Errors, Intentions, and Explanations –

Feedback Generation for Language Tutoring Systems

Wolfgang Menzel

Fachbereich Informatik

Universität Hamburg



A Vision



- A Vision
- Where we are?



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity
- Constraint-Based Error Diagnosis



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity
- Constraint-Based Error Diagnosis
- Late Hypothesis Selection



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity
- Constraint-Based Error Diagnosis
- Late Hypothesis Selection
- Structural Constraints



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity
- Constraint-Based Error Diagnosis
- Late Hypothesis Selection
- Structural Constraints
- Intentions for Hypothesis Selection



- A Vision
- Where we are?
- Error Diagnosis and Ambiguity
- Dealing with Diagnostic Ambiguity
- Constraint-Based Error Diagnosis
- Late Hypothesis Selection
- Structural Constraints
- Intentions for Hypothesis Selection
- Conclusions



- having a system which . . .
 - ... facilitates goal-oriented interaction
 - ... encourages the student to actively produce language
 - ... puts few constraining limitations on language use
 - ... provides helpful feedback for the student to improve



clicking isn't enough



clicking isn't enough

Consequently, a session with one of the currently fashionable multimedia packages is like trying to learn to speak a foreign language from a tutor who is stone deaf . . .



clicking isn't enough

Consequently, a session with one of the currently fashionable multimedia packages is like trying to learn to speak a foreign language from a tutor who is stone deaf . . .

The reason is that there exists no software intelligent enough to process ill-formed sentences other than by simply refusing them.



clicking isn't enough

Consequently, a session with one of the currently fashionable multimedia packages is like trying to learn to speak a foreign language from a tutor who is stone deaf . . .

The reason is that there exists no software intelligent enough to process ill-formed sentences other than by simply refusing them.

Brian Farrington (Eurocall 2000)



- We can provide helpful error explanations . . .
 - ... under strong domain restrictions
 - ... for language with limited lexical and grammar coverage
 - ... using approximative linguistic knowledge
 - ... making simplifying assumptions on error types and maximum error complexity



- We can provide helpful error explanations . . .
 - ... under strong domain restrictions
 - ... for language with limited lexical and grammar coverage
 - ... using approximative linguistic knowledge
 - ... making simplifying assumptions on error types and maximum error complexity
- We know that . . .
 - ... existing system solutions are too restricted
 - ... explanation quality is sometimes poor



- We tend to neglect that . . .
 - ... every explanation is targeted at a particular correction proposal
 - ... the *only* criterion available to determine the optimum proposal is the "similarity" of the corrected to the original student solution
 - ... "similarity" is relative to the knowledge captured by a particular model
 - ... the quality of error descriptions depends on knowledge being available at the *right point in time*



- We believe that explanation quality could be improved by using . . .
 - ... more precise linguistic descriptions
 - ... strong support from extra-linguistic knowledge
 - ... reference to the non-linguistic context



- We believe that explanation quality could be improved by using . . .
 - ... more precise linguistic descriptions
 - ... strong support from extra-linguistic knowledge
 - ... reference to the non-linguistic context
- We should become aware of the fact that . . .
 - ... the ideal point of reference is the intention of the learner.



• Why intentions?



- Why intentions?
 - language teacher use it routinely



- Why intentions?
 - language teacher use it routinely
 - it can be inferred
 - from past behaviour of the student
 - from domain knowledge
 - from the discourse context



- Why intentions?
 - language teacher use it routinely
 - it can be inferred
 - from past behaviour of the student
 - from domain knowledge
 - from the discourse context
 - it can be directly elicited from the student
 - e.g. asking "What did you want to say?"
 - can be communicated by linguistic and non-linguistic means



- Why intentions?
 - language teacher use it routinely
 - it can be inferred
 - from past behaviour of the student
 - from domain knowledge
 - from the discourse context
 - it can be directly elicited from the student
 - e.g. asking "What did you want to say?"
 - can be communicated by linguistic and non-linguistic means
 - could provide a strong support for guiding the diagnosis



ambiguity

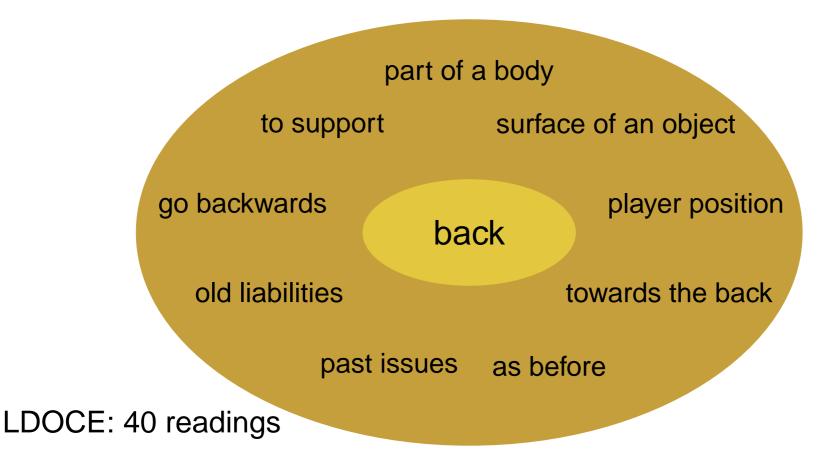


ambiguity





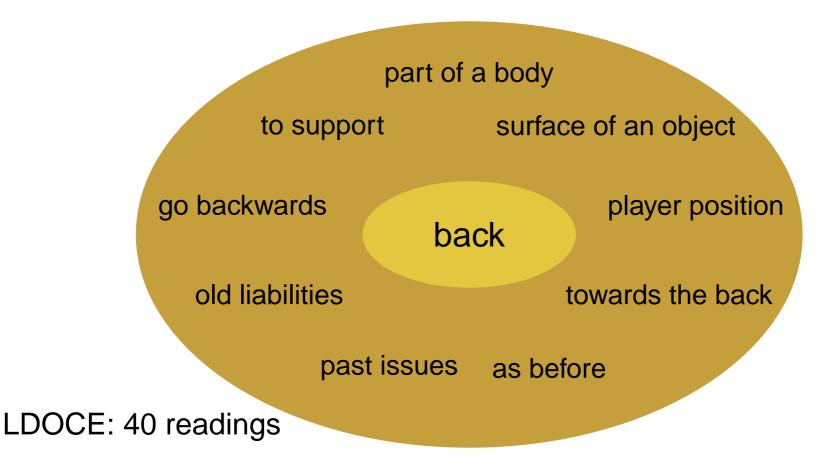
ambiguity





ambiguity

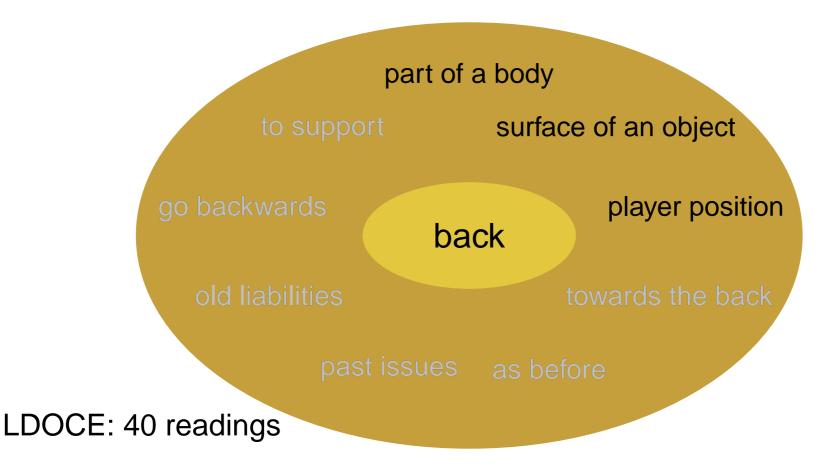
He painted the back twice.





ambiguity

He painted the back twice.

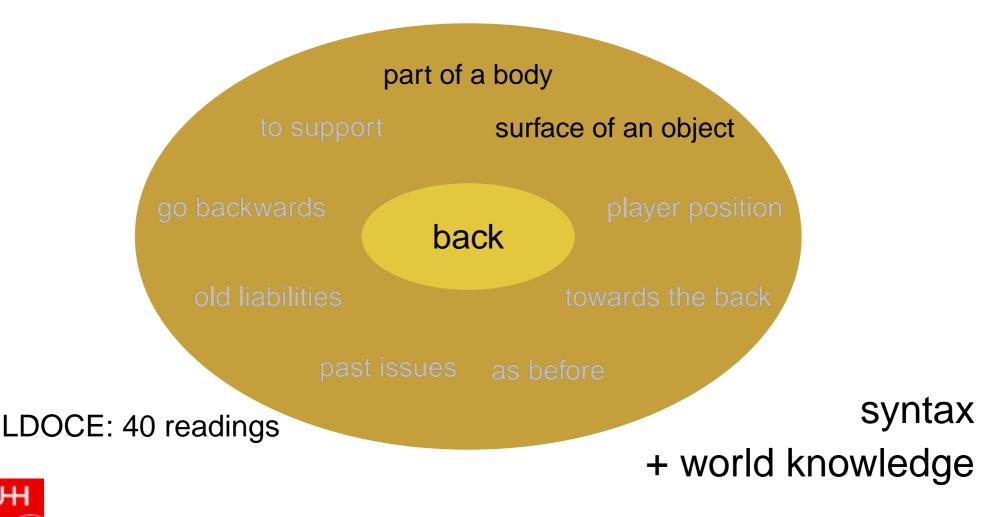




syntax

ambiguity

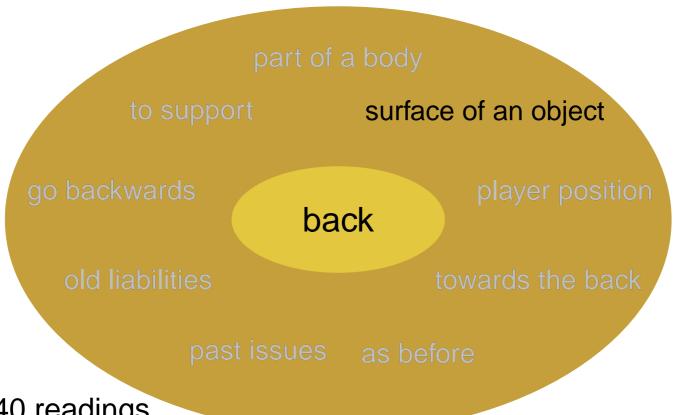
He painted the back twice.





ambiguity

He painted the back twice.





syntax

- + world knowledge
- + domain knowledge



language errors introduce additional ambiguity



- language errors introduce additional ambiguity
- example: spelling correction



- language errors introduce additional ambiguity
- example: spelling correction

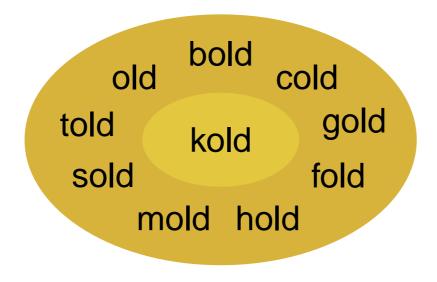
*It is very kold in here.





- language errors introduce additional ambiguity
- example: spelling correction

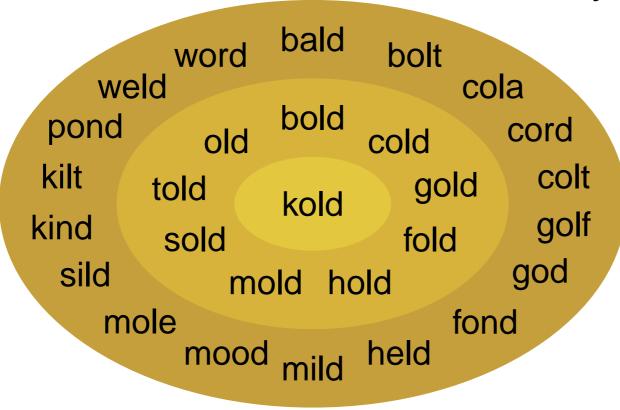
*It is very kold in here.





- language errors introduce additional ambiguity
- example: spelling correction

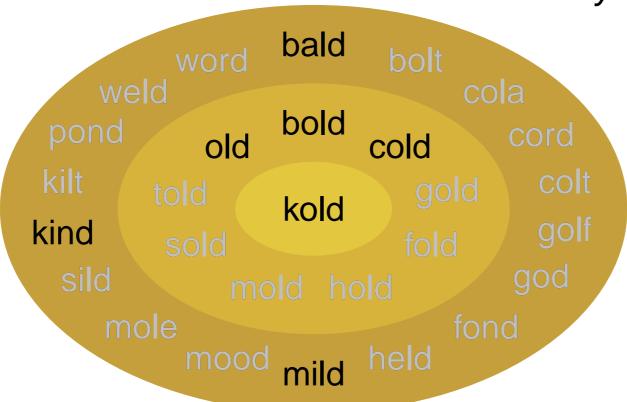
*It is very kold in here.





- language errors introduce additional ambiguity
- example: spelling correction

*It is very kold in here.

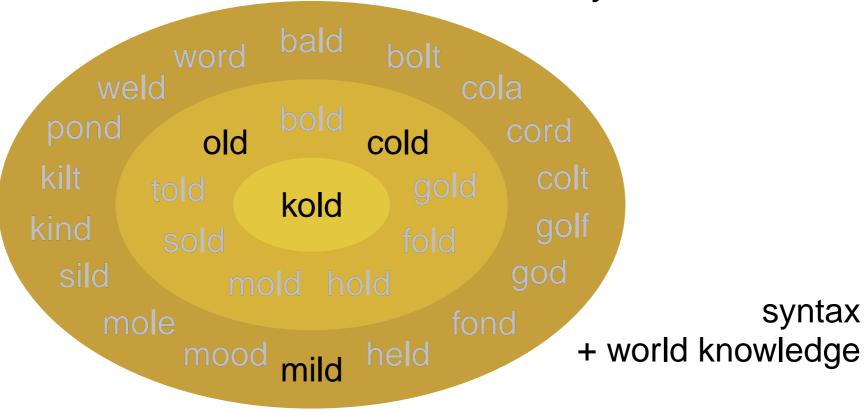


syntax



- language errors introduce additional ambiguity
- example: spelling correction

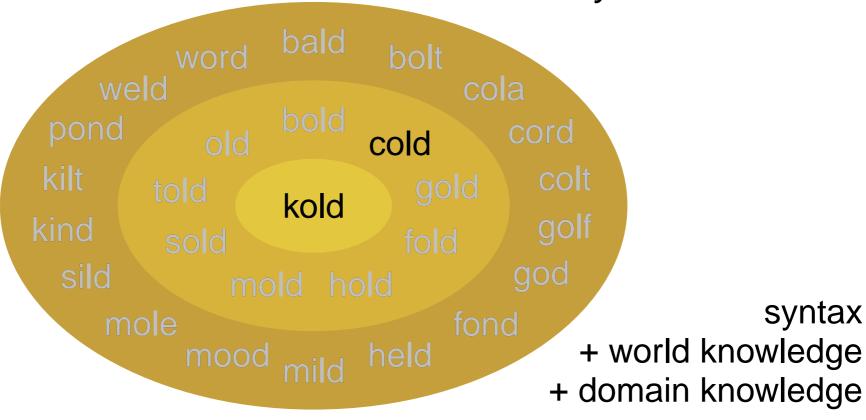
*It is very kold in here.





- language errors introduce additional ambiguity
- example: spelling correction

*It is very kold in here.





 minimal error heuristics: prefer simpler correction proposals over more complex one



- minimal error heuristics: prefer simpler correction proposals over more complex one
 - can be misleading

*It is very kolds in here.

 $kolds \rightarrow holds$.



- minimal error heuristics: prefer simpler correction proposals over more complex one
 - can be misleading

*It is very kolds in here.

 $kolds \rightarrow holds$.

 error assumptions are even necessary for correct word forms if the utterance is syntactically inacceptable

*It is very told in here.



- character-based correction proposals are rather poor explanations no indication
 - what might have caused the error
 - how to avoid a similar error in the future



- character-based correction proposals are rather poor explanations no indication
 - what might have caused the error
 - how to avoid a similar error in the future
- possible causes:
 - substitution or insertion of neighboring keys
 - phonetic similarity
 - interference from another language



- different error perspectives
 - the same error can be explained in different ways
 - introduces yet another type of diagnostic ambiguity
 - provides also additional criteria for hypothesis selection



- different error perspectives
 - the same error can be explained in different ways
 - introduces yet another type of diagnostic ambiguity
 - provides also additional criteria for hypothesis selection
- kold → cold: phonetic confusion is most plausible



 different perspectives might lead to differently complex error descriptions

*It was there fault.

- no error assumption on a purely phonetic level
- character-based explanation with two substitutions
 - → minimal error heuristics fails if phonetic similarity is ignored



 different perspectives might lead to differently complex error descriptions

*It was there fault.

- no error assumption on a purely phonetic level
- character-based explanation with two substitutions
 - → minimal error heuristics fails if phonetic similarity is ignored
 - cheaper character-based corrections available:

there → here

do not remove the syntactic inconsistency



- 1. The necessity to consider erroneous input increases (local) ambiguity.
 - enumerating all possible correction possibilities is neither feasible nor desirable.



- 1. The necessity to consider erroneous input increases (local) ambiguity.
 - enumerating all possible correction possibilities is neither feasible nor desirable.
- 2. Errors can be explained from different perspectives.
 - the perspective might influence the decision on the most plausible explanation
 - usually plausibility is a gradual notion



- 1. The necessity to consider erroneous input increases (local) ambiguity.
 - enumerating all possible correction possibilities is neither feasible nor desirable.
- 2. Errors can be explained from different perspectives.
 - the perspective might influence the decision on the most plausible explanation
 - usually plausibility is a gradual notion
- 3. Least effort corrections do not always yield satisfying error explanations.



- 4. The diagnosis procedure should be aware of alternative explanation/correction possibilities
 - the alternatives can be more plausible from another perspective.



- 4. The diagnosis procedure should be aware of alternative explanation/correction possibilities
 - the alternatives can be more plausible from another perspective.
- 5. Considering an error might even be necessary if the input seems (locally) acceptable.



 huge number of error hypotheses even for relatively simple problems



- huge number of error hypotheses even for relatively simple problems
- needed: strong constraints to narrow down the space of possible alternatives



- huge number of error hypotheses even for relatively simple problems
- needed: strong constraints to narrow down the space of possible alternatives
- two approaches:
 - 1. artificially constrain the sublanguage
 - 2. use (dynamic) domain knowledge about the current state of affairs



- huge number of error hypotheses even for relatively simple problems
- needed: strong constraints to narrow down the space of possible alternatives
- two approaches:
 - 1. artificially constrain the sublanguage
 - 2. use (dynamic) domain knowledge about the current state of affairs



two architectures



- two architectures
 - 1. late selection:
 - diagnosis produces an as large as possible set of potential error explanations
 - a subsequent selection step selects the most plausible ones



- two architectures
 - 1. late selection:
 - diagnosis produces an as large as possible set of potential error explanations
 - a subsequent selection step selects the most plausible ones
 - 2. early integration:
 - the domain knowledge is directly integrated into the diagnosis procedure
 - guides it towards the most plausible explanation

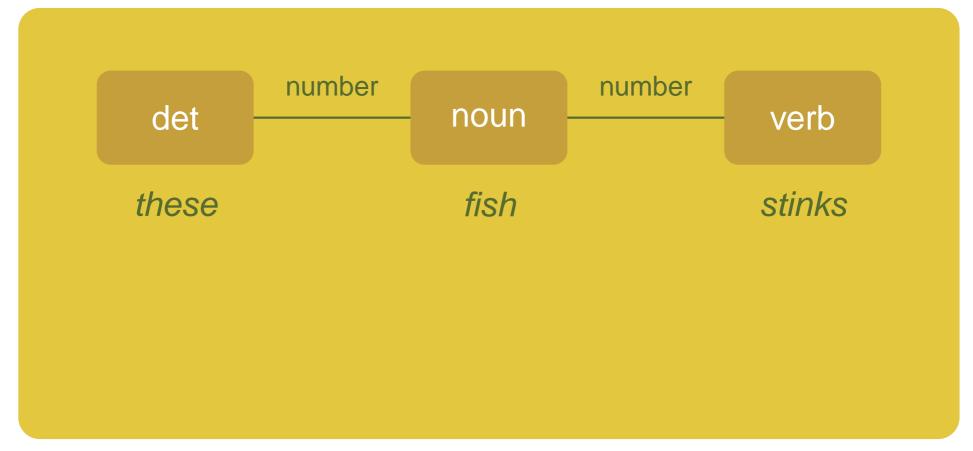


- example: morpho-syntactic regulartities
 - constraints model the compatibility of feature assignments
 - feature assignments need not be unique

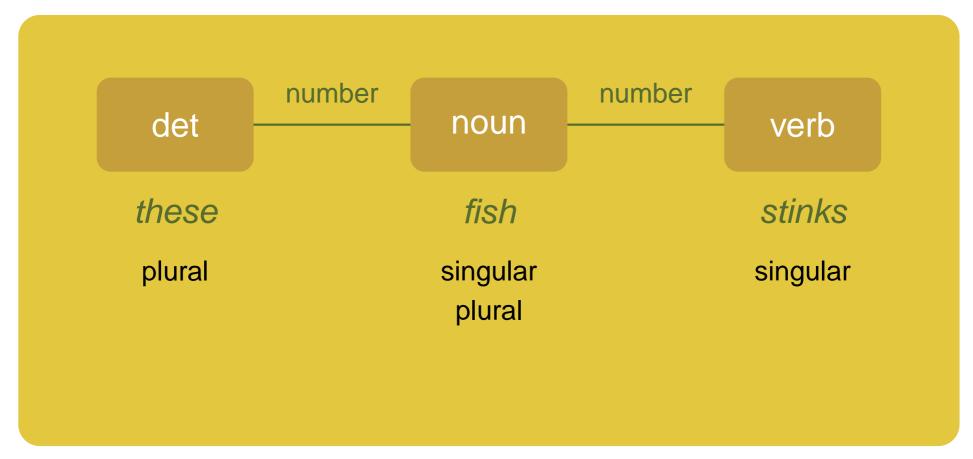


- example: morpho-syntactic regulartities
 - constraints model the compatibility of feature assignments
 - feature assignments need not be unique
- for all variable assignments find the constraints that are violated
- if an assignment with no constraint violations is found, signal "ok"
- else find the assignment with a minimum number of constraint violations
- output the explanation(s) connected to the violated constraint(s)

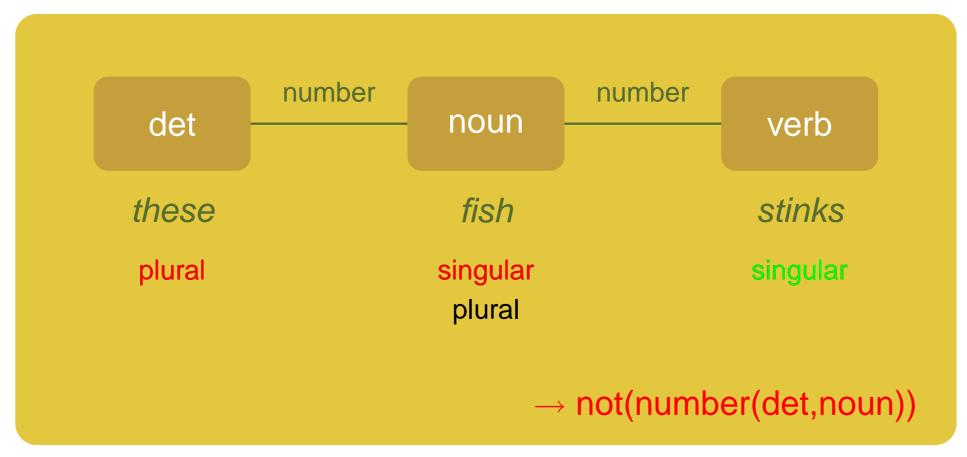




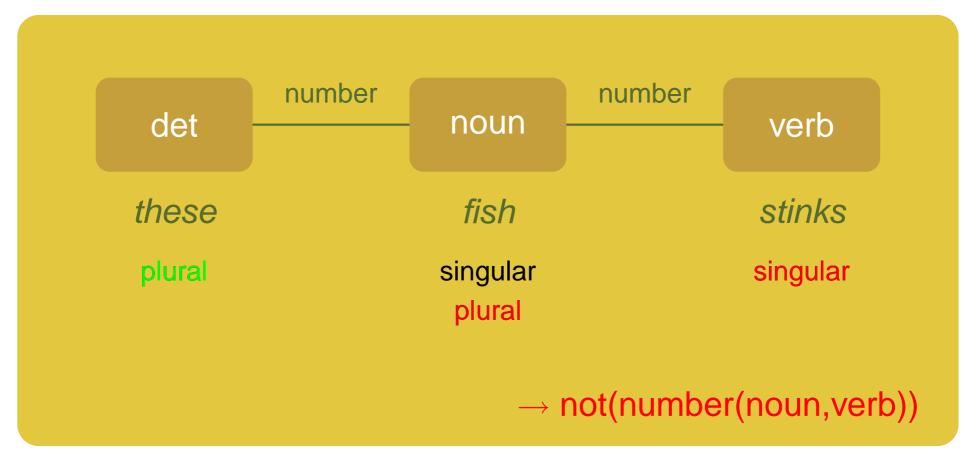














constraint system of English is rather impoverished



- constraint system of English is rather impoverished
- richer constraint systems in other languages

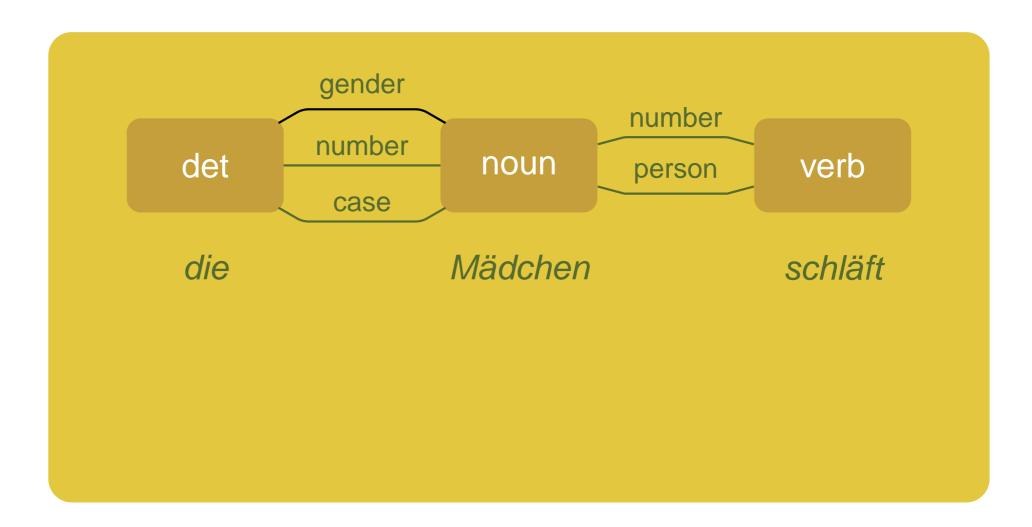


- constraint system of English is rather impoverished
- richer constraint systems in other languages
- e.g. German
 - subject-verb: person, number
 - noun phrase: number, gender, case

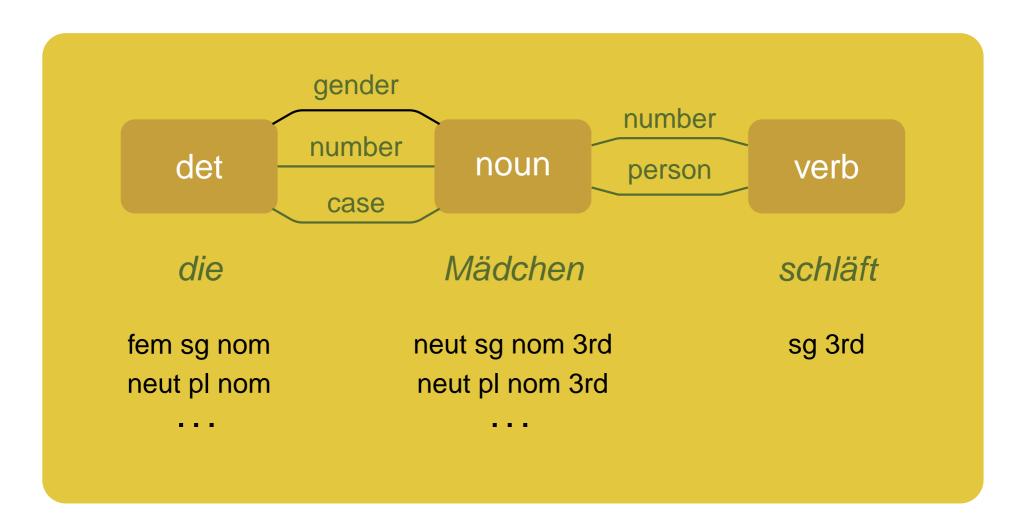


- constraint system of English is rather impoverished
- richer constraint systems in other languages
- e.g. German
 - subject-verb: person, number
 - noun phrase: number, gender, case
- e.g. Russian
 - subject-verb: person, number, gender
 - noun phrase: number, gender, case, animatedness

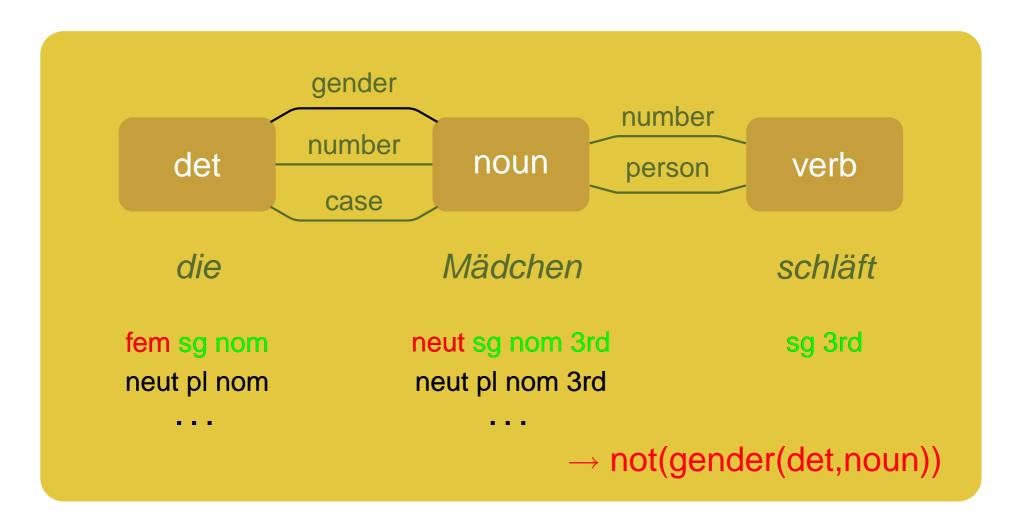




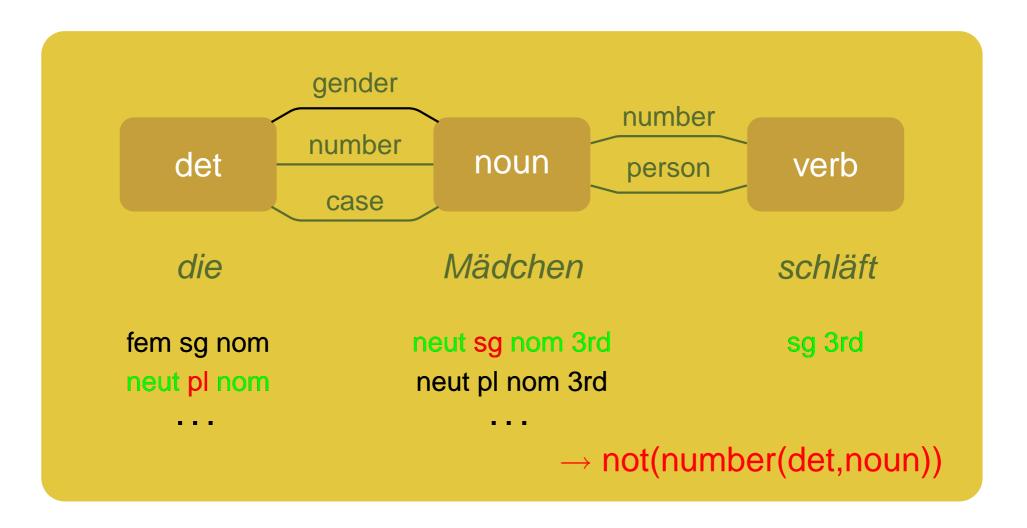




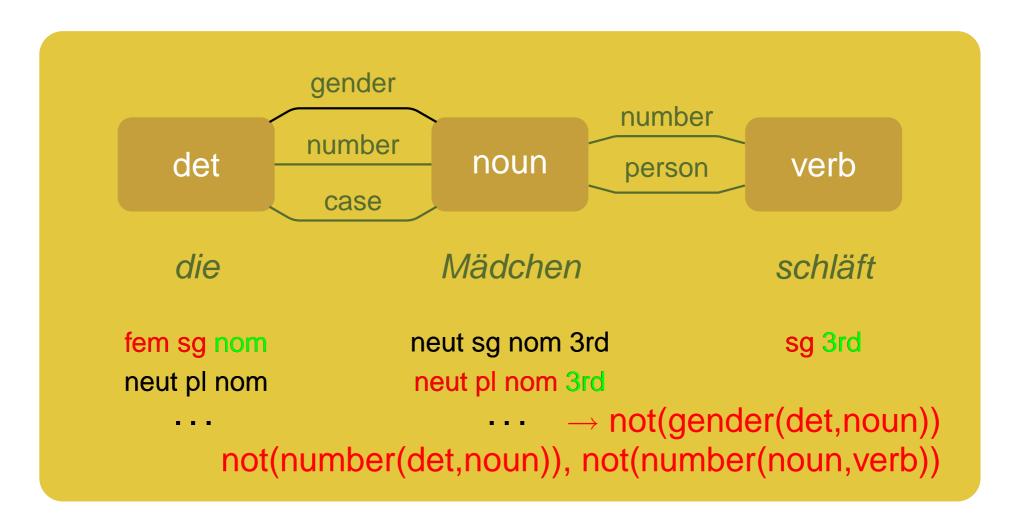




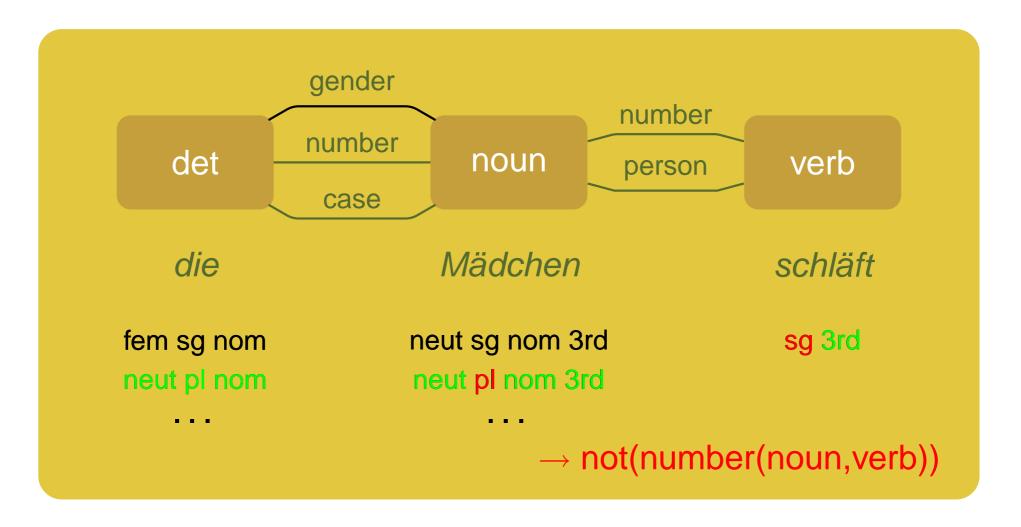














- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses



- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses
- Schwind (1994): constraint checking within the scope of phrase structure rules
 - alternative (more plausible) diagnoses are lost



- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses
- Schwind (1994): constraint checking within the scope of phrase structure rules
 - alternative (more plausible) diagnoses are lost

Der Götter zürnen



- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses
- Schwind (1994): constraint checking within the scope of phrase structure rules
 - alternative (more plausible) diagnoses are lost



correct genitive case NP



- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses
- Schwind (1994): constraint checking within the scope of phrase structure rules
 - alternative (more plausible) diagnoses are lost



correct genitive case NP

subject NP without nominative case



- global consistency required
 - locally restricted contraint checking leads to incomplete diagnoses
- Schwind (1994): constraint checking within the scope of phrase structure rules
 - alternative (more plausible) diagnoses are lost



correct genitive case NP

subject NP without nominative case

special treatment proposed



- Holland (1995)
 - partial diagnoses trigger contingent errors



- Holland (1995)
 - partial diagnoses trigger contingent errors

Wir stehen auf die Berg



- Holland (1995)
 - partial diagnoses trigger contingent errors

Wir stehen auf die Berg missing gender agreement



- Holland (1995)
 - partial diagnoses trigger contingent errors

Wir stehen auf die Berg missing gender agreement

Wir stehen auf der Berg

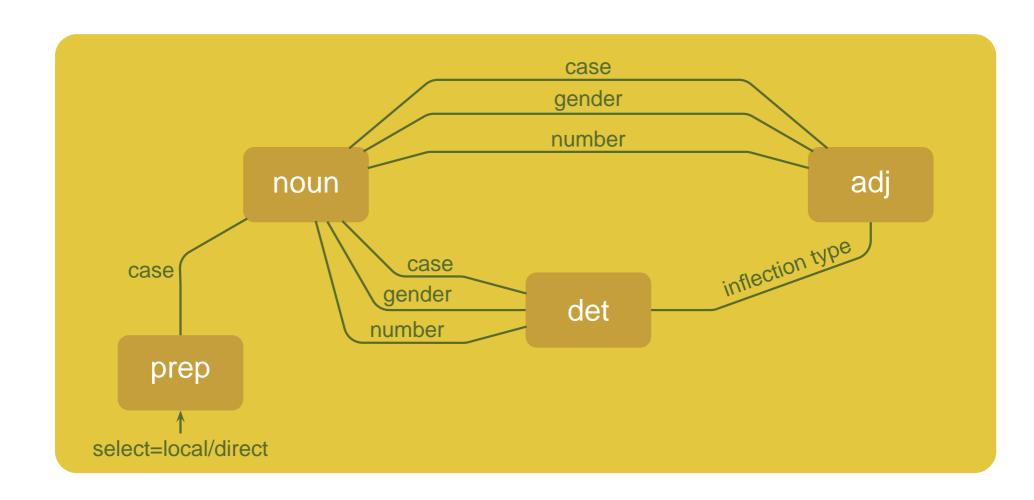


- Holland (1995)
 - partial diagnoses trigger contingent errors



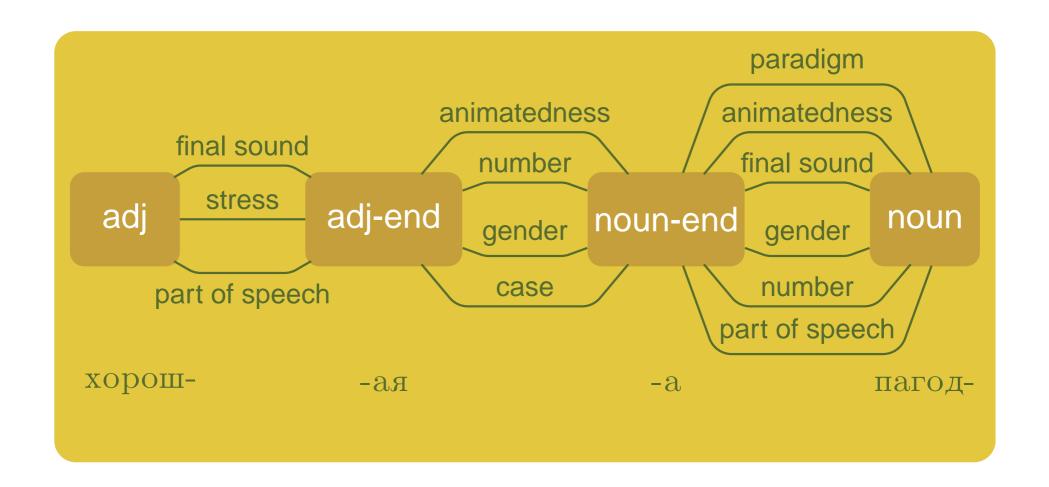


Agreement in a German PP



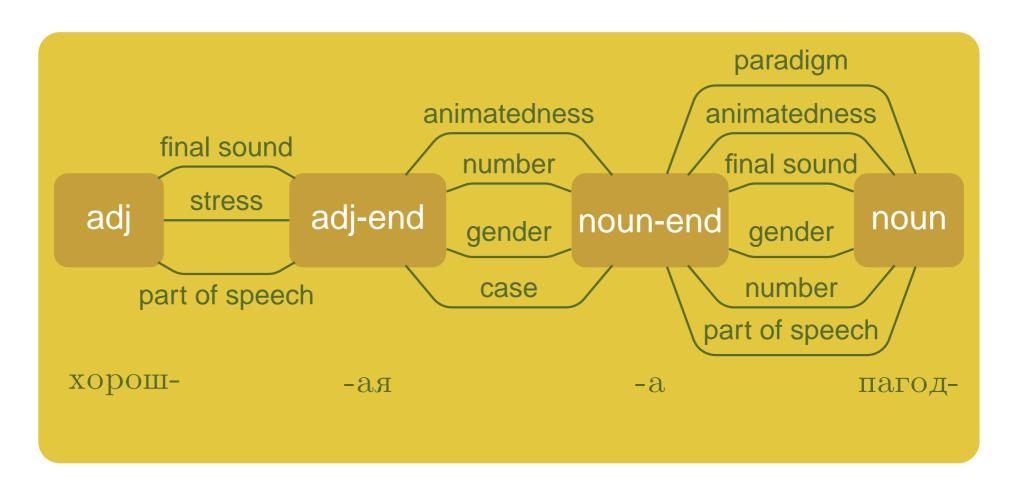


Inflection in a Russian NP





Inflection in a Russian NP



 additional diagnostic capabilities: detection of erroneous inflection patterns



- alternative error perspective: fact errors
 - the student might have thought that "child" is not a singular form:

not(value(noun, singular))



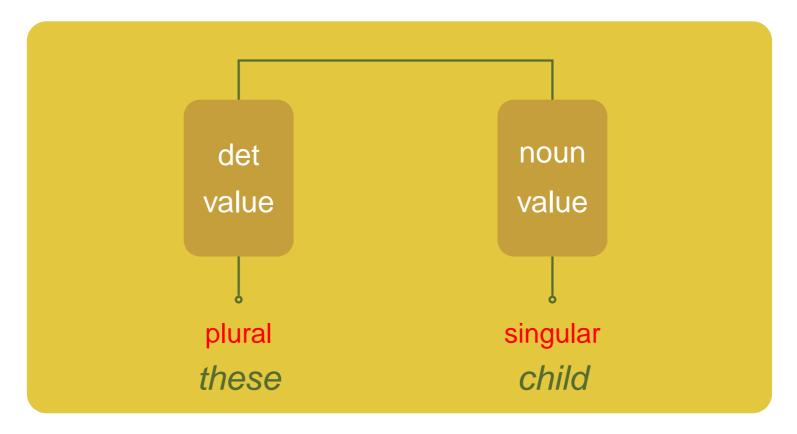
- alternative error perspective: fact errors
 - the student might have thought that "child" is not a singular form:
 - not(value(noun, singular))
- assuming the ignorance of lexical information
 - different view on the same error
 - sometimes yields more concise explanations



- alternative error perspective: fact errors
 - the student might have thought that "child" is not a singular form:
 - not(value(noun, singular))
- assuming the ignorance of lexical information
 - different view on the same error
 - sometimes yields more concise explanations
- performing a separate error simulation with lexical value assignment components



 without an disambiguating context fact diagnoses are always ambiguous





the two error perspectives are complementary



- the two error perspectives are complementary
- provide alternative information about an error
 - grammar rules vs. correction proposals



- the two error perspectives are complementary
- provide alternative information about an error
 - grammar rules vs. correction proposals
- descriptions can be differently complex



- the two error perspectives are complementary
- provide alternative information about an error
 - grammar rules vs. correction proposals
- descriptions can be differently complex
- both are required



- the two error perspectives are complementary
- provide alternative information about an error
 - grammar rules vs. correction proposals
- descriptions can be differently complex
- both are required
- Schwind (1994): rule errors only



- the two error perspectives are complementary
- provide alternative information about an error
 - grammar rules vs. correction proposals
- descriptions can be differently complex
- both are required
- Schwind (1994): rule errors only
- Heift (1999): fact errors only

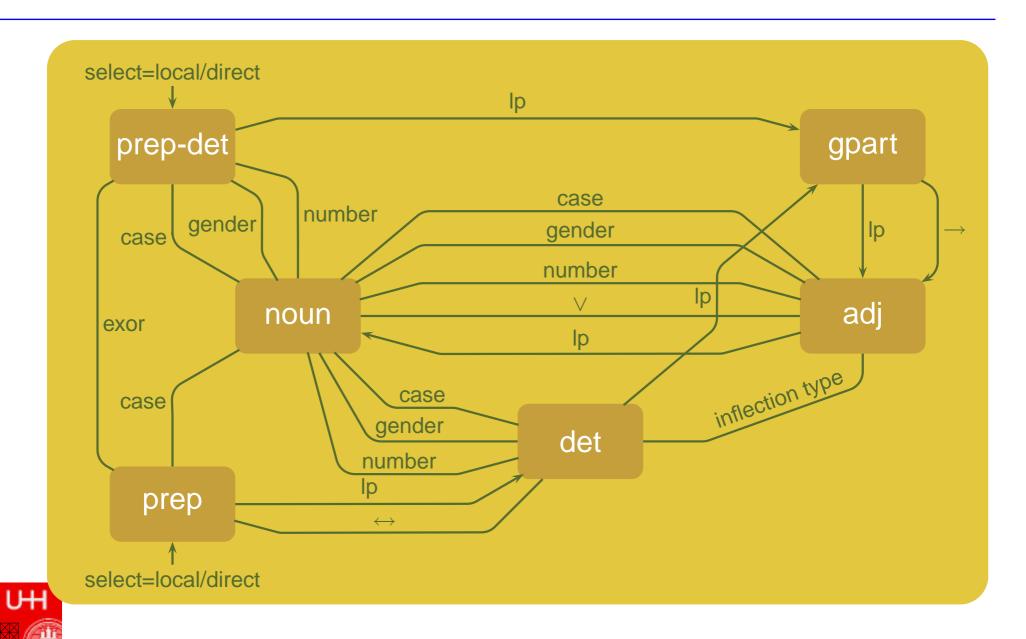


- approach can be extended to
 - linear precedence regularities
 - co-occurrence constraints

within the limits of a fixed structural pattern



Constraints for a German PP



- flexible exercises
 - free lexical choice (within the limitations of the dictionary)



- flexible exercises
 - free lexical choice (within the limitations of the dictionary)
- highly precise diagnoses in limited exercises
 - diagnostic results can be used to retrieve alternative forms from the dictionary → correction proposals
 - explorative learning-by-doing experiments becomes possible



- flexible exercises
 - free lexical choice (within the limitations of the dictionary)
- highly precise diagnoses in limited exercises
 - diagnostic results can be used to retrieve alternative forms from the dictionary → correction proposals
 - explorative learning-by-doing experiments becomes possible
- no diagnostic bias
- multitude of diagnostic information
 - hypothesis selection required
 - selection can be sensitive to a didactic goal and / or the desires UH of the student

Late Hypothesis Selection

system architecture



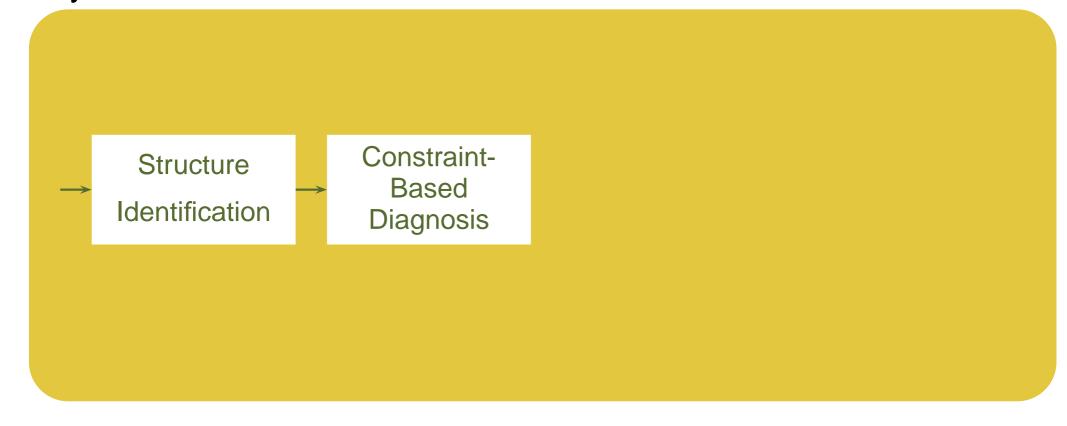
Late Hypothesis Selection

system architecture

Structure Identification

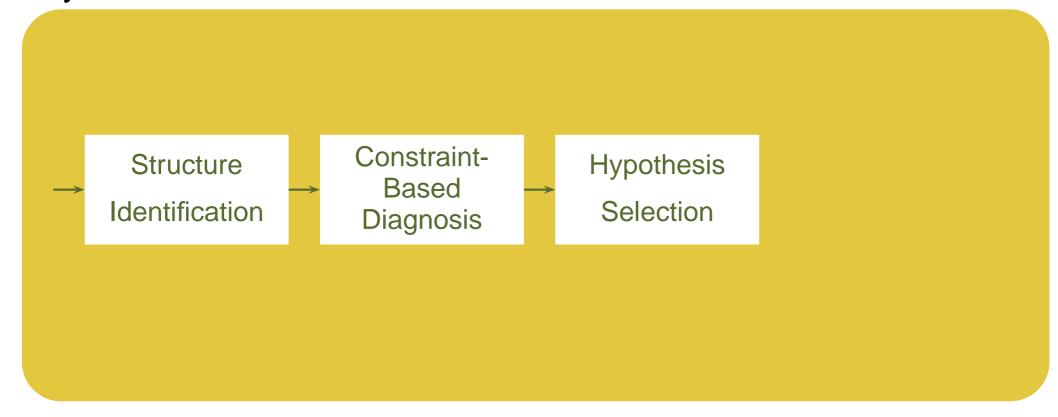


system architecture



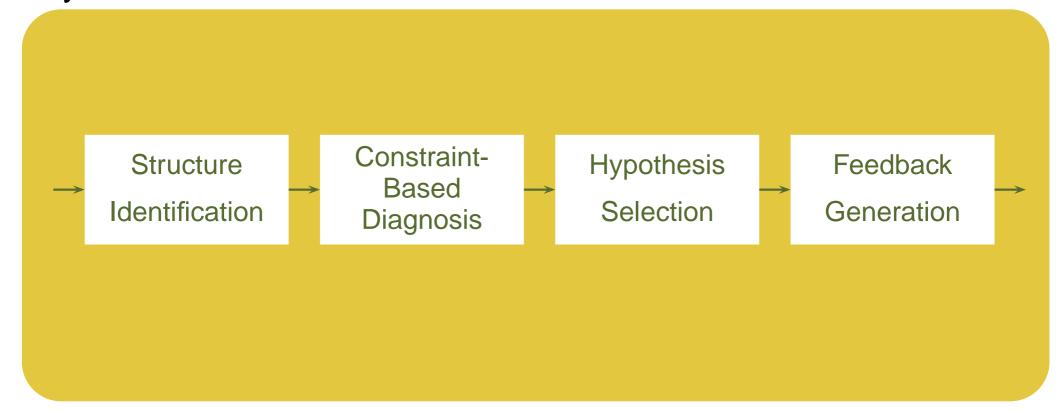


system architecture





system architecture





selection heuristics



- selection heuristics
 - minimality



- selection heuristics
 - minimality
 - higher up in a syntactic structure:
 - better reflects the violated grammar rule



- selection heuristics
 - minimality
 - higher up in a syntactic structure:
 - better reflects the violated grammar rule
 - deeper down in a syntactic structure:
 - better indicates a correction possibility



- selection heuristics
 - minimality
 - higher up in a syntactic structure:
 - better reflects the violated grammar rule
 - deeper down in a syntactic structure:
 - better indicates a correction possibility
 - preference for constraint violations:
 - better reflects the violated grammar rule



- selection heuristics
 - minimality
 - higher up in a syntactic structure:
 - better reflects the violated grammar rule
 - deeper down in a syntactic structure:
 - better indicates a correction possibility
 - preference for constraint violations:
 - better reflects the violated grammar rule
 - preference for lexical error descriptions:
 - better indicates a correction possibility



selection heuristics (cont.)



- selection heuristics (cont.)
 - conjunctive or disjunctive combinability:
 - results in more compact error descriptions



- selection heuristics (cont.)
 - conjunctive or disjunctive combinability:
 - results in more compact error descriptions
 - citation form preference:
 - explanations referring to the default case are more plausible



- selection heuristics (cont.)
 - conjunctive or disjunctive combinability:
 - results in more compact error descriptions
 - citation form preference:
 - explanations referring to the default case are more plausible
 - (L1 dependent) error type preference:
 - typical errors are more likely (e.g. gender in German)



late selection is only possible for limited exercises



- late selection is only possible for limited exercises
- full enumeration of alternative hypotheses for more complex models is infeasible
 - direct integration of diagnosis and selection is necessary



extending the idea of constraint retraction to syntactic structures



- extending the idea of constraint retraction to syntactic structures
- instead of having rules to construct hierarchical representations use constraints to describe the space of possible structural descriptions



- extending the idea of constraint retraction to syntactic structures
- instead of having rules to construct hierarchical representations use constraints to describe the space of possible structural descriptions
- Constraint Dependency Grammar (MARUYAMA 1990)



- extending the idea of constraint retraction to syntactic structures
- instead of having rules to construct hierarchical representations use constraints to describe the space of possible structural descriptions
- Constraint Dependency Grammar (MARUYAMA 1990)
- initial space of hypotheses:
 - fully underspecified structural descriptions
 - every node modifies every other with all possible labels
 - containing all possible dependency trees for an utterance

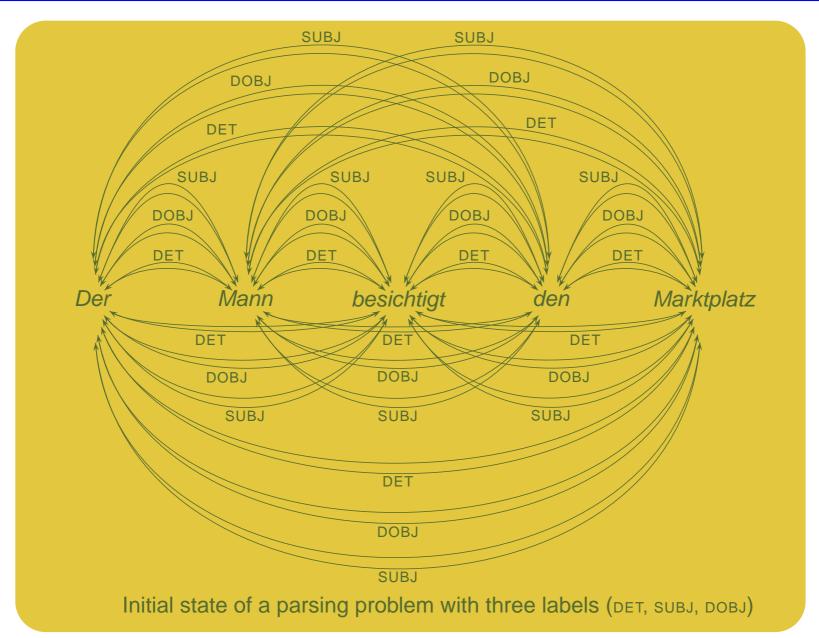


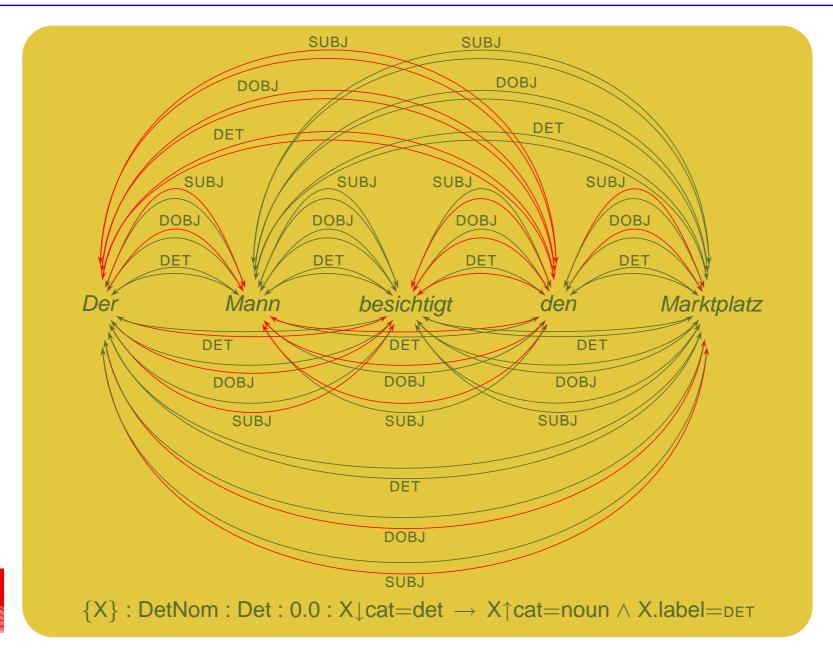
- constraints license certain dependency edges or combinations thereof
 - dependency edges which violate a constraint can be removed from the space of structural hypotheses

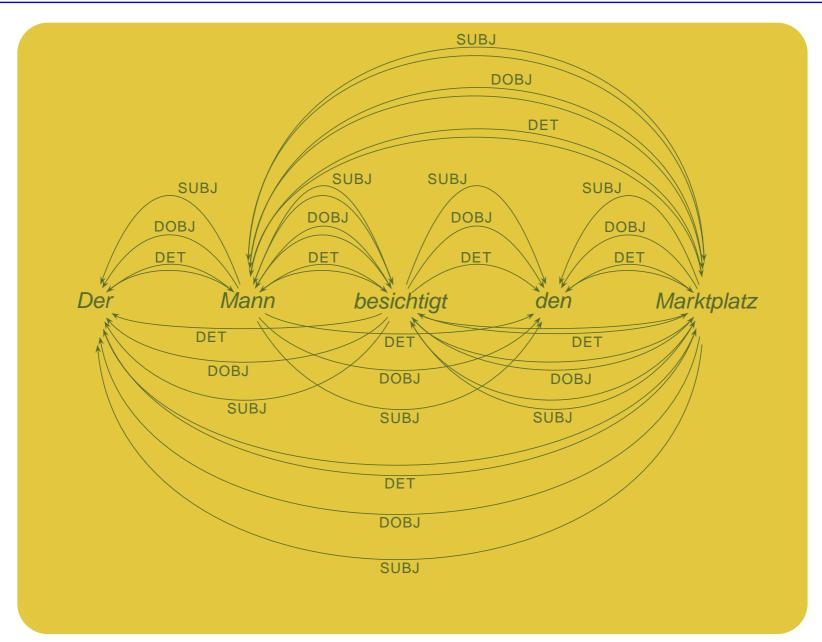


- constraints license certain dependency edges or combinations thereof
 - dependency edges which violate a constraint can be removed from the space of structural hypotheses
- constraints can be weighted
 - edges which violate a constraint are deprecated not removed
 - parsing becomes a constraint optimization problem
 - uncertain and preferntial knowledge can be included
 - e.g. the subject usually precedes the object

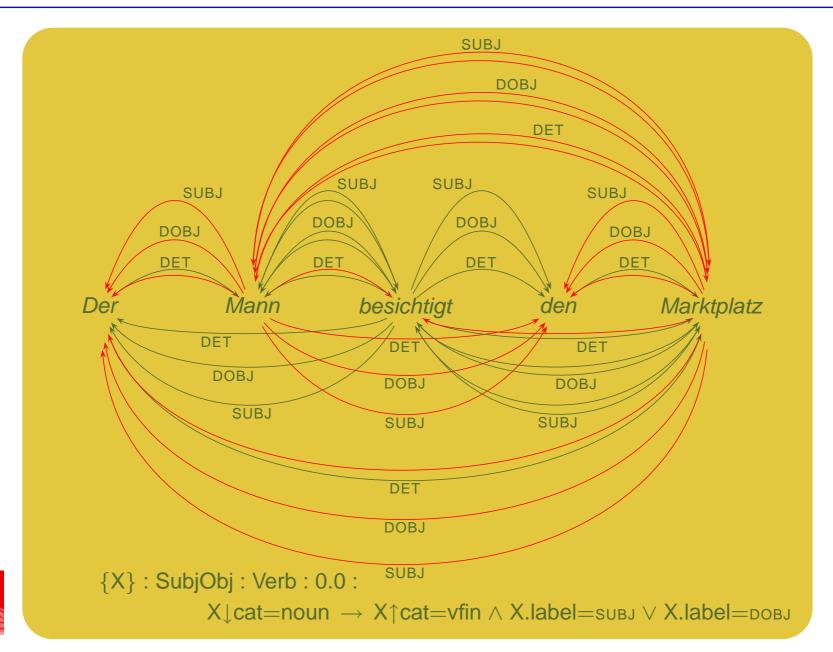




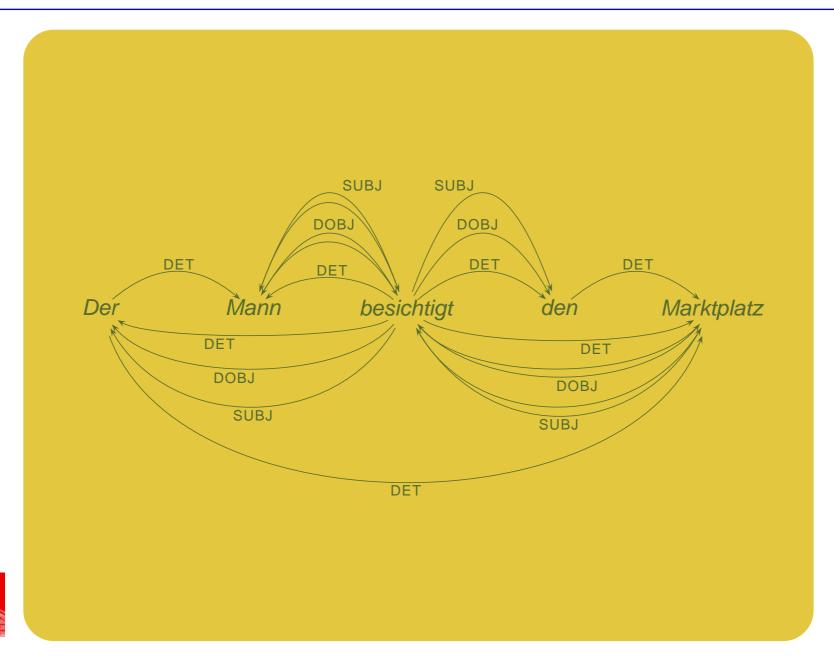


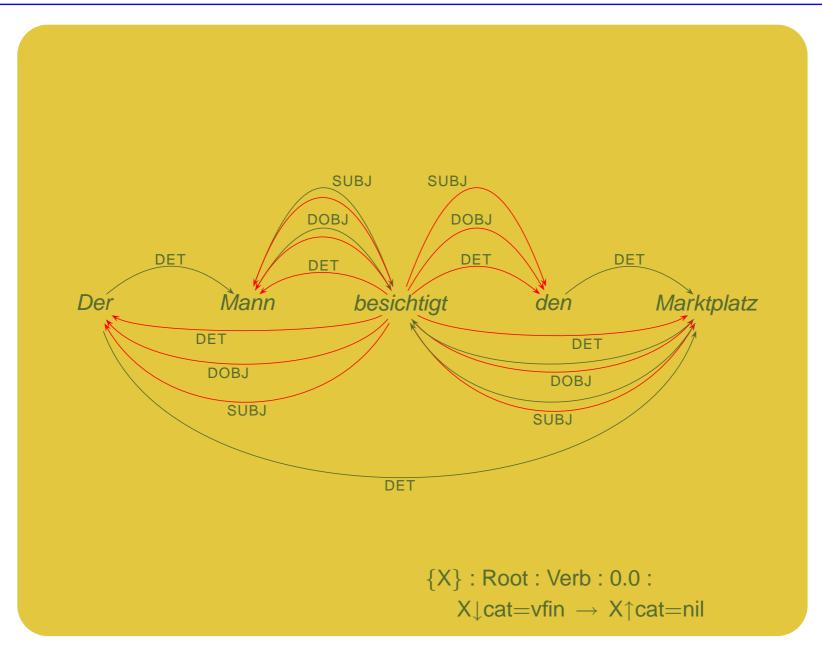




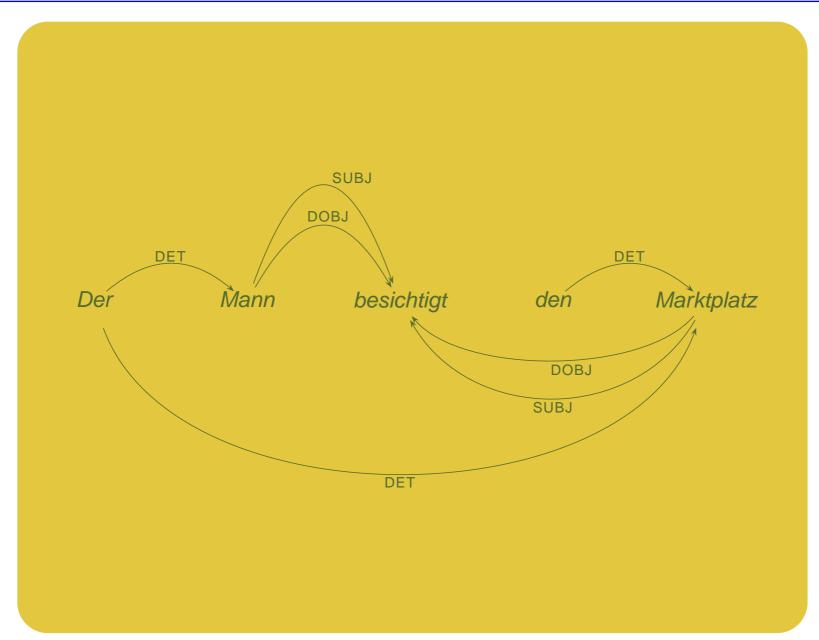


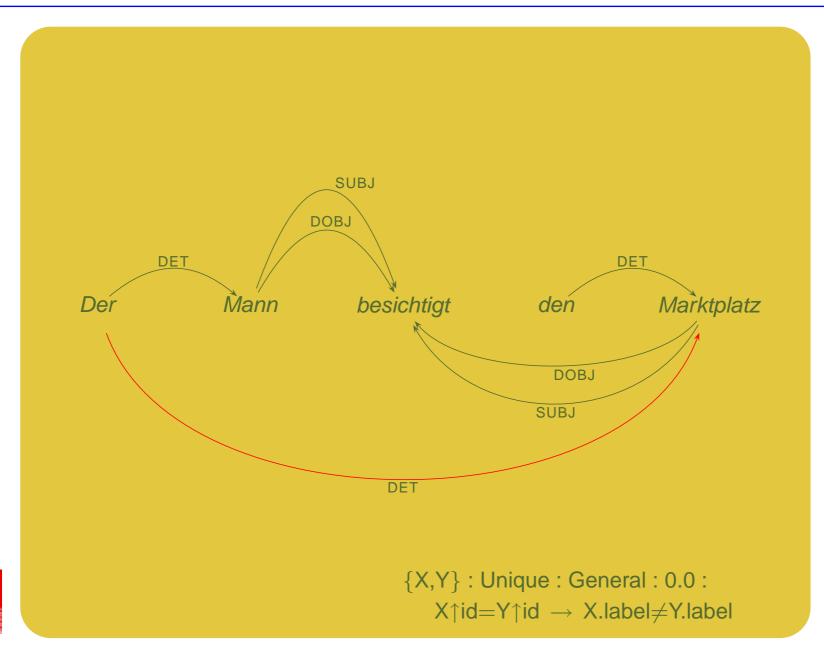




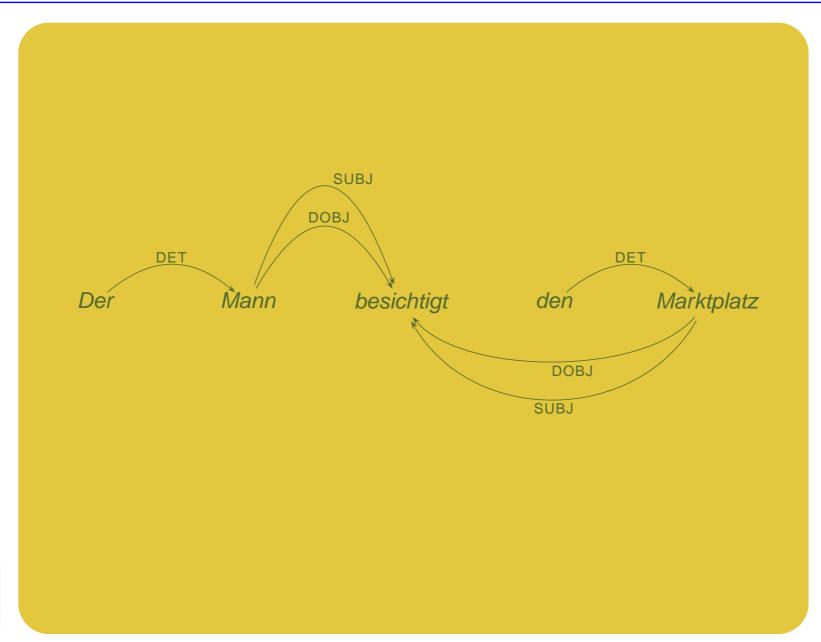


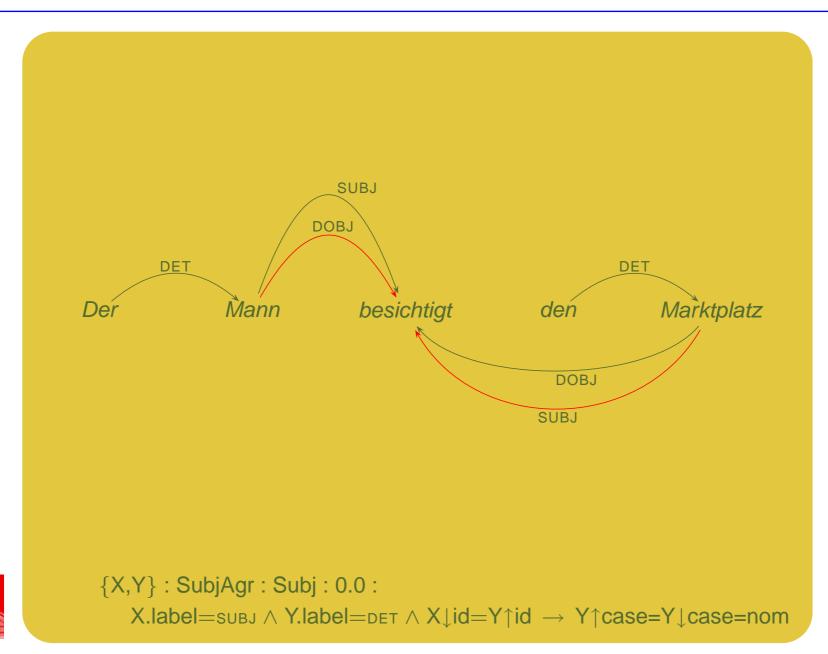




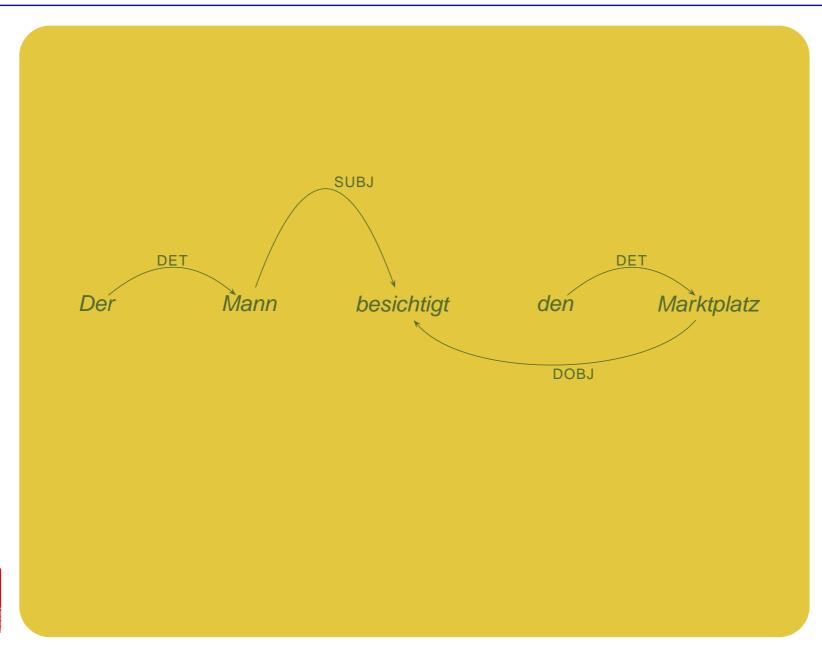












many constraints are defeasible



- many constraints are defeasible
- almost arbitrary input can be analysed



- many constraints are defeasible
- almost arbitrary input can be analysed
- constraint violations in the optimum solution can be interpreted as diagnoses



- many constraints are defeasible
- almost arbitrary input can be analysed
- constraint violations in the optimum solution can be interpreted as diagnoses
- modelling of transitive closures is only approximative
 - correction proposal cannot be derived reliably



combination with error simulation in a two phase-diagnosis



- combination with error simulation in a two phase-diagnosis
- constraint-based error simulation
 - highly precise and supports multiple explanation perspectives
 - but requires a syntactic structure being given



- combination with error simulation in a two phase-diagnosis
- constraint-based error simulation
 - highly precise and supports multiple explanation perspectives
 - but requires a syntactic structure being given
- parsing as constraint optimization
 - (so far) no precise error diagnosis in transitive correctness conditions (like agreement)
 - but determines a syntactic structure

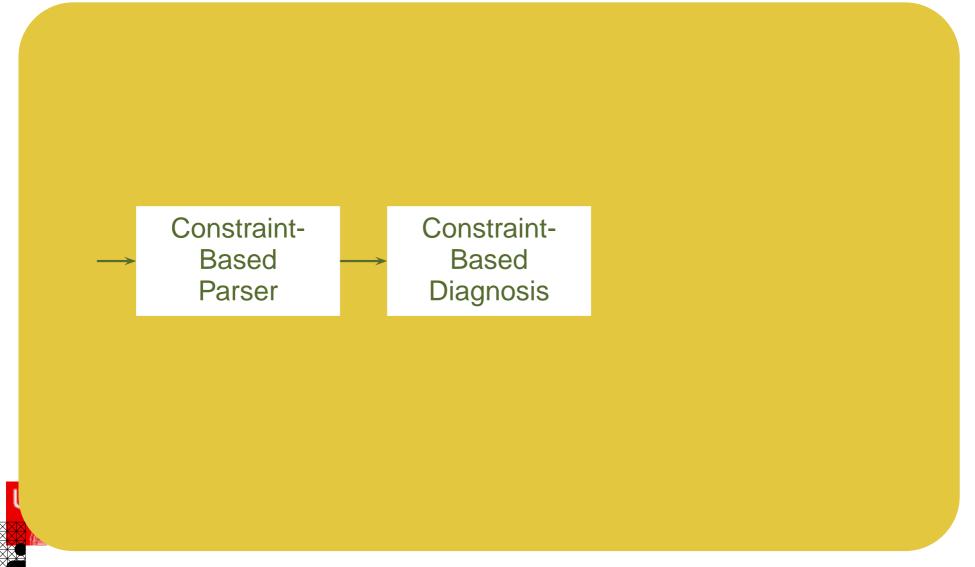


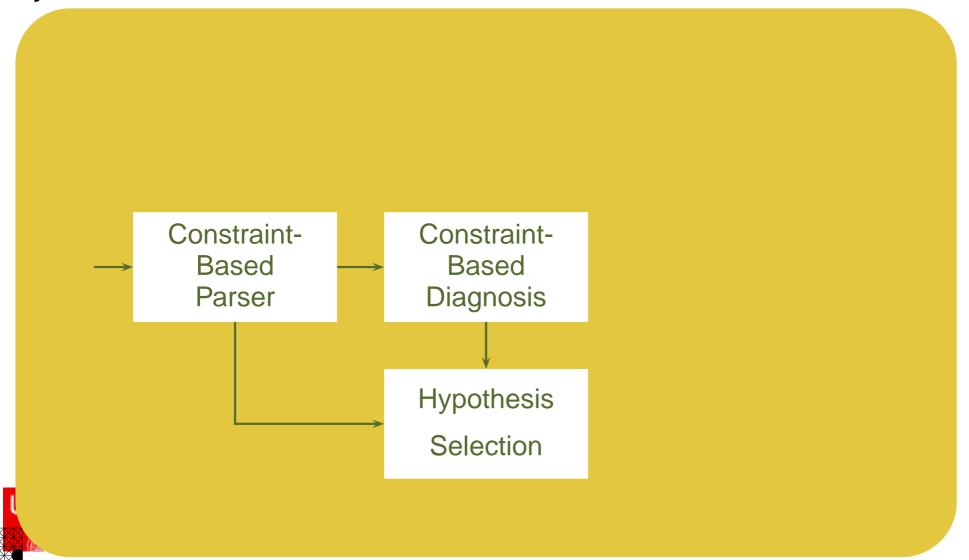
- combination with error simulation in a two phase-diagnosis
- constraint-based error simulation
 - highly precise and supports multiple explanation perspectives
 - but requires a syntactic structure being given
- parsing as constraint optimization
 - (so far) no precise error diagnosis in transitive correctness conditions (like agreement)
 - but determines a syntactic structure
- good synergy when used in combination

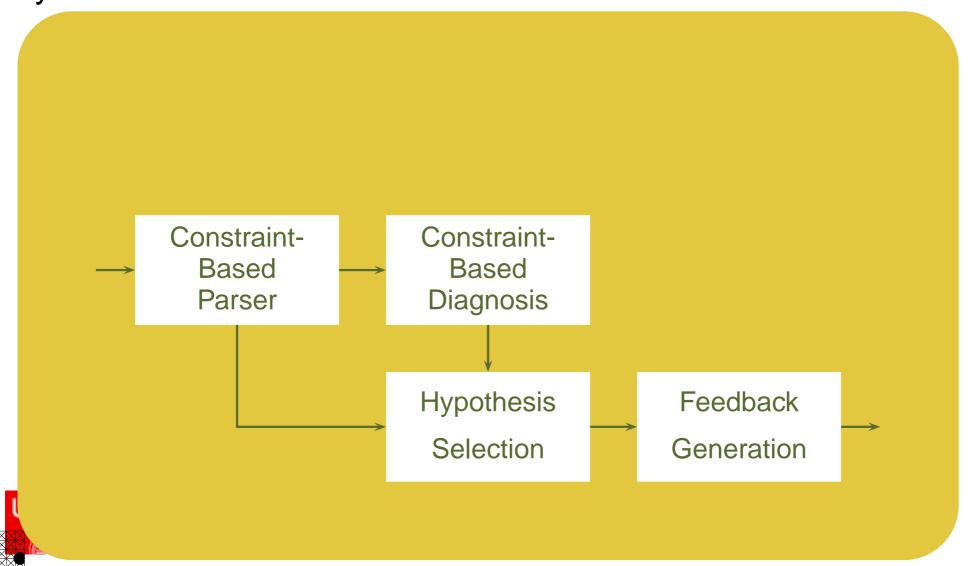


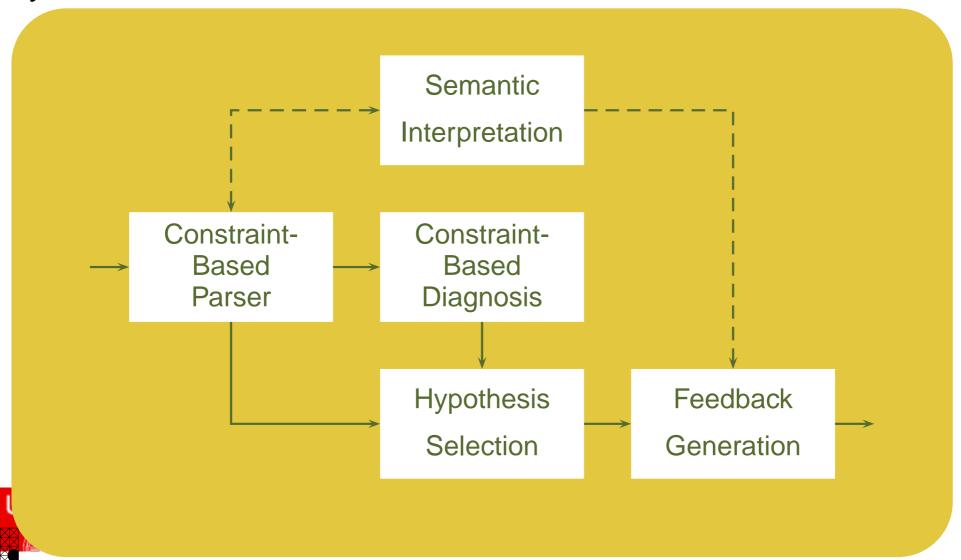
system architecture

Constraint-Based Parser

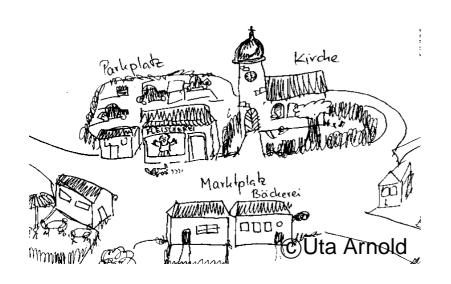








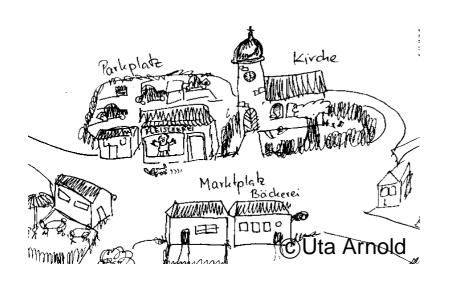
 semantic preferences, world knowledge and context information can be integrated into the optimisation process



```
in_front_of(church,marketplace).
left_of(church,parking_lot).
at(market_place,bakers)
count(church,1).
```



 semantic preferences, world knowledge and context information can be integrated into the optimisation process



in_front_of(church,marketplace). left_of(church,parking_lot). at(market_place,bakers) count(church,1).

• if the "world" contains just a single church, prefer the singular reading



Where does the constraining information come from?



- Where does the constraining information come from?
- simplifying assumption: the obedient student
 - provide a scenario and a task
 - assume the student complies with the given limitations
 - static scenarios
 - dynamic scenarios



- Where does the constraining information come from?
- simplifying assumption: the obedient student
 - provide a scenario and a task
 - assume the student complies with the given limitations
 - static scenarios
 - dynamic scenarios
- alternatively
 - let the student take the initiative
 - communicate with the student about her intentions



a static scenario: Meister Albrecht



a static scenario: Meister Albrecht

 (static) domain knowledge can be integrated into the error sensitive parsing



a static scenario: Meister Albrecht

- (static) domain knowledge can be integrated into the error sensitive parsing
- in rich scenarios the domain knowledge does not provide enough constraining information



a static scenario: Meister Albrecht

- (static) domain knowledge can be integrated into the error sensitive parsing
- in rich scenarios the domain knowledge does not provide enough constraining information
- dynamic scenarios allow to focus on the changing aspects of a scene (Reuer 2003)



 strongest constraints could be derived from the intention of the student



- strongest constraints could be derived from the intention of the student
- How to gain access to intentions?



- strongest constraints could be derived from the intention of the student
- How to gain access to intentions?
 - verbal: asking back

Do you mean several fish or only one?



- strongest constraints could be derived from the intention of the student
- How to gain access to intentions?
 - verbal: asking back

Do you mean several fish or only one?

non-verbal: select from a menue





more complex tasks



- more complex tasks
 - direct manipulation environment (HAMBURGER, 1995)



- more complex tasks
 - direct manipulation environment (HAMBURGER, 1995)

e.g. given:

- the cupboard
- the fridge
- the table
- the task: Tell me how to prepare breakfast.



- more complex tasks
 - direct manipulation environment (HAMBURGER, 1995)

e.g. given:

- the cupboard
- the fridge
- the table
- the task: Tell me how to prepare breakfast.
- in case of difficulties

Show me what you wanted to say!



- more complex tasks
 - direct manipulation environment (HAMBURGER, 1995)
 - e.g. given:
 - the cupboard
 - the fridge
 - the table
 - the task: Tell me how to prepare breakfast.
 - in case of difficulties
 - Show me what you wanted to say!
 - virtual world can become an alternative communication channel if the verbal communication breaks down

 available diagnostic techniques can produce a great variety of diagnostic information



- available diagnostic techniques can produce a great variety of diagnostic information
- precise and helpful feedback can be provided
 - as long as sufficient support from background information is available



- available diagnostic techniques can produce a great variety of diagnostic information
- precise and helpful feedback can be provided
 - as long as sufficient support from background information is available
- strongest support can be expected if the system has access to the student's intention



- available diagnostic techniques can produce a great variety of diagnostic information
- precise and helpful feedback can be provided
 - as long as sufficient support from background information is available
- strongest support can be expected if the system has access to the student's intention
- intentions can be elicited from the student by means of multi-modal interactions
 - verbal and non-verbal means can be used



- available diagnostic techniques can produce a great variety of diagnostic information
- precise and helpful feedback can be provided
 - as long as sufficient support from background information is available
- strongest support can be expected if the system has access to the student's intention
- intentions can be elicited from the student by means of multi-modal interactions
 - verbal and non-verbal means can be used



→ requires integrated system solutions