

SDRT 2: Glue

Julian J. Schlöder

Formal Pragmatics, Lecture 4, Jan 17th

Constructing Logical Form

Overview

- We have now:
- A language to express the truth-conditions of a discourse.
- A logic to express “typically, this is that”.
- We want:
- To formally derive the logical form of a discourse from its surface form.
- That is, to say “typically, this unstructured discourse has the following structure”.

Pragmatic Principles

- To construct logical form, we use pragmatic principles.
- Specific principles:
 - Postulates about under which conditions one typically assigns a discourse relation.
- A general principle:
 - Discourses are interpreted as to maximise coherence.
 - If a discourse is ambiguous as to its structure, choose the most coherent structure.

Missing Component

- We need a (formal) language to reason from **incomplete forms** to **complete (truth-conditional) forms**.
- We would like to keep the defeasible logic **decidable**.
 - Commonsense Entailment works on top of propositional logic.
 - We'll now use quantifier-free predicate logic (static).
- So, the **construction** does not have access to the truth conditions of what it constructs.
- This makes sense: speakers usually agree about meaning, but not so much about truth.
 - So we don't want to have the truth-conditions help construct meaning, unless particular ones have been agreed upon or can count as "world knowledge".

Underspecified Logical Form

- The idea is this: we construct a language for **incomplete descriptions of logical forms**.
- Make it so that a fully specified logical form is a **model** of a description.
- That is, if \mathcal{K} is a description for K ($\in \text{LIC}$), then $K \models \mathcal{K}$.

Describing Information Content

- So what are the bits and pieces of the LIC?
- DPL formulae
 - Take a constant symbol f_φ for each DPL wff φ .
 - (This is the “toy version” of underspecification)
- Labels
 - Take a constant symbol l_λ for each label λ .
 - Plus corresponding variable symbols l_1, l_2, \dots
- Discourse relations
 - Take a constant symbol D_R for each discourse relation R
 - Plus corresponding variable symbols D_1, D_2, \dots

Underspecification

- We underspecify:
 - What the contents are.
 - Which contents are connected.
 - How they are connected.
- Take two predicate symbols to describe assignment:
 - $labels(l, f)$
 - $relates(l_1, l_2, l_3, D)$
- And two to describe structure:
 - $outscopes(l_1, l_2)$
 - $accessible(l_1, l_2)$

Anaphora

- Anaphora are a type of underspecification.
- So take a constant symbol v_x for each variable in DPL (do this for every type of variable).
- And add a predicate symbol:
→ $anaphor(l, v)$
- (If you extend the language to describe DPL formulae, you can write anaphora as $x = ?$ to indicate something like “x is a free variable”.)

Examples

- ULFs are constructed from surface form.

(1) There is a woman.

$$labels(l_1, f_{\exists x. woman(x)})$$

(2) She runs.

$$labels(l_2, f_{run(y)}) \wedge anaphor(l_2, v_y)$$

(3) There is a woman. She runs.

$$\begin{aligned} & labels(l_1, f_{\exists x. woman(x)}) \\ & \wedge labels(l_2, f_{run(y)}) \wedge anaphor(l_2, v_y) \\ & \wedge relates(l_0, l_1, l_2, D) \end{aligned}$$

Cue Phrases

- Add an (empirically sourced) vocabulary of linguistic cues to this language.
- `[[therefore]] = therefore(I)`
- `[[and then]] = and-then(I)`
- `[[I hereby command]] = command(I)`
- `[[I hereby assert]] = inform(I)`
- Including grammatical features:
 - `indicative(I)`
 - `interrogative(I)`
 - `imperative(I)`
- Plus tense, aspect, anything useful from the grammar...

Reminder: SDRSs

- Recall: A **Segmented DRS** is a triple (Π, \mathcal{F}, L) such that:
- Π is a set of labels.
- $\mathcal{F} : \Pi \rightarrow \text{LIC}$ is a function mapping labels to LIC wffs
- $L \in \Pi$ (the “last” added label).

“Model Theory”: Assignment

- The underspecified language has the logical constants $=$, \neg , \vee and \wedge .
- Call a formulae in this language an ULF (underspecified logical form).
- Let $K = (\Pi, \mathcal{F}, L)$ be an SDRS and A be a function s.t.:
 - for each variable l_i , $A(l_i) \in \Pi$
 - for each variable D_i , $A(D_i)$ is some discourse relation.
 - $A(f_\varphi) = \varphi$, $A(l_\pi) = \pi$, $A(D_R) = R$, $A(v_x) = x$ for all formulae φ , labels π , relations R and DPL-variables x .

“Model Theory”: Satisfaction

- $K, A \models x = y$ iff $A(x) = A(y)$ (for any variables or constants x, y)
- $K, A \models \text{labels}(l, f)$ iff $A(f)$ is a conjunct of $\mathcal{F}(A(l))$
- $K, A \models \text{relates}(l_1, l_2, l_3, D)$ iff $A(D)(A(l_2), A(l_3))$ is a conjunct of $\mathcal{F}(A(l_1))$.
- $K, A \models \text{outscopes}(l_1, l_2)$ iff $A(l_2) < A(l_1)$.
- $K, A \models \text{accessible}(l_1, l_2)$ iff $A(l_1)$ is accessible from $A(l_2)$.
- $K, A \models \text{anaphor}(l, v)$ iff there is a DPL variable z introduced in some segment λ such that λ is accessible to $A(l)$ and $\mathcal{F}(A(l))$ has a conjunct $A(v) = z$.
- If $\text{cue}(l)$ is a linguistic cue predicate, $K, A \models \text{cue}(l)$ always.
- Negation, disjunction and conjunction as usual

Glue Language

- The Glue Language is obtained from the underspecified language by adding the connectives \rightarrow and $>$.
- Moreover, the Glue Language contains additional predicates for **world knowledge**.
- For instance $\text{cause}(f_\varphi, f_\psi)$ for " $\varphi \therefore \psi$ is a valid enthymeme".

Back-flow of semantic information

- The following are Glue logic axioms:

$$(relates(l_0, l_1, l_2, D_{Explanation}) \wedge labels(l_1, f_\varphi) \wedge labels(l_2, f_\psi)) \\ \rightarrow cause(f_\psi, f_\varphi).$$

$$(relates(l_0, l_1, l_2, D_{Narration}) \rightarrow before(l_1, l_2)).$$

- That is, we can set things up so that **deep semantic** information from the logic of information content is available as **shallow** information in the Glue logic.
 - ("shallow" because abstracted from semantic structure to constant symbols)
- We do this by encoding our knowledge about meaning postulates in such Glue axioms.

A More Complex Case

- Let $\text{occasion}(l_1, l_2)$ describe that the event labelled l_2 is occasioned by the one labelled l_1 .
- A **script for occasion** is a Glue formula of the following form:

$$(\text{relates}(l_0, l_1, l_2, D) \wedge \text{labels}(l_1, f_\varphi) \wedge \text{Info}(f_\varphi) \wedge \text{labels}(l_2, f_\psi) \wedge \text{Info}(f_\psi)) \\ > \text{occasion}(l_1, l_2)$$

A More Complex Case

- Let $\text{occasion}(l_1, l_2)$ describe that the event labelled l_2 is occasioned by the one labelled l_1 .
- A **script for occasion** is a Glue formula of the following form:

$$(\text{relates}(l_0, l_1, l_2, D) \wedge \text{labels}(l_1, f_\varphi) \wedge \text{Info}(f_\varphi) \wedge \text{labels}(l_2, f_\psi) \wedge \text{Info}(f_\psi)) \\ > \text{occasion}(l_1, l_2)$$

- One suggested by Asher & Lascarides:

$$\begin{aligned} & \text{relates}(l_0, l_1, l_2, D) \\ & \wedge \text{labels}(l_1, f_\varphi) \wedge \text{subformula}(f_\varphi, f_{\text{fall}}(e_1, x_1)) \\ & \wedge \text{labels}(l_2, f_\psi) \wedge \text{subformula}(f_\psi, f_{\text{help-up}}(e_2, x_2, x_3)) \\ & > \text{occasion}(l_1, l_2) \end{aligned}$$

This Seems Very Tedious

- The Big Problem of Formal Pragmatics: how do these things generalise?
- At the current state of research, we can describe mechanisms for pragmatic inference.
- But we need to hard code world knowledge, lexical knowledge etc.
- Part of our mechanisms is also a language for hard-coding.

Inferring Relations: sufficiency

- This is the “sufficiency principle” from our study of enthymemes:
- $labels(l_1, f_\varphi) \wedge labels(l_2, f_\psi) \wedge relates(l_0, l_1, l_2, R) \wedge cause(f_\psi, f_\varphi) > R = D_{Explanation}.$
- $labels(l_1, f_\varphi) \wedge labels(l_2, f_\psi) \wedge relates(l_0, l_2, l_1, R) \wedge cause(f_\psi, f_\varphi) > R = D_{Result}.$

Inferring Relations: sufficiency

- This is the “sufficiency principle” from our study of enthymemes:
- $labels(l_1, f_\varphi) \wedge labels(l_2, f_\psi) \wedge relates(l_0, l_1, l_2, R) \wedge cause(f_\psi, f_\varphi) > R = D_{Explanation}$.
- $labels(l_1, f_\varphi) \wedge labels(l_2, f_\psi) \wedge relates(l_0, l_2, l_1, R) \wedge cause(f_\psi, f_\varphi) > R = D_{Result}$.
- Hereinafter, I will make our lives a bit easier, where possible:
 - $R(\alpha, \beta) \wedge cause(\alpha, \beta) > R = Explanation$.
- Typical abbreviation in SDRT papers:
 - $\lambda : ?(\alpha, \beta) \wedge cause(K_\alpha, K_\beta) > \lambda : Explanation(\alpha, \beta)$.

Inferring Relations: lexical knowledge

- $R(\alpha, \beta) \wedge \text{occasion}(\alpha, \beta) > R = \textit{Narration}$.

(4) a. Max fell.
b. John helped him up.]-Narration

- $\text{cause}(\beta, \alpha) \rightarrow \neg \text{occasion}(\alpha, \beta)$

(5) a. Max fell.
b. John pushed him.]-Explanation

Inferring Relations: aspectual knowledge

(6) a. There is a man.
b. He knows that it is raining.]-Background

(7) a. There is a man.
b. He is walking.]-Background

$$- R(l_1, l_2) \wedge (\text{state}(l) \vee \text{activity}(l)) > \text{Background}(l_1, l_2)$$

Inferring Relations: logic-ish knowledge

- $R(\alpha, \beta) \wedge \text{subtype}(\alpha, \beta) > R = \textit{Elaboration}.$
- $\text{labels}(l_1, f_\varphi) \wedge \text{labels}(l_2, f_\psi) \wedge \text{relates}(l_0, l_1, l_2, R) \wedge \text{subtype}(f_\psi, f_\varphi) > R = D_{\textit{Elaboration}}.$
- Where $\text{subtype}(f_\psi, f_\varphi)$ means that any situation described by ψ can also be described as φ .
 - (Type theory helps us to push the boundaries of decidability a little bit)

Inferring Relations: logic-ish knowledge

- $R(\alpha, \beta) \wedge \text{subtype}(\alpha, \beta) > R = \text{Elaboration}.$
- $\text{labels}(l_1, f_\varphi) \wedge \text{labels}(l_2, f_\psi) \wedge \text{relates}(l_0, l_1, l_2, R) \wedge \text{subtype}(f_\psi, f_\varphi) > R = D_{\text{Elaboration}}.$
- Where $\text{subtype}(f_\psi, f_\varphi)$ means that any situation described by ψ can also be described as φ .
 - (Type theory helps us to push the boundaries of decidability a little bit)
- A heuristic approach: for any n , \vdash^n ("FOL-provable in n or less steps") is decidable. This can be encoded as a predicate.
 - (A computational approach would use an automated theorem prover with a time limit)

Inferring Relations: Cue Phrases

- Monotonic cues:

$(R(\alpha, \beta) \wedge \text{therefore}(\alpha)) \rightarrow R = \textit{Result}$

$(R(\alpha, \beta) \wedge \text{and-then}(\alpha)) \rightarrow R = \textit{Narration}$

- Performatives:

$\text{inform}(\pi) \rightarrow ((R(?, \pi) \wedge \text{right-veridical}(R)) \vee (R(\pi, ?) \wedge \text{left-veridical}(R)))$.

- Defeasible cues:

$\text{indicative}(\alpha) > \text{inform}(\alpha)$

Inferring Relations: Rationality Principles

- It is rational to try to interpret a response to a question as an answer:

$$(R(\alpha, \beta) \wedge \text{interrogative}(\alpha) \wedge \text{spk}(\alpha) \neq \text{spk}(\beta)) > R = \text{IQAP}$$

(8) a. A: Is John going out tonight?
b. B: I saw him get dressed earlier. }]-IQAP

(9) a. A: Why is seaweed good for you?
b. B: Lots of vitamins. }]-IQAP

Construction of Discourse (overview)

- A context may contain underspecifications, or things that can be revised.
- Thus, the context is a big ULF formula Γ (possibly empty).
 - Alternatively, let the context be set of SDRSs. Then define Γ to be the set of all ULFs that are true for all contextual SDRSs.
- Now, let \mathcal{K} a ULF representing new information. Let π be a label not used in Γ . Then define:
- $update(\Gamma, \pi : \mathcal{K})$ is the set of all (and only) those SDRSs where $L = \pi$ and that satisfy the defeasible consequences of attaching \mathcal{K} to some available segment α in Γ .
- If $update(\Gamma, \pi : \mathcal{K}) = \emptyset$, then \mathcal{K} is incoherent in Γ .

Construction of Discourse (formal)

- Let π be a label not used in Γ .
- Let R_n, l_1 and l_2 be variables not used in Γ .
- Let α be the “last” label in Γ (i.e. the π from the last update).
 - Can also define this as the “accessibility-minimal” label.
- Then: $K \in \text{update}(\Gamma, \pi : \mathcal{K})$ iff
 $K = (\Pi, \mathcal{F}, L)$, is a coherent SDRS with $L = \pi$, and for all formulae φ of the underspecified language: (if $\Gamma \neq \emptyset$):
If $\Gamma \wedge \mathcal{K} \wedge \text{relates}(l_1, l_2, l_\pi, R_n) \wedge \text{accessible}(l_2, l_\alpha) \vdash \varphi$, then $K \models \varphi$.
(if $\Gamma = \emptyset$): If $\mathcal{K} \vdash \varphi$, then $K \models \varphi$.

Maximise Discourse Coherence

- There may be a *lot* of SDRSs in $update(\Gamma, \pi : \mathcal{K})$.
- We want to pick out the “best” ones.
- Intuitively, some ways of structuring a discourse “tell a better story” than others.
- We’ll call the good ones “most coherent” and formalise conditions on what such coherence might be.

MDC

An SDRS K is at least as coherent as an SDRS K' , $K' \leq^c K$, if and only if all of the following hold:

1. *Prefer consistency*: If K' is consistent, then so is K .
2. *Prefer rich structure*: K has at least as many coherence relations as K' .
3. *Prefer resolution*: K binds (over accommodates) at least as many presuppositions as K' does.
4. *Prefer better relations*: For every rhetorical relation $R(\pi_1, \pi_2)$ that K' and K share: $R(\pi_1, \pi_2)$ is at least as coherent in K as it is in K' .
5. *Prefer flat structure*: K has at most as many labels as K' unless K' has a *semantic clash* and K does not.

- A semantic clash is a conflict of veridicality.

- (10)
- a. π_1 : If a shepherd goes to the mountains,
 π_2 : he normally brings his dog.
 π_3 : He brings a good walking stick too.
 - ✓b. π_0 : *Consequence*(π_1, π)
 π : *Parallel*(π_2, π_3)
 - ✗c. π_0 : *Consequence*(π_1, π_2) \wedge *Parallel*(π_2, π_3)

MDC: Quality of Relations

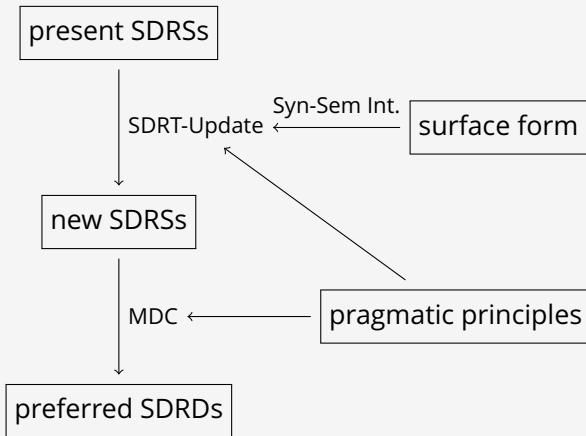
- Some Contrasts sound better than others; some Parallels sound better than others.

(11) a. John loves opera, but hates musicals.
??b. John loves opera, but hates rap music.

(12) a. John loves opera and likes musicals, too.
??b. John loves opera and likes to go swimming, too.

Implicature

- Anything entailed by the most coherent SDRs (might be multiple) is **implicated**.



MDC: Lexical Disambiguation

- *bank* can be *financial institution* and *area near water*.

(13) a. Sue was wondering where the fisherman is.
b. Max said he was out getting cash.
c. She found him at the bank (*financial institution*).

(14) a. Sue was wondering where the fisherman is.
b. Max said he was out getting cash.
c. But she found him at the bank (*area near water*).

(15) a. Sue was wondering where the fisherman is.
b. Max said he was out getting cash (*at an ATM*).
c. But she found him at the bank (*financial institution*).

Constructing Logical Form

Case Study

Constructing Logical Form

Case Study

Attachment of *Why*?

- This is again some of my own work.

(16) a. Richard: They'll check every single doctor.
b. Anon Why is that? $\left. \vphantom{\begin{array}{l} a. \\ b. \end{array}} \right] \text{-Explanation}_q$

(17) a. Gillian: Do you want mum to come to Argos?
b. Robert: Why are you asking me? $\left. \vphantom{\begin{array}{l} a. \\ b. \end{array}} \right] \text{-Explanation}_q^*$

Bare *Why*?

(18) a. Brenda: He's in hospital.

b. Carla: Why?

[*Why is he in hospital?* ~~*Why are you telling me?*~~]

c. Brenda: Because he's not very well

(19) a. Anon: Do you love me $\langle \text{unclear} \rangle$?

b. Bnon: Why?

[*Why are you asking?*]

c. Anon: $\langle \text{unclear} \rangle$ I love you so much.

- It seems that assertoric antecedents are Explanation_q and non-assertoric antecedents are Explanation_q^* .

Counterexample

- (With thanks to Robin Cooper)

(20) a. Amy: I'll have you know that I'm upset.

b. Bob: Why?

[*Why are you upset?* OR *Why are you saying that?*]

c. Amy: I had a terrible day at work.

(21) a. Amy: I'll have you know that I'm upset.

b. Bob: Why?

[*Why are you upset?* OR *Why are you saying that?*]

c. Amy: So you be careful around me today.

- So, sometimes, bare *Why?* is ambiguous with assertoric antecedents.

Performatives Matter

- (22) a. Amy: I'm upset.
b. Bob: Why?
[*Why are you upset?* ~~OR *Why are you saying that?*~~]
c. Amy: I had a terrible day at work.

- (23) a. Amy: I'm upset.
b. Bob: Why?
[*Why are you upset?* ~~OR *Why are you saying that?*~~]
#c. Amy: So you be careful around me today.

– So, the *I'll have you know* (\approx *I am hereby telling you*) matters.

Coding it in the Glue Logic

- Introduce a predicate prop in the underspecified language such that:
 $K, A \models \text{prop}(I)$ iff the content labelled by $A(I)$ is a propositional formula (not a question or command).
- Take *Why?* to be a monotonic linguistic cue for
 $(R = \text{Explanation}_q \vee R = \text{Explanation}_q^*)$

Inferring the Right Relation

Glue Axioms for *Why*?

- a. $(R(\alpha, \pi) \wedge (R = \textit{Explanation}_q \vee R = \textit{Explanation}_q^*) \wedge \textit{prop}(\alpha) > R = \textit{Explanation}_q).$
- b. $(R(\alpha, \pi) \wedge (R = \textit{Explanation}_q \vee R = \textit{Explanation}_q^*) \wedge (\textit{inform}(\alpha) \vee \textit{interrogative}(\alpha) \vee \textit{imperative}(\alpha)) > R = \textit{Explanation}_q^*).$

– Both (a) and (b) apply for *I'll have you know that p*.

World Knowledge, again

- (with thanks to Jonathan Ginzburg)
- World knowledge can override these defaults.

(24) a. Amy: You're upset.

b. Bob: Why?

~~[Why am I upset?]~~ OR *Why are you saying that?*

- Only I have knowledge of my internal states.
- Amy cannot *know* that Bob is upset, let alone *why*.
- This rules out Explanation_q .

Case Study

Presupposition

Case Study

Presupposition

Presupposition Triggers (1)

- I use ∂ to denote presuppositions.

(25) John *knows* that it is raining.
 ∂ It is raining.

(26) John *realised* that it is raining.
 ∂ It is raining.

(27) John *stopped* smoking.
 ∂ John smoked.

(28) John *started* smoking.
 ∂ John didn't smoke.

Presupposition Triggers (2)

(29) John is smoking *again*.
 ∂ John smoked before.

(30) John *started* to smoke *again*.
 ∂ John smoked once, then didn't.

(31) *It was* John, *who* stole the cookies.
 ∂ Someone stole the cookies.

(32) John's son is bald.
 ∂ John has a son.

(33) *The* king of France is bald.
 ∂ there is a king of France

Presupposition Projection

(34) If John is bald, then John's son is bald.
 ~~∂ John has a son.~~

(35) If John has a son, then John's son is bald.
 ~~∂ John has a son.~~

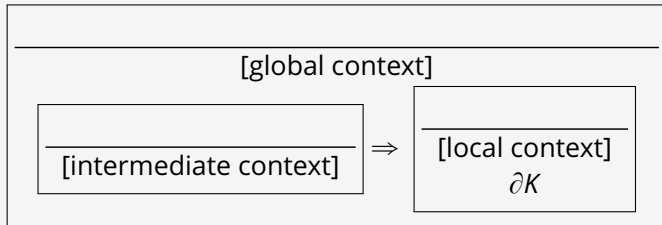
(36) If John went diving before, he'll bring his wetsuit.
 ~~∂ John has a wetsuit.~~

Binding and Accommodation

- It is a syntax-semantics interface job to produce a presupposition.
- But it is a pragmatics job to resolve it.
- If the context entails the presupposed content, our life is easy: we can just delete it.
- If not, then we need to **add** the content somewhere in the context.

Accommodation: Options

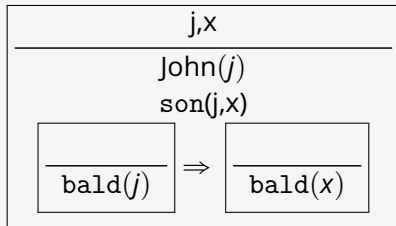
- Traditional division:



- In SDRT: any accessible segment

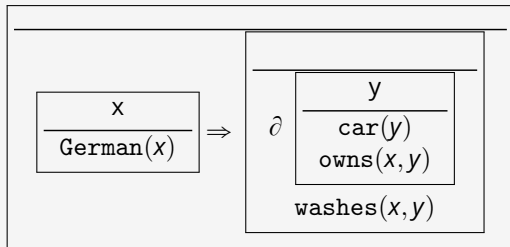
Global Accommodation

(37) If John is bald, then John's son is bald.



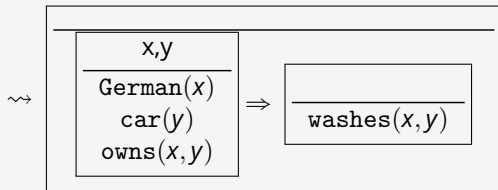
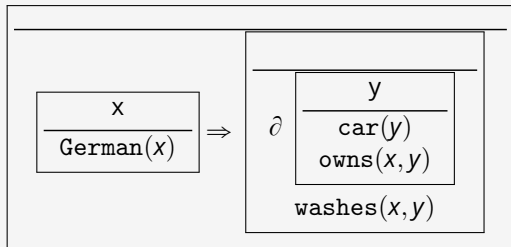
Intermediate Accommodation

(38) All Germans wash their cars.



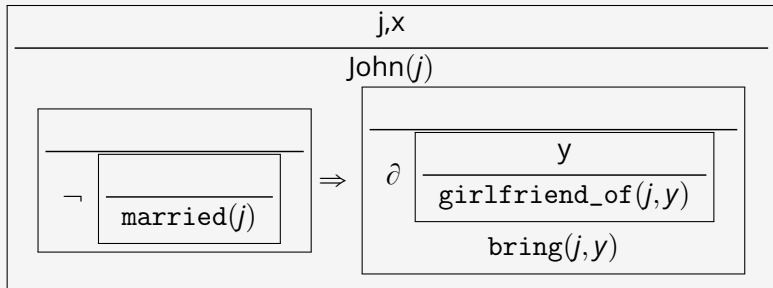
Intermediate Accommodation

(38) All Germans wash their cars.



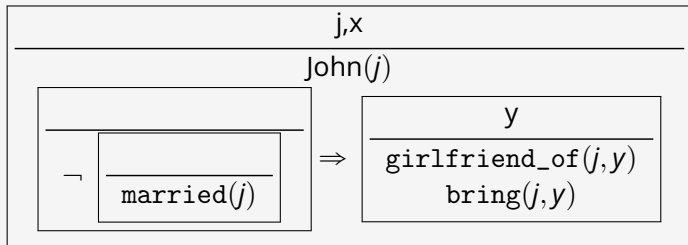
Local Accommodation

- (39) John said he'd bring someone.
If John is not married, he will bring his girlfriend.



Local Accommodation

- (40) John said he'd bring someone.
If John is not married, he will bring his girlfriend.



DRT: informativeness + consistency

- The DRT story: accommodate as **globally as possible** where **consistent** and **informativeness is preserved**.

(41) Either this house has no bathroom, or the bathroom
is in an odd place.
~~∅ there is a bathroom~~

- Global accommodation would render the first disjunct uninformative.

Trouble for DRT

(42) Either John didn't make cookies, or Mary stole the cookies.

∂ ~~there are cookies~~

- Informativeness doesn't apply here: accommodating globally that *there are cookies* does not render *John didn't make cookies* uninformative.

(43) John had an accident.

The car hit him.

∂ there was a car

- Informative + Consistent.

- Presuppositions need to be **coherent** in their context.

(44) John had an accident.
?? There was a car.

- So, SDRT presupposition is simple:
 - Except an odd construct with FBP to make anaphora really work.
- If the grammar produces $\mathcal{K}_1 \partial \mathcal{K}_2$ from a clause, update first with \mathcal{K}_2 and then with \mathcal{K}_1 (except in null contexts).
- *Binding* is just attachment as *Consequence*.
- *Accommodation* is attachment as anything else.
 - typically *Background*

MDC can override global binding 1

- (45) a. π_1 : If John went diving before,
 π_2 : he'll bring his wetsuit.
 π_3 : John has a wetsuit.
- ✓b. π_0 : *Consequence*(π, π_2)
 \wedge *Def-Consequence*(π_1, π_3) \wedge *Background*(π_2, π_3)
- ✗c. π_0 : *Consequence*(π_1, π_2)
 π : *Background*(π_0, π_3)

– MDC: more relations, flatter structure.

≈ If John went diving before then he owns a wetsuit and will bring it.

MDC can override global binding 2

- (46) a. π_1 : Either John didn't make cookies,
 π_2 : or Mary stole the cookies.
 π_3 : there are cookies.
- ✓b. π_0 : *Contrast*(π_3, π_2) \wedge *Narration*(π_3, π_2)
 π : *Alternation*(π_1, π_0)
- ✗c. π_0 : *Alternation*(π_1, π_2)
 π : *Background*(π_0, π_3)

– MDC: more labels.

\approx Either John didn't make cookies, or there are cookies but Mary stole them.

- Reading for Friday:
- Hunter, J & Abrusán, M. (2017). Rhetorical Structure and QUDs. In: *JSAI International Symposium on Artificial Intelligence*.
- And have a think about papers for the reading group.
- I suggest *Strategic Conversation* (Asher & Lascarides, 2013) for something on noncooperative dialogue.
- And I'll pick out a nice (informal) paper on irony.