Assessing the Reliability of an Annotation Scheme for Indefinites

Maria Aloni & Raquel Fernández
ILLC/University of Amsterdam
m.d.aloni/raquel.fernandez@uva.nl

First meeting MoL January Projects

Amsterdam, 5-1-2011

Short description

- Research in (computational) linguistics and cognitive science often relies on subjective judgements
- Various techniques have been developed to assess reliability of subjective judgements by measuring the extent to which several subjects agree on their judgements
- In this project:
  - we study methods to assess reliability of linguistic judgements;
  - apply them to evaluate a methodology developed for a corpus study on indefinite descriptions
- The MoL project will consist in
  1. designing an evaluation experiment for the methodology of the corpus study;
  2. annotating some data according to the proposed design;
  3. evaluating the results by calculating inter-annotator agreement and by drawing conclusions on the reliability of the methodology under investigation.

Preliminary schedule

- **1st week:**
  - Introduction to the linguistic phenomenon: range of functions that indefinite expressions can fulfill and possible hypotheses as to their synchronic and diachronic distribution across different languages.
  - Design of the annotation experiment: formulation of guidelines and decision trees for the elicitation of judgements; selection of the data to be annotated.
- **2nd week:**
  - Design of a web interface for carrying out the annotation; short pilot experiment.
  - Annotation of one single data set by several independent annotators.

Preliminary schedule

- **3rd week:**
  - Introduction to different methods for measuring inter-annotator agreement; students with some programming knowledge will be encouraged to write scripts to measure agreement automatically.
  - Calculation of inter-annotator agreement.
- **4th week:**
  - Evaluation of the results obtained and writing up of a final (joint) paper.

Plan for today

- Intro on indefinites
- Training on annotation: *any* and *some*
- Division of labour: 3 tasks for next Tuesday
  1. Read and present one or two papers on *any* or *some* (everybody)
  2. Write (first version) guidelines for annotation (3 people)
  3. Preparation of the data for pilot (2 people)

Why indefinites?

Their study has inspired among the most important theoretical breakthroughs in logical semantics

- Classical analysis (Montague style): Indefinites as existentially quantified NPs

  (1) a. some man $\mapsto \lambda P(\exists x(man(x) \wedge P(x)))$
  b. some man walks $\mapsto \lambda P(\exists x(man(x) \wedge P(x)))(walk) = \exists x(man(x) \wedge walk(x))$

- Divergence in meaning between *some* and $\exists \mapsto$ original motivation for Grice's notion of an implicature

  (2) *Some of the deaths* were accidental.
  a. Intuitive interpretation: some and not all
  b. Classical logical rendering: some and perhaps all
  c. Gricean implicature: not all

Why indefinites?

- Discourse properties of indefinites $\mapsto$ FCS, DRT, dynamics

  (3) *A/*every man* walks. He whistles.

  (4) If *a/*every farmer* owns *a/*every donkey*, he beats it.

- Exceptional scope of indefinites $\mapsto$ choice-functional analyses

  (5) Every linguists studied every analysis that solves *some/*most problem*.

  a. narrow scope: $\forall\forall\exists$ (unproblematic)
  b. wide scope: $\exists\forall\forall$ (referential-quant. ambiguity?)
  c. intermediate scope: $\forall\exists\forall$ (no!)

- Branching readings $\mapsto$ game-theoretical semantics (Hintikka)

  (6) *Some relative* of each villager and *some relative* of each townsman hate each other.

- Variety of indefinites $\mapsto$ ...

The variety of indefinites

- Indefinite (roughly): an expression with indefinite reference
- Wealth of indefinite forms:
  - English: *a*, *some*, *any*, ...
  - German: *ein*, *irgendein*, ...
  - Dutch: *een*, *enige*, *een of andere*, ...
  - Italian: *uno*, *nessuno*, *(un) qualche*, *(uno) qualsiasi*, ...
  - Lezgian: *sa wuc jat'ani*, *x̂ajit'ani*, ...
  - Russian: *-nibud*, *-to*, *ljuboj*, ...
- Open questions:
  1. Why so much variety in indefinite forms?
  2. What is their common core? What is specific to each of them?
- Core idea of my VIDI project: Via fossilization a principled answer to question 1
- Two recent answers to question 2: Kratzer's alternative semantics & Chierchia's implicature account
- But first: is *any* really an indefinite?

Basic facts on *any*

- Restricted distribution:

  (7) #I saw any pigs.

- And two main uses:

  (8) a. I didn't see any pigs. [Negative Polarity (NPI)]
  b. $\neg\exists x\phi$

  (9) a. I can catch any raven. [Free Choice (FC)]
  b. $\forall x\Diamond\phi$

  (10) I cannot catch any raven.
  a. $\neg\Diamond\exists x\phi$ NPI
  b. $\neg\forall x\Diamond\phi$ FC

- Various analyses (based on Horn 2004):
  - Unitarian-universalist approach
  - Ambiguist view
  - The indefinite analysis

Unitarian-universalist approach

- *Any* as wide scope universal. Eg. Quine (1960).
- More precisely, a wide scope universal taking immediate scope wrt its *trigger* or *licensor* (Horn 1972):

(11) a. I didn't see any pigs.
b. $\forall x \neg \phi$

(12) a. I can catch any raven.
b. $\forall x \Diamond \phi$

(13) I cannot catch any raven.
a. $\forall x \neg \Diamond \phi$ NPI
b. $\neg \forall x \Diamond \phi$ FC

- Advantages: parsimonious ('Senses are not to be multiplied beyond necessity' Grice 1989)
- Disadvantage: empirically flawed

Against universalism

- Genuine existential character of *any*:

(14) I wonder if Susan married anybody. [Fauconnier 1979]

a. I wonder if ($\exists x$ Mary married x)

b. #$\forall x$ (I wonder if Mary married x)

(15) Maria rarely fails any student. [Higginbotham]

a. Rarely it is the case that ($\exists x$ Maria fails x)

b. #$\forall x$ (rarely it is the case that Maria fails x)

Double nature of *any*

- FC *any* compatible with A-adverbs (absolutely/almost), which consort with universals and tend to exclude existentials:

(16) a. Almost (everybody/anybody/# somebody) can swim.
b. #I didn't talk to almost anybody.
c. Can almost anybody swim? [FC only]

- NPI *any* felicitous in *there*-insertion contexts, notorious for affinity to existentials:

(17) a. There is (somebody/# everybody) in the garden.
b. There isn't anybody that can swim. [NPI only]
c. #There is anybody that can swim.

Ambiguist view

- Two *any*'s (e.g. Dayal 1998):
  - NPI *any*: an existential in monotonic decreasing contexts
  - FC *any*: a wide scope universal in modal or characterizing statements
- Problems: FC *any* less universal than it first appeared:

(18) a. Pick any card! [Horn 1972]
b. To continue, push any key! [Giannakidou 2001]

The indefinite analysis

- Both *any*'s must be indefinites (Partee 1986)
- Both *any*'s differ from universals and existentials:

(19) a. I didn't see anybody/# everybody/#somebody whatsoever. NPI
b. Anybody/# everybody/#somebody whatsoever can come to the party. FC

- Various implementations:
  - Scalar analysis (Fauconnier, Horn): *any* as minimal element in a scale
  - Widening and strengthening analyses (Kadmon & Landman): *any* existential with two pragmatic characteristics:
    1. It widens the domain;
    2. is licensed only if domain widening leads to a stronger statement.
  - **Alternative semantics analyses** (Menendez-Benito, Aloni)
  - **'Obligatory implicature' accounts** (Chierchia)
  - See also Giannakidou, Tovena and Jayez, ...

Alternative semantics for indefinites

- GOAL: Explain variety of indefinite forms
  - What is their common core?
  - What is specific to each of them?
- MAIN IDEAS (Kratzer & Shimoyama 02)
  - Indefinites 'introduce' sets of propositional alternatives;
  - These are bound by propositional operators: [∃], [∀], [¬], [Q];
  - Different indefinites associate with different operators.
- EXAMPLES

(20)
a. [∃] (someone fell)
b. [Q] (who fell)
c. [∀] (anyone$_{FC}$ fell)
d. [¬] (anyone$_{NPI}$ fell)

e.

(only) d_1 fell	(only) d_2 fell	...

Obligatory implicature framework

- GOAL: Explain variety of indefinite forms
  - What is their common core?
  - What is specific to each of them?
- MAIN IDEAS (e.g. Chierchia 2010)
  - Indefinites are existential NPs;
  - Existential sentences give rise to many implicatures (generated by reasonings on alternative forms)
  - Differences between indefinite forms in terms of different alternative forms they generate or implicature mechanisms they allow/force, ...
- EXAMPLES

(21)
a. FCI/NPI *any*: exhaustified D-alternatives + weak exhaustification
b. FCI *qualunque*: exhaustified D-alternatives + strong exhaustification
c. NPI *ever*: simple D-alternatives + weak exhaustification

Division of tasks

Read and present on 11-1-11 one of the following (pair of) papers:

- On **any**:
  1. Gilles Fauconnier (1975) Pragmatic Scales and Logical Structure. Linguistic Inquiry 6. [**scalar analyses**]
     Larry Horn (2005) Airport '86 revisited: Toward a unified indefinite *any*. In G. Carlson & F.J. Pelletier (eds.) The Partee Effect.
  2. Nirit Kadmon & Fred Landman (1993). *Any*. Linguistics and Philosophy 16. [**widening & strengthening**]
     Maria Aloni (2007). Free Choice, Modals and Imperatives. Nat Lang Semantics 15.
  3. Paula Menendez-Benito (2010) On Universal Free Choice Items. Nat Lang Semantics 18. [**alternative semantics**]
     Maria Aloni (2007) Free choice and exhaustification: an account of subtrigging effects. In Estela Puig-Waldmueller (ed.), Proceedings of Sinn und Bedeutung 11.
  4. Chierchia (2010) Universal FC chapter [**implicatures**]
- On **some**:
  5. Donka Farkas (2002) Varieties of Indefinites, in Proceedings of SALT 12.

Corpus studies on indefinites: Motivation

- **Formal pragmatics:** Use of plain indefinites (e.g. *somebody*) can give rise to different pragmatic effects:
  - *Free choice implicature*: each individual is a permissible option (E.g. 'You may invite somebody')
  - *Ignorance implicature*: speaker doesn't know who (E.g. 'Somebody called')
  - ...
- **Typology:** Many languages have developed specialized forms for such enriched meanings:
  - *Free choice indefinites*: Italian *-unque*-series, Czech *koli*-series,
  - *Ignorance indefinites*: Russian *to*-series, German *irgend*-series,
  - ...

Corpus studies on indefinites: Motivation

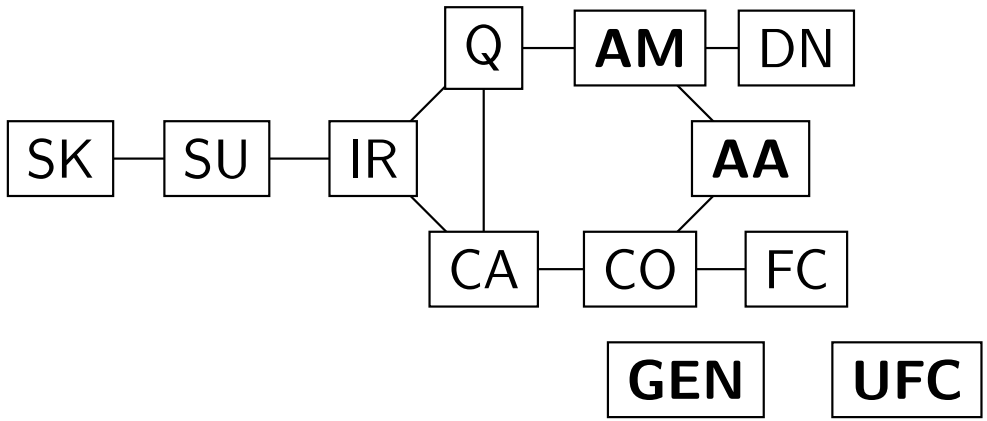
- **Formal pragmatics:** Use of plain indefinites (e.g. *somebody*) can give rise to different pragmatic effects:
  - *Free choice implicature*: each individual is a permissible option (E.g. 'You may invite somebody')
  - *Ignorance implicature*: speaker doesn't know who (E.g. 'Somebody called')
  - ...
- **Typology:** Many languages have developed specialized forms for such enriched meanings:
  - *Free choice indefinites*: Italian *-unque*-series, Czech *koli*-series,
  - *Ignorance indefinites*: Russian *to*-series, German *irgend*-series,
  - ...
- Main hypothesis: Different indefinites as conventionalization (or fossilization) of different pragmatic effects

Corpus studies on indefinites: Motivation

- **Formal pragmatics:** Use of plain indefinites (e.g. *somebody*) can give rise to different pragmatic effects:
  - *Free choice implicature*: each individual is a permissible option (E.g. 'You may invite somebody')
  - *Ignorance implicature*: speaker doesn't know who (E.g. 'Somebody called')
  - ...
- **Typology:** Many languages have developed specialized forms for such enriched meanings:
  - *Free choice indefinites*: Italian *-unque*-series, Czech *koli*-series,
  - *Ignorance indefinites*: Russian *to*-series, German *irgend*-series,
  - ...
- Main hypothesis: Different indefinites as conventionalization (or fossilization) of different pragmatic effects
- Main objective of corpus studies: Full understanding of
  - what is fossilized (synchronic) [languages: Ge,Cz,It,Sp,Du]
  - how it happened (diachronic) [languages: Ge,Sp,Du]

Corpus study on indefinites: Methodology

- Classification of randomly selected occurrences of indefinites according to a number of categories
- Starting point: Haspelmath's functional map



- Some annotated examples:

(22) *Somebody* called. Guess who? (SK)

(23) I heard *something*, but I couldn't tell what it was. (SU)

(24) In Freiburg the weather is nicer than *anywhere* in Italy. (CO)

(25) John didn't see *anybody*. (DN)

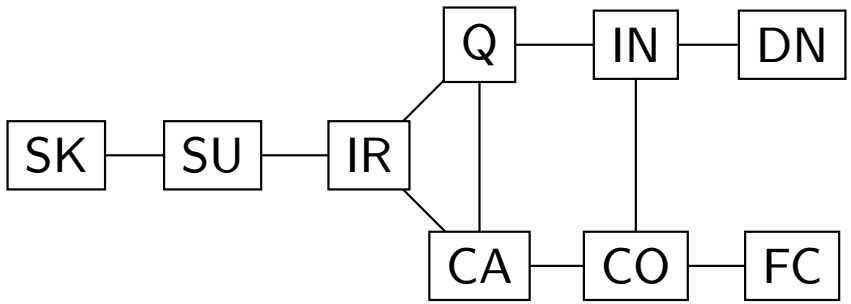
- Diagnostic tests used during annotation organized in a decision tree
- Main goal of MoL project: evaluation of this methodology via measurement of inter-annotator agreement

Haspelmath's nine functions

Haspelmath's (1997) typological study identifies nine core functions for indefinite forms:

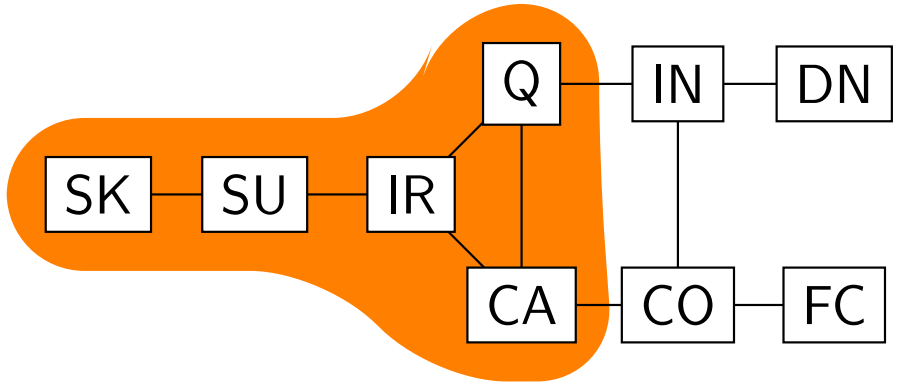
(26)	*Somebody* called. Guess who?	(SK: specific known)
(27)	I heard *something*, but I couldn't tell what it was.	(SU: sp. unknown)
(28)	You must try *something* else.	(IR: irrealis non-specific)
(29)	Did *anybody* call?	(Q: question)
(30)	If Joe sees *anything*, he will call.	(CA: antecedent of conditional)
(31)	John is taller than *anyone* else.	(CO: comparative)
(32)	It is not necessary that *anybody* call.	(IN: indirect negation)
(33)	John didn't see *anybody*.	(DN: direct negation)
(34)	*Anybody* can solve this problem.	(FC: free choice)

Haspelmath's functional map



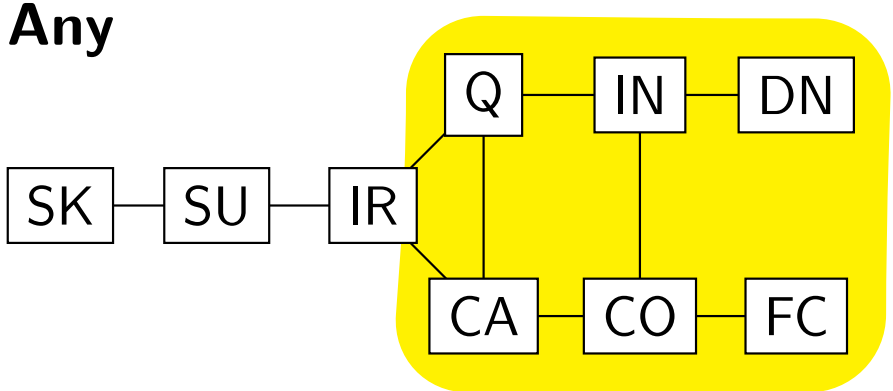
- An indefinite form will always express a set of functions that are contiguous on the map;
- Items which acquire new functions will develop first those functions that are contiguous to the original function.

- **Some**



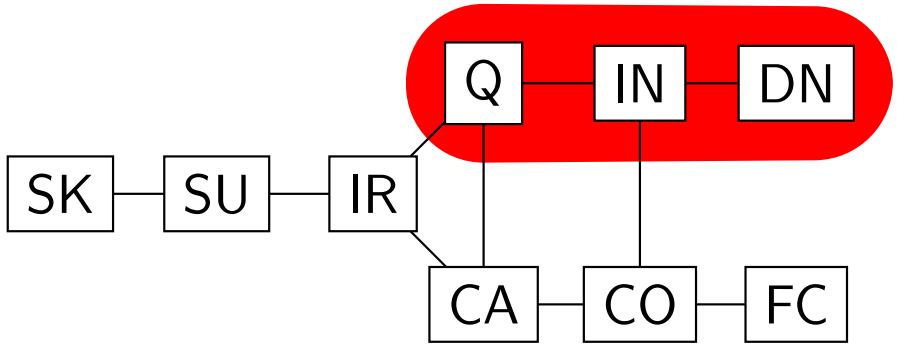
(35) #CO: John is taller than somebody else ≠ John is taller than anybody else

- **Any**

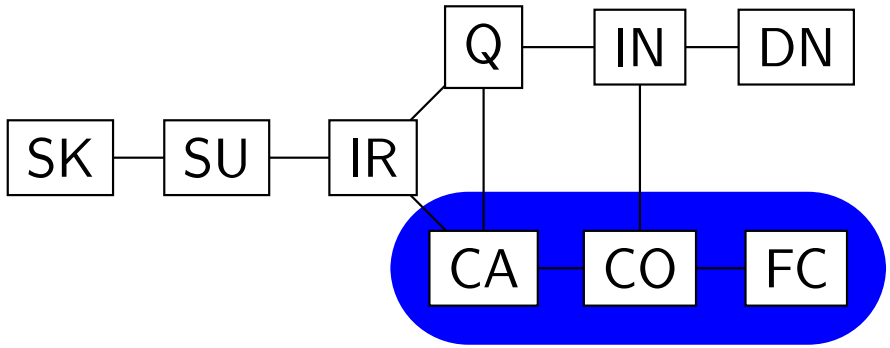


(36) #IR: # You must try *anything* else.

- Italian **nessuno**



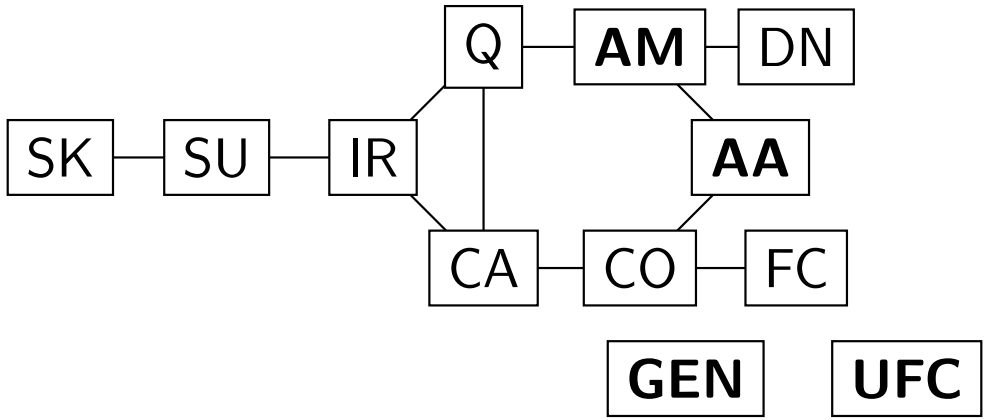
- Italian **qualunque**



Challenge for semanticists

- Try to identify the semantic properties that underlie the implicational universals;
- For each indefinite in each language,
  - either there is a single generalization that can account for what unites its possible functions;
  - or the series has to be split into multiple homonymic series (e.g. any-1 and any-2).

Our extended map

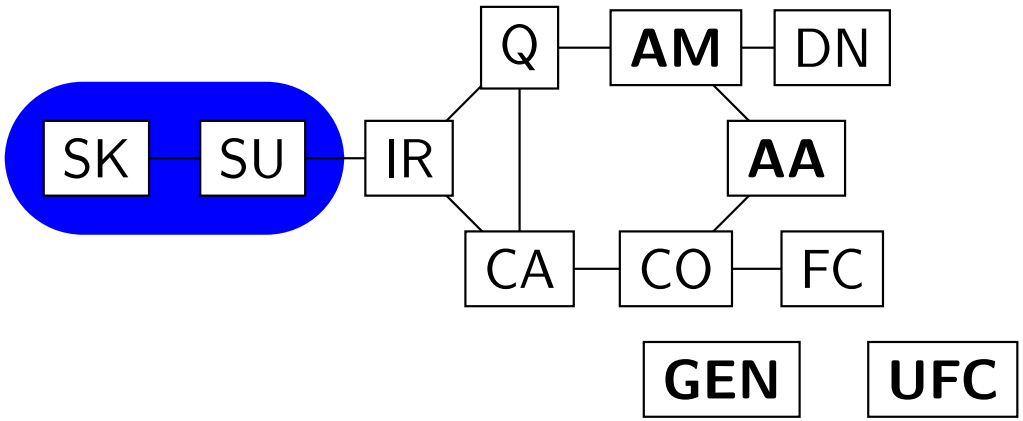


(37) *New functions on the map*

Abbr	**Label**	**Example**
AM	anti-morphic	I don't think that *anybody* knows the answer.
AA	anti-additive	The bank avoided taking *any* decision.
UFC	universal free choice	John kissed *any* woman with red hair.
GEN	generic	*Any* dog has four legs.

Specific–non specific

- Specificity area:



- Continuation test: (. . . indefinite$_i$ . . .). (. . . pronoun$_i$ . . .)

(38) SK/SU: I heard *something*. It was very loud. [specific]

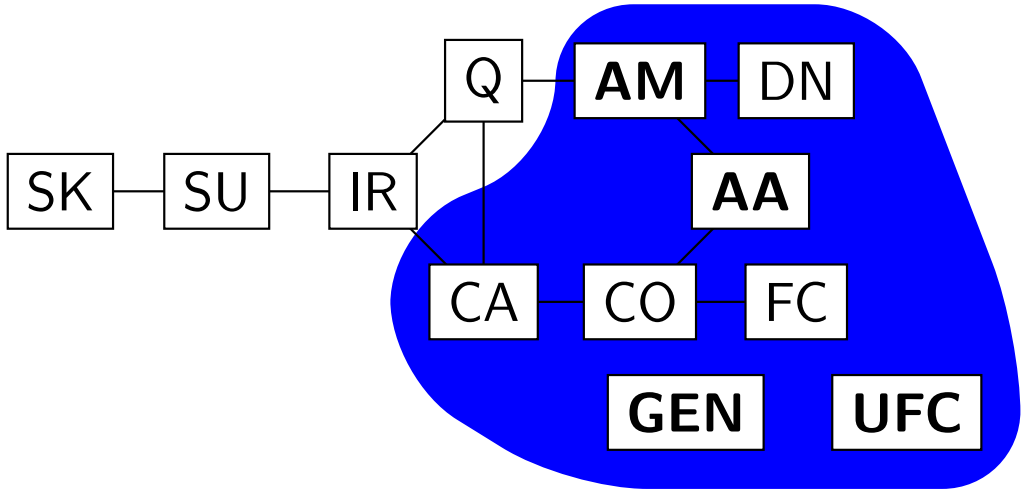
(39) IR: You must try *something* else. # It is very nice. [non specific]

- Standard Analysis (Farkas' *scopal specificity*):

(40) a. Specific uses: wide scope existential
b. Non-specific uses: narrow scope existential

Existential–wide scope universal

- Wide scope universal area:



- Test: **Op** (. . . indefinite . . .) $\Rightarrow \forall x$ (**Op**. . . x . . .)

(41) IR: You must try *somewhere* else $\not\Rightarrow$ for all places x: you must try x [NO]

(42) Q: Did *anybody* tell you anything about it? $\not\Rightarrow$ for every x: did x tell you about it? [NO]

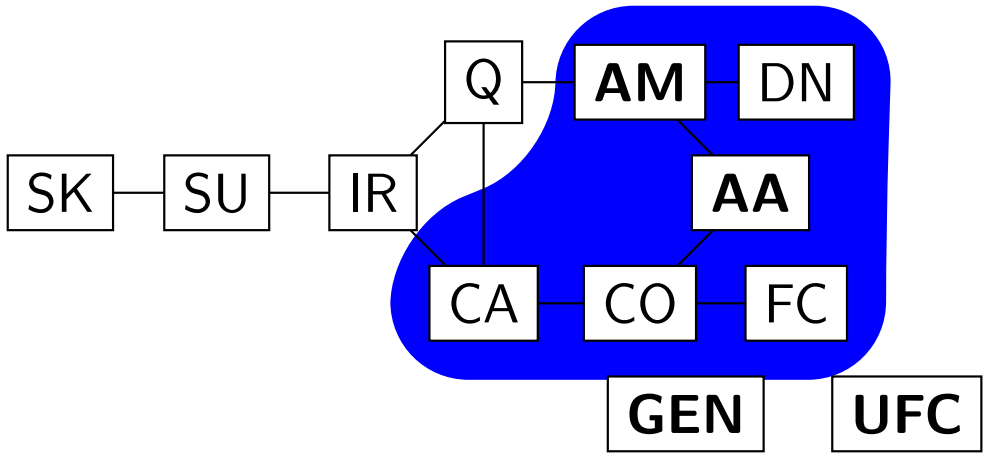
(43) DN: I didn't see *anybody* $\Rightarrow$ for every x: I didn't see x [YES]

(44) FC: You may kiss *anybody* $\Rightarrow$ for every x: you may kiss x [YES]

(45) CA: If you see *anybody*, tell me immediately $\Rightarrow$ for every x: if you see x, tell me immed. [YES]

Anti-additivity

- Anti-additive area:

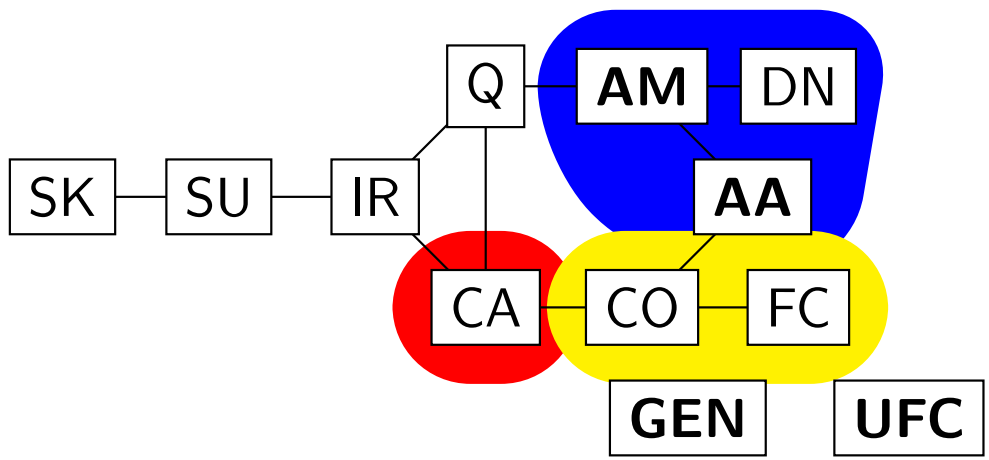


- Anti-additivity test: $Op(a \vee b) \Rightarrow Op(a) \wedge Op(b)$

(46)
a. FC: You may kiss John or Mary $\Rightarrow$ you may kiss John and you may kiss Mary [YES, but not in classical modal logic]
b. UFC: [John kissed *any* woman with red hair] John kissed Lee or Bea $\not\Rightarrow$ John kissed Lee and John kissed Bea [NO]
c. DN: I didn't see John or Mary. $\Rightarrow$ I didn't see John and I didn't see Mary [YES]
d. CO: Bill is taller than John or Mary. $\Rightarrow$ Bill is taller than John and Bill is taller than Mary [YES]

- Within anti-additive area we can distinguish:
  - Negative area (blue): $Op(a \vee \neg a)$ is $\bot$
  - Restrictor area (red): $Op(a \vee \neg a)$ is $\top$
  - Free choice area (yellow): $Op(a \vee \neg a)$ is neither

(47)
a. DN: The door is not open or close. (inconsistent)
b. IN: It is not necessary that (the door is open or close) (inconsistent)
c. CA: If the door is open or close, I will go to the party. (antecedent is trivial)
d. FC: The door may be open or close. (informative)
e. CO: ?Drinking is better than smoking or non-smoking.



Division of tasks

Design of the annotation experiment:

1. Formulation of guidelines and decision trees for the elicitation of judgements (3 students)
   First version by Tuesday 11-1-11
2. Preparation of the data to be annotated (2 students)
   Data inserted in data base (?) for pilot by Tuesday 11-1-11
   randomly selected examples from British National Corpus (BYU-BNC): 40 for *any* + 10 for singular *some*

Timeline

- Tuesday 11-1: short student presentations (20 minutes each) + begin pilot study annotation (50 examples)
- Thursday 13-1 : discussion pilot study annotation
- Tuesday 18-1: introduction to different methods for measuring inter-annotator agreement + calculation of inter-annotator agreement of pilot study