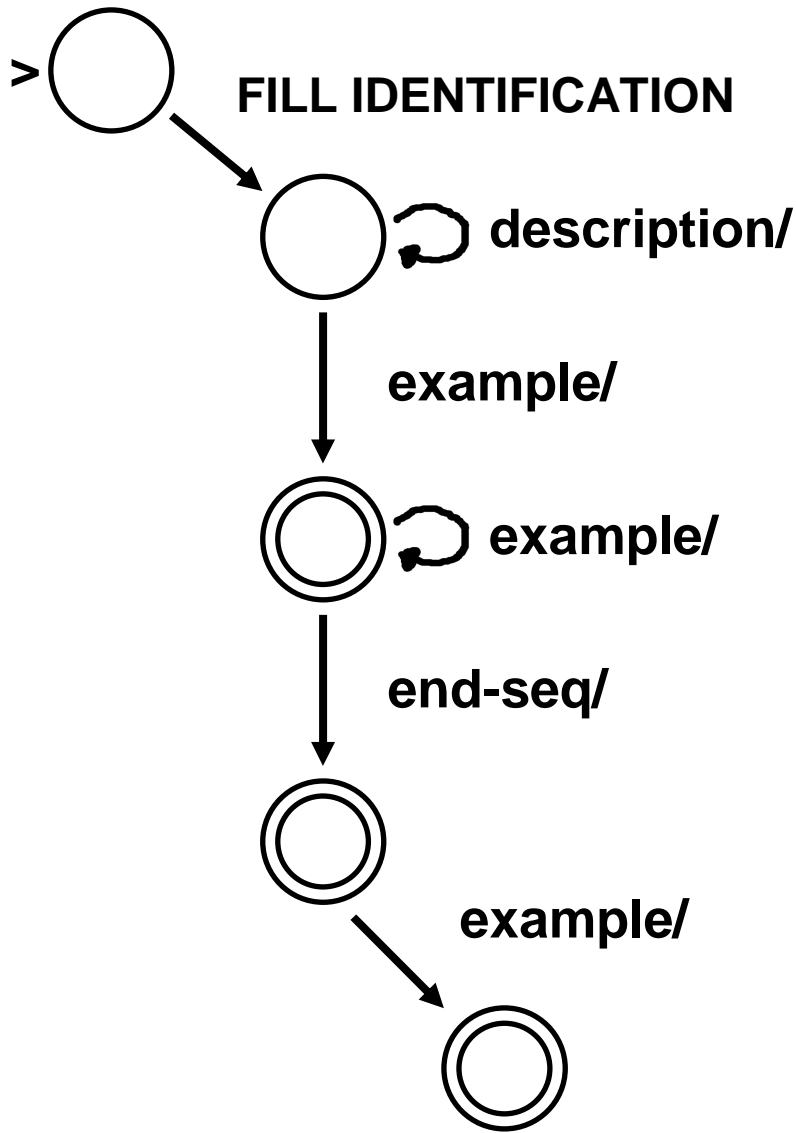
{Particular-illustration/Evidence}

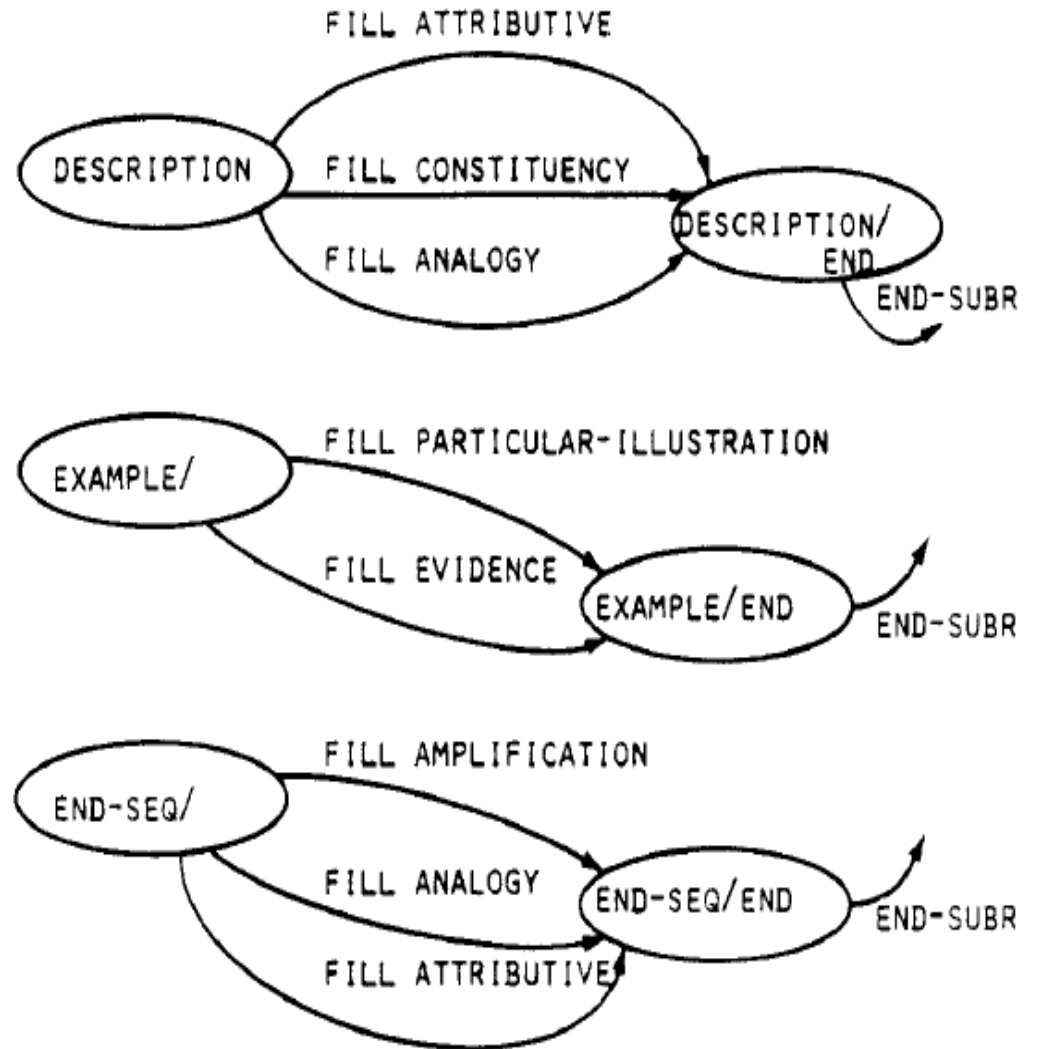
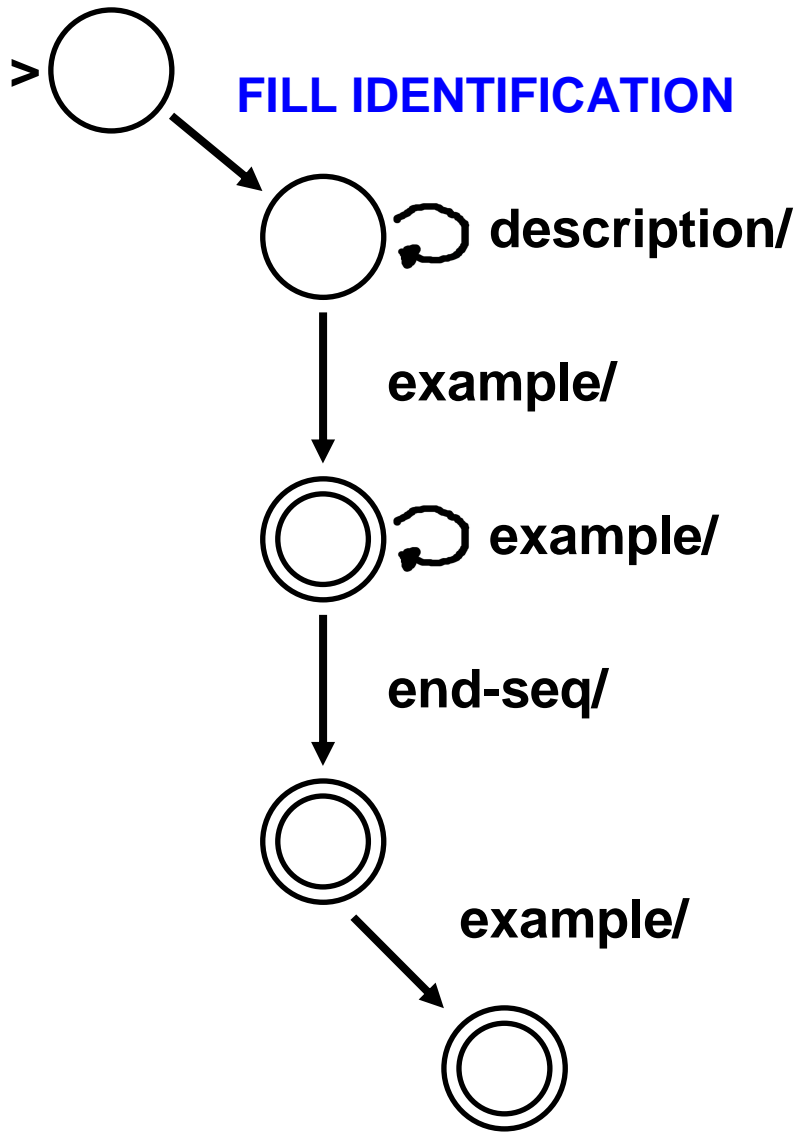
„Eltville (Germany) (1) An important wine village of the Rheingau region. (2) The vineyards make wines that are emphatically of the Rheingau style, (3) with a considerable weight for a white wine. (4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.“

1. Identification (class&attribute)
2. Attributive
3. Amplification
4. Particular-illustration

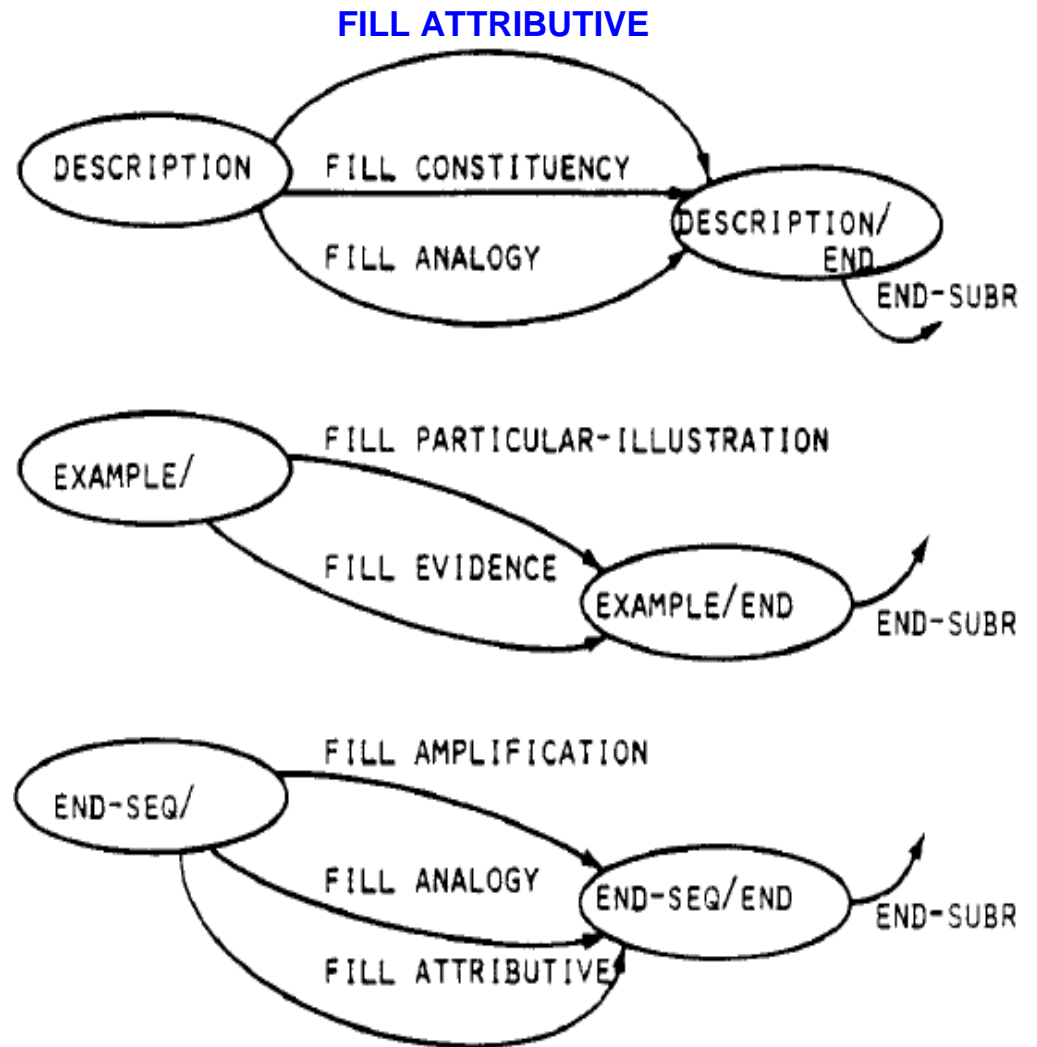
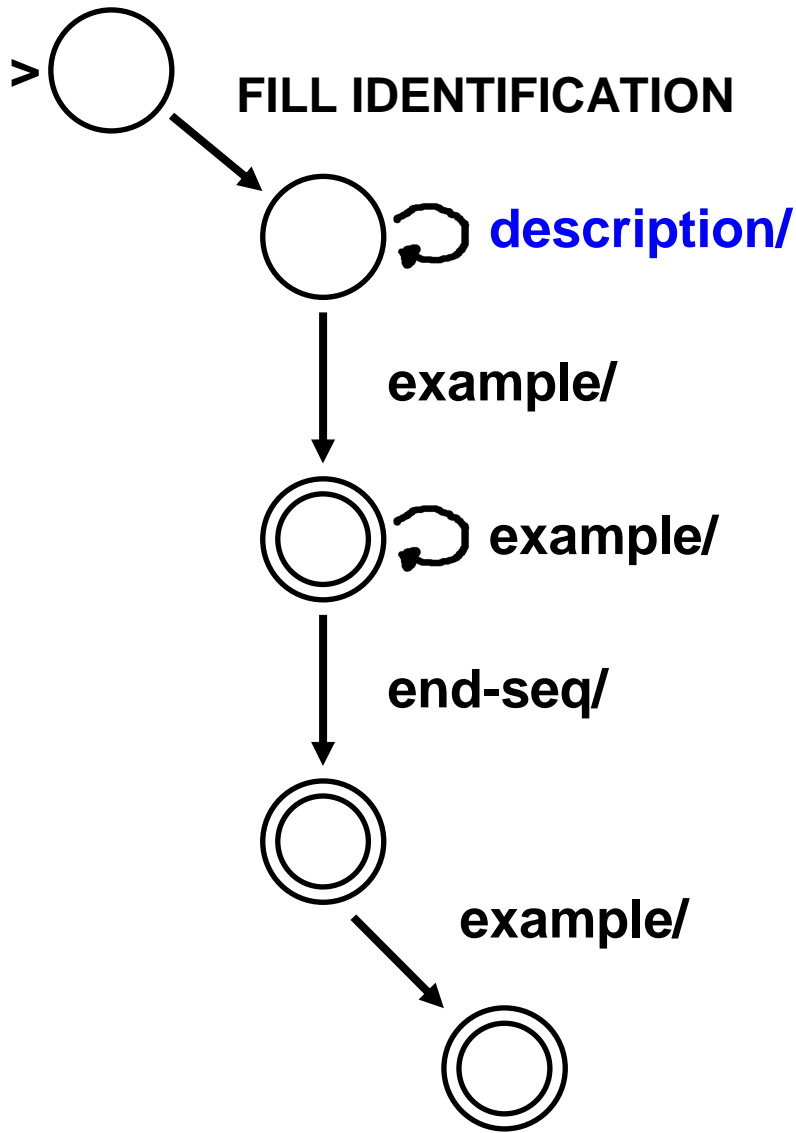
Schemata

- viele Alternativen
- Deskriptiv: zeigen normalerweise benutzte Mittel um Diskursziele zu erreichen
- Können vom Generierungssystem benutzt werden

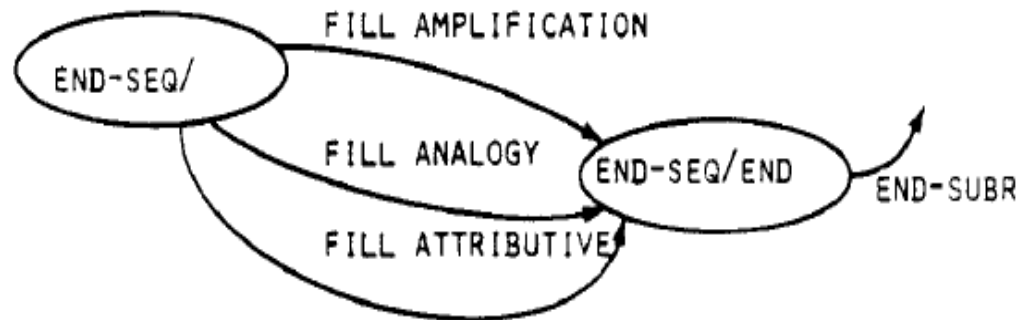
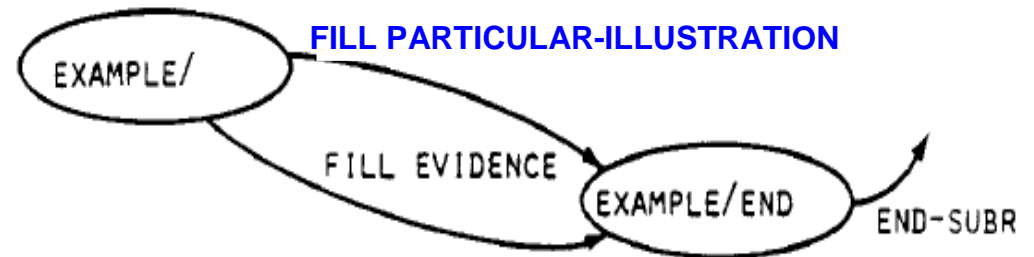
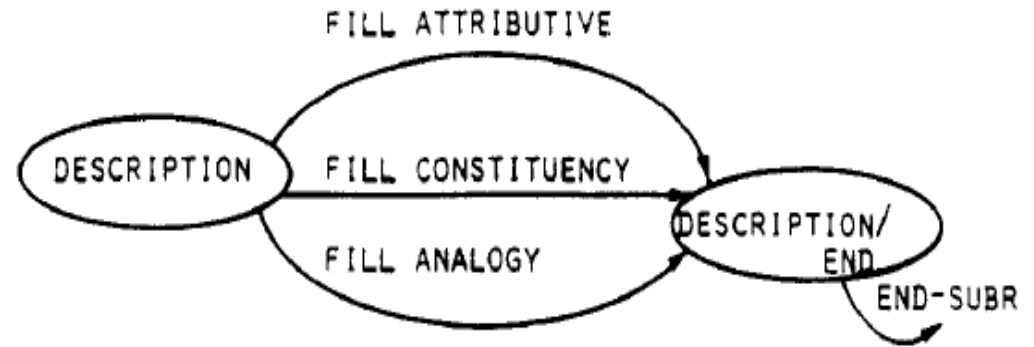
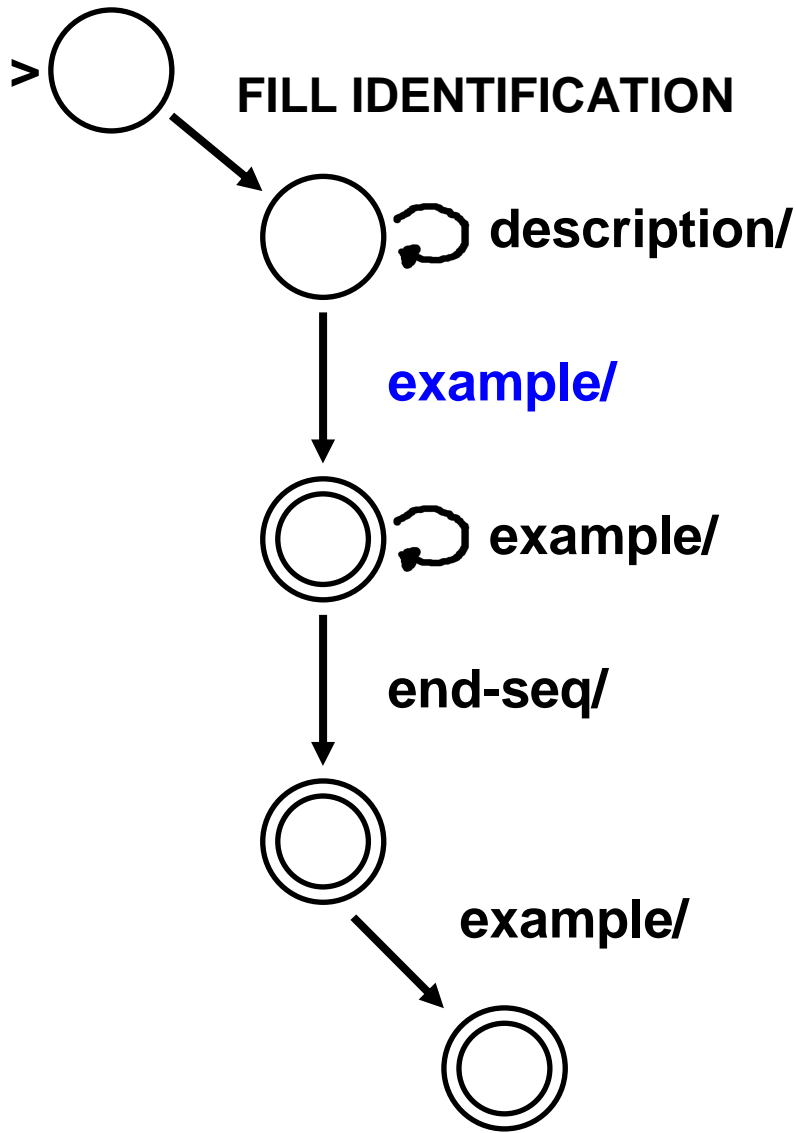




Flagey-Echezeaux (France) is an important wine township in the Cote de Nuits



Flagey-Echezeaux (France) is an important wine township in the Cote de Nuits whose vineyards produce a fine rich round wine.

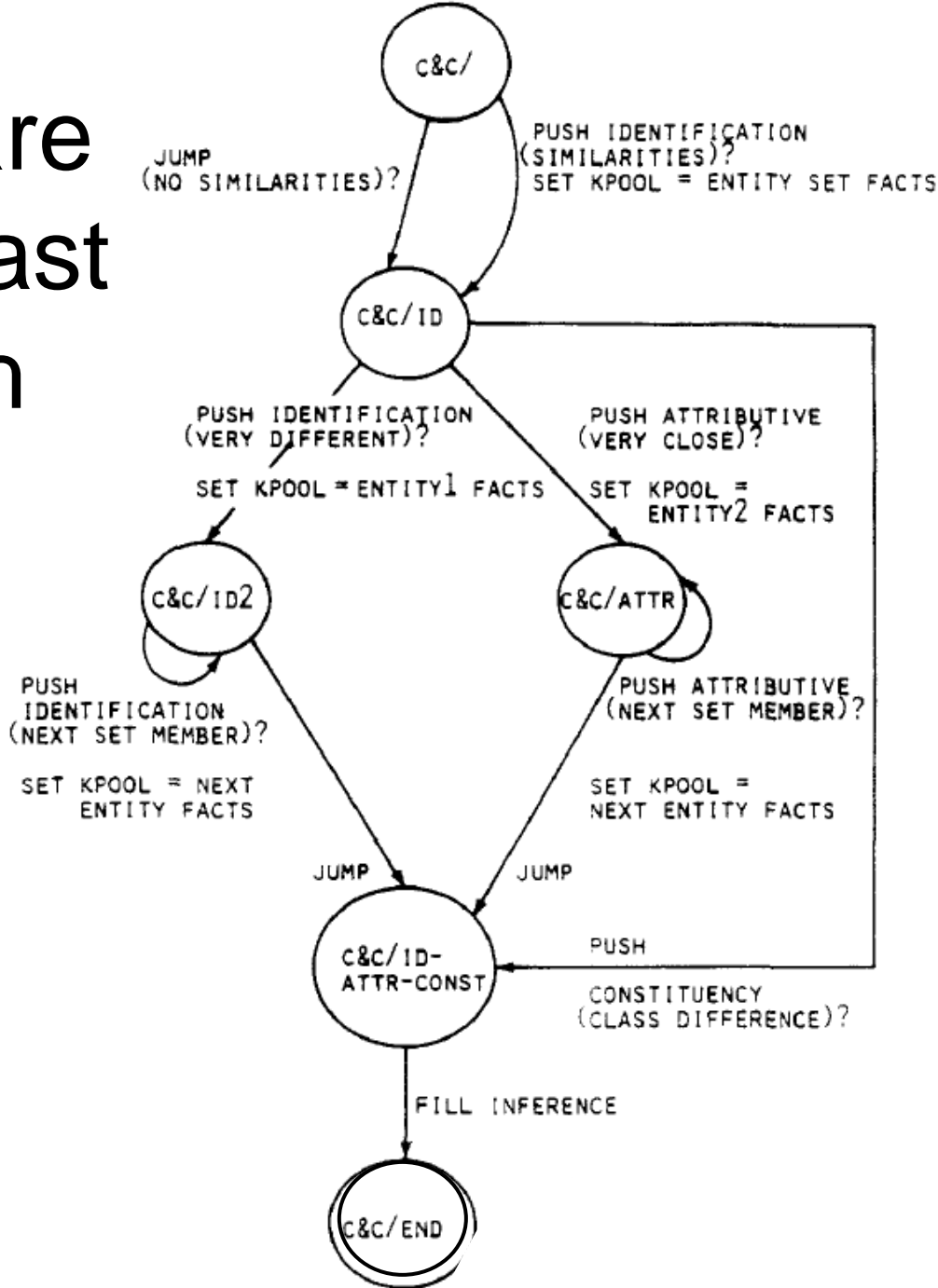


Flagey-Echezeaux (France) is an important wine township in the Cote de Nuits whose vineyards produce a fine rich round wine. **The most famous of the vineyards is Echezeaux.**

Schemata in TEXT

- Implementiert auf Basis von ATNs (augmented transition network)
 - Vorbedingungen
 - Aktionen vor & nach einem Übergang
 - Erlauben Rekursion

Compare & Contrast Graph



(1) A Hobie Cat is a brand of catamaran, (2) which is a kind of sailboat. (3) Catamarans have sails and a mast like other sailboats, (4) but they have two hulls instead of one. (5) That thing over there is a catamaran. (6) Hobie Cats have a canvas cockpit connecting the two pontoons and one or two sails. (7) The 16-ft. Hobie Cat has a main and a jib and the 14-ft. Hobie Cat has only a main.

Identification Schema

Identification Schema

Analogy Schema

Identification

Identification
Analogy

Similarities
Differences

Particular-illustration

Attribute

Particular-illustration

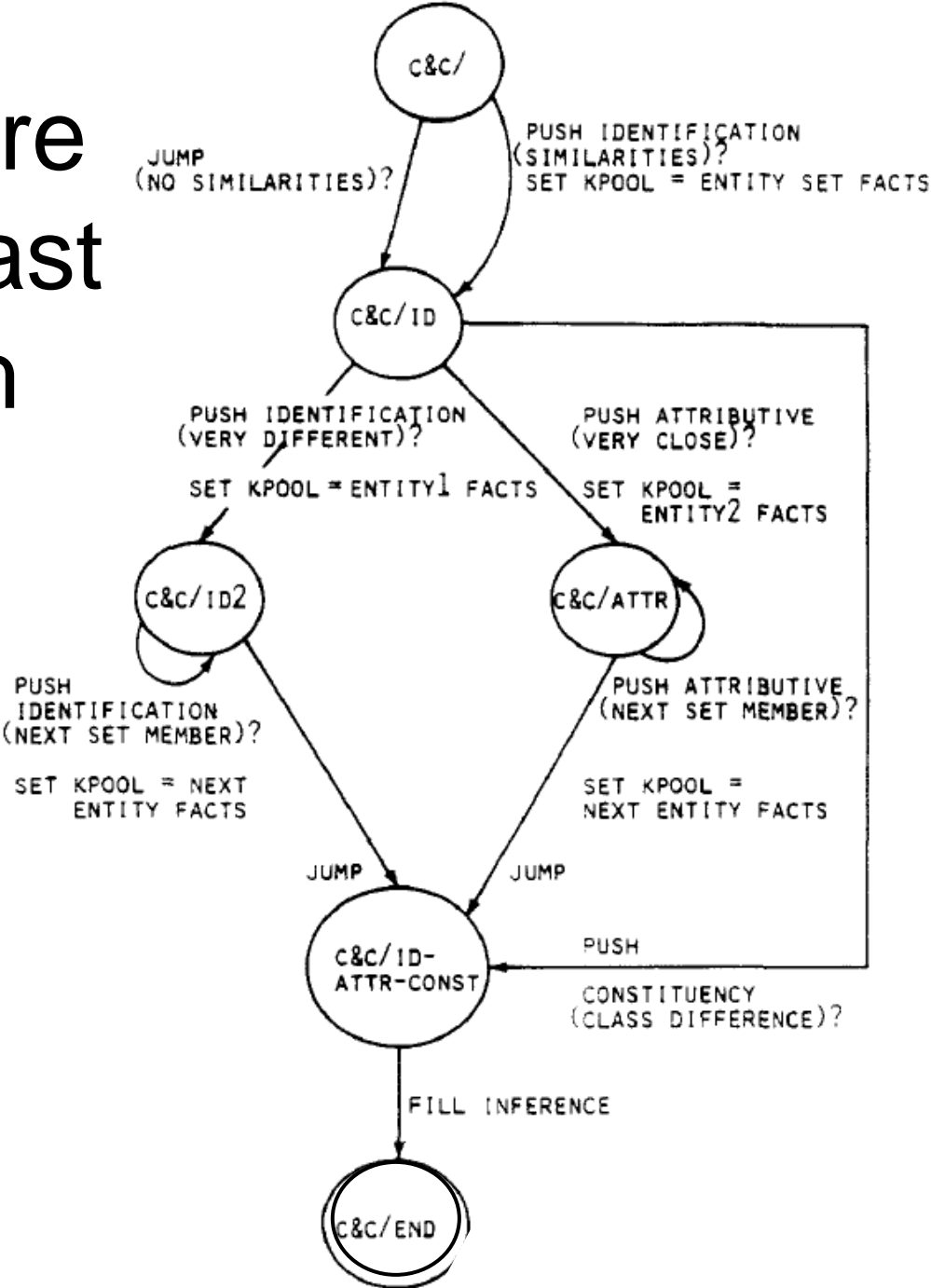
Schemarekursion

- Rhetorische Prädikate funktionieren rekursiv
- Jedes Prädikat kann zu einer Proposition oder einem Schema ausgeweitet werden

Schemarekursion

- In TEXT ist wenig Rekursion implementiert
- Frage: Wann ist Rekursion notwendig?
 - Abhängig vom Benutzer
 - Abhängig von der Menge an Informationen, die vorhanden ist und die gegeben werden soll
 - Frage nach Unterschied

Compare & Contrast Graph



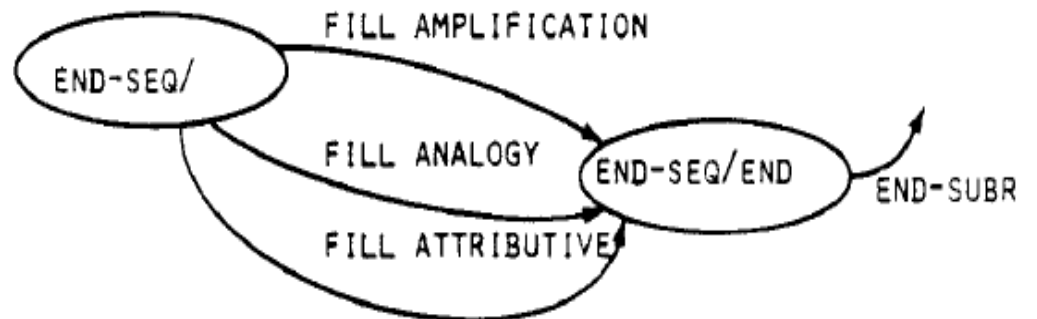
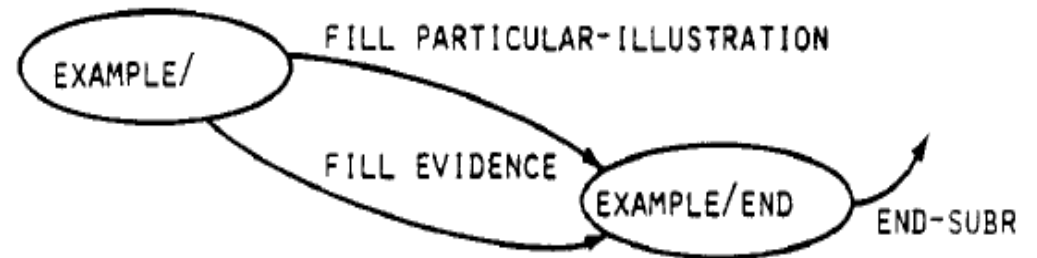
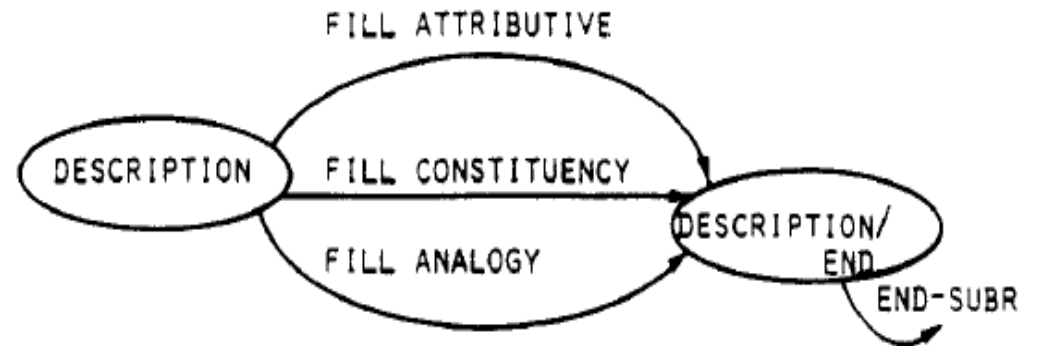
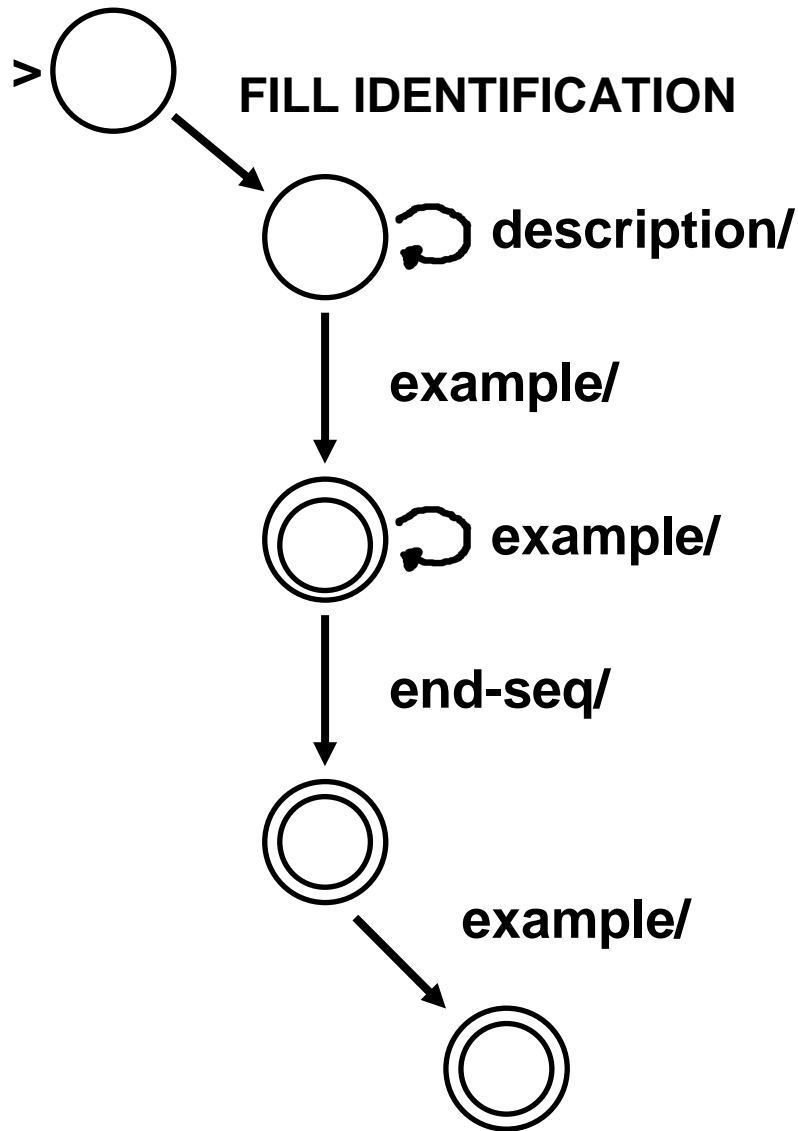
Das Generierungsmodell TEXT

- Datenbank von militärischen Fahrzeugen und Waffen (Office of Naval Research)
- beantwortet Fragen nach
 - Definition
 - „*What is a frigate?*“
 - Information in der Datenbank
 - „*What do you know about submarines?*“
 - Unterschieden
 - „*What is the difference between an ocean escort and a cruiser?*“

Wahl des Schemas

- Frage nach Definition
 - Identification
 - Costituency
- Frage nach vorhandener Information
 - Attribute
- Frage nach Unterschieden zweier Objekte
 - Compare&Contrast

Füllen des Schemas



Fokus

Ich habe mir ein Buch gekauft. Vorgestern war ich im Kino. Liest du auch gerne Krimis?

- der Gegenstand, um den es in einem Satz geht
- ein Objekt aus dem Pool
- Ermöglicht Kohärenz

Fokussiermechanismus

... Aktueller Fokus
Ich habe mir ein **Buch** gekauft.
Es ist von Allende.
...

Focus constraints (Sidner)

Ich habe mir ein **Buch** gekauft.

Es ist von **Allende**.

- Wechsel zu gerade eingeführtem Gegenstand
 - *Das ist eine chilenische Autorin.*
- Behalte den aktuellen Fokus bei
 - *In dem **Buch** geht es um...*
- Kehre zu einem vorherigen Fokus zurück
 - *Ansonsten habe ich noch viel für die Uni gemacht. (Fokus: **Wochenende**)*
- Wechsel zu einem Gegenstand, der implizit mit dem aktuellen in Zusammenhang steht
 - *Ich habe noch mehr Bücher von ihr zu Hause.*

Focus constraints in TEXT

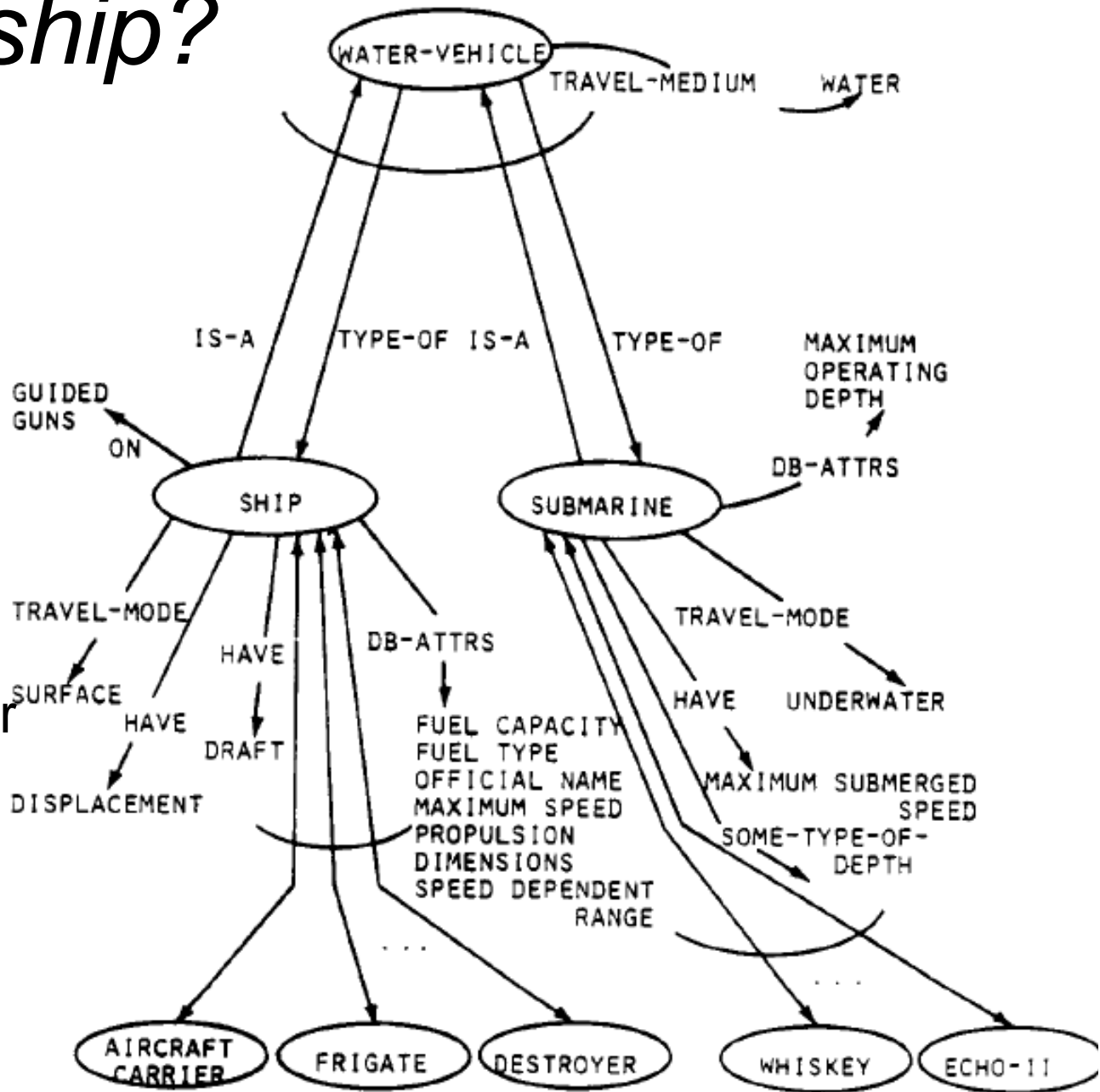
- Ordnung:
 1. Wechsel zu gerade eingeführtem Gegenstand
 2. Behalte den aktuellen Fokus bei
 3. Kehre zum Gegenstand der vorherigen Diskussion zurück
 4. (Wechsel zu einem Thema, das implizit im Zusammenhang steht) → Thema, das die meisten impliziten Verknüpfungen zur vorherigen Proposition hat

What is a ship?

1. Relevant knowledge pool

2. Wahl des Schemas

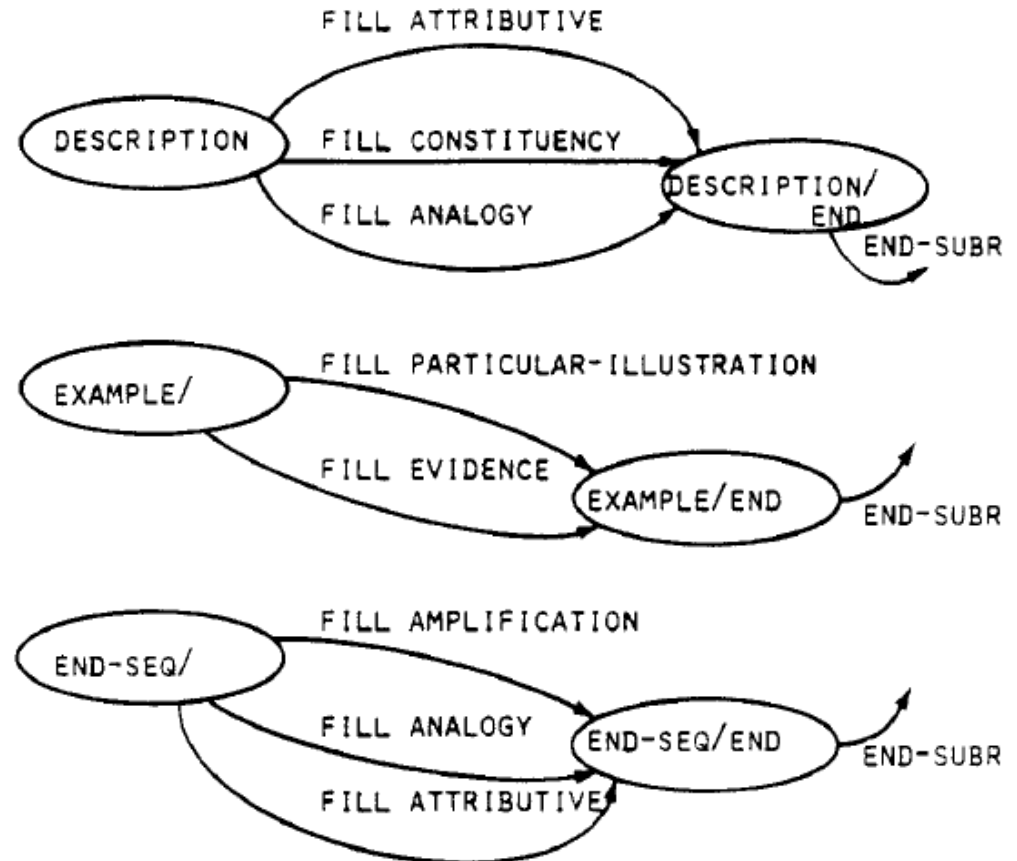
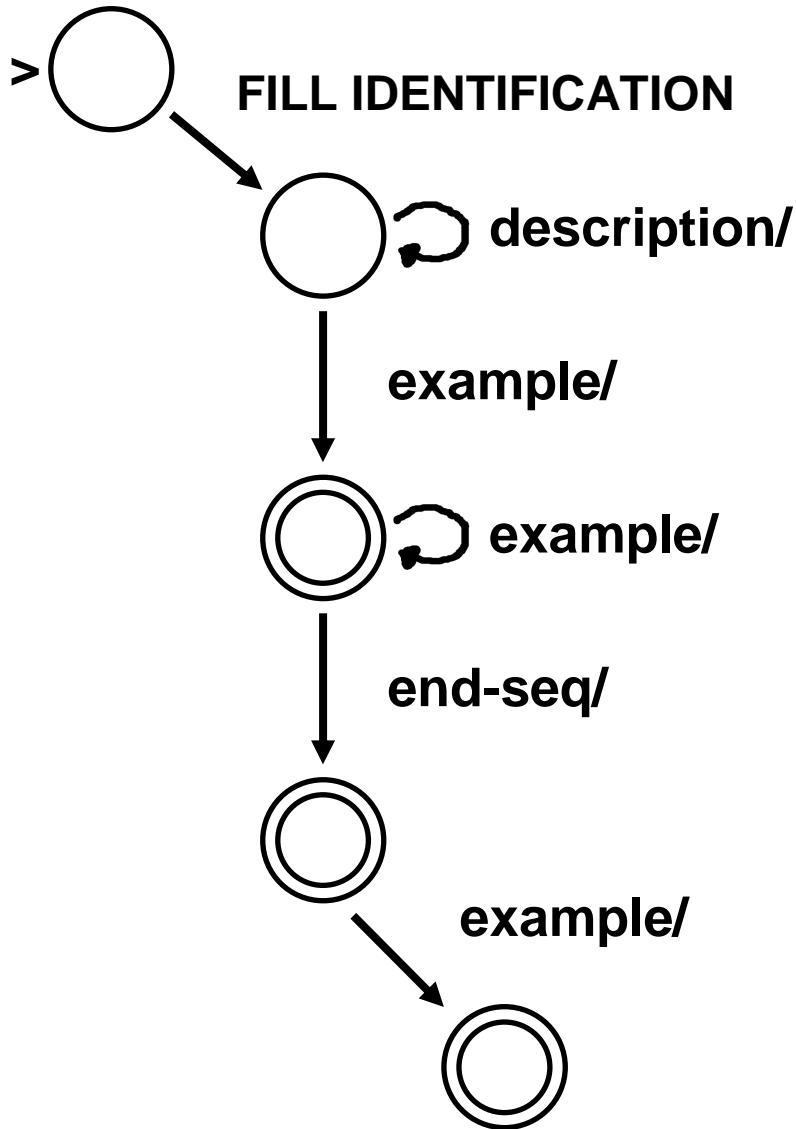
- Welcher Fragetyp?
- Identification oder constituency Schema? → identification



What is a ship?

3. Füllen des Schemas

- Graph durchlaufen



- (attributive db SHIP (name OFFICIAL_NAME) (topics SPEED_DEPENDENT_RANGE DIMENSIONS) (duplicates (FUEL FUEL_TYPE FUEL_CAPACITY)) (attrs PROPULSION MAXIMUM_SPEED))
-

- Other DB attributes of the ship include MAXIMUM SPEED, PROPULSION, FUEL (FUEL CAPACITY and FUEL TYPE), DIMENSIONS, SPEED DEPENDENT RANGE and OFFICIAL NAME.

(definition SHIP)

:

What is a ship?

:

Schema selected: identification

1. A ship is a water-going vehicle that travels on the surface. 2. Its surface-going capabilities are provided by the DB attributes DISPLACEMENT and DRAFT. 3. Other DB attributes of the ship include MAXIMUM SPEED, PROPULSION, FUEL (FUEL CAPACITY and FUEL TYPE), DIMENSIONS, SPEED DEPENDENT RANGE and OFFICIAL NAME. 4. The DOWNES, for example, has MAXIMUM SPEED of 29, PROPULSION of STMTURGRD, FUEL of 810 (FUEL CAPACITY) and BNKA (FUEL TYPE), DIMENSIONS of 25 (DRAFT), 46 (BEAM), and 438 (LENGTH) and SPEED.

Evaluierung von Antworten, die TEXT generiert hat

What is a guided projectile? (präzisionsgelenkte Munition)

(1) A guided projectile is a projectile that is self-propelled. (2) There are 2 types of guided projectiles in the ONR database: torpedoes and missiles. (3) The missile has a target location in the air or on the earth's surface. (4) The torpedo has an underwater target location. (5) The missile's target location is indicated by the DB attribute DESCRIPTION and the missile's flight capabilities are provided by the DB attribute ALTITUDE. (6) The torpedo's underwater capabilities are provided by the DB attributes under DEPTH (for example, MAXIMUM OPERATING DEPTH). (7) The guided projectile has DB attributes TIME TO TARGET & UNITS, HORZ RANGE & UNITS and NAME.

Überblick

- Probleme beim Textplanen
- Diskursstruktur
- Schemata (McKeown, TEXT)
- Fokus
- Beispiel