

ΟΥΧΟΥ (εν αναμονη καλυτερου ονοματος)

1. Task Description

The task is to populate a collection of samples for a fine-grained linguistic inference dataset.

Each sample consists of:

- a **pair of sentences**, a **Premise** and a **Hypothesis**
- a non-empty set of **inference labels**, with elements from **Entailment**, **Contradiction** and **Unknown**
 - the set must contain **Entailment** if at least one semantic reading of the premise supports at least one reading of the hypothesis
 - the set must contain **Contradiction** if at least one reading of the premise refutes the hypothesis (or supports its negation)
 - the set must contain **Unknown** if at least one reading of the premise neither supports nor refutes the hypothesis
- a non-empty set of **linguistic tags** characterizing the sample and justifying the labels assigned (see below for a specification of the tags considered)

A valid sample would look like:

Premise

"Δεν ξαναπατάω εκεί, χτεσινό φαγητό μας σέρβιραν!", παραπονέθηκε ο μπαμπάς της Φανής.

Hypothesis

Ο μπαμπάς της Φανής δήλωσε ότι το φαγητό ήταν φρεσκομαγειρεμένο.

Labels

- Contradiction

Tags

- Lexical Semantics:Lexical Entailment:Antonymy

Explanation

- The lexical semantics of the antonyms χτεσινό/φρεσκομαγειρεμένο are responsible for the label

A sample might be characterized by multiple relevant tags:

Premise

Προσπαθώντας να κάτσει, ο Τάκης έσπασε το πόδι της καρέκλας κι έπεσε στο πάτωμα.

Hypothesis

Η καρέκλα έσπασε.

Labels

- Entailment

Tags

- Lexical Semantics:Lexical Entailment:Meronymy
- Predicate-Argument Structure:Core Arguments

A sample can also have multiple labels (e.g. in the presence of syntactic ambiguities):

Premise

Η Καλλιόπη είδε την πάπια με τα κυάλια.

Hypothesis

Η πάπια είχε κυάλια.

Labels

- Entailment
- Unknown

Tags

- Predicate-Argument Structure:Syntactic Ambiguity
-

Linguistic Tags

Linguistic tags are organized hierarchically from least to most specific.

When annotating a sample, categorization must proceed all the way down to the most specific entry level available; that is, **Logic:Quantification** is not a valid tag, because it has children tags **Universal**, **Existential** and **Non-Standard**, whereas **Common Sense/Knowledge** is valid, seeing as it has no internal subcategorization.

1. Lexical Semantics
 1. Lexical Entailment
 1. Hyponymy
 2. Hypernymy
 3. Synonymy
 4. Antonymy
 5. Meronymy
 2. Morphological Modification
 3. Factivity
 1. Factive
 2. Non-Factive
 4. Symmetry/Collectivity
 5. Redundancy
 6. FAO
2. Predicate-Argument Structure
 1. Syntactic Ambiguity
 2. Core Arguments
 3. Alternations
 4. Ellipsis
 5. Anaphora/Coreference
 6. Intersectivity
 1. Intersective
 2. Non-Intersective
 7. Restrictivity
 1. Restrictive
 2. Non-Restrictive
3. Logic
 1. Single Negation
 2. Multiple Negations
 3. Conjunction
 4. Disjunction
 5. Conditionals
 6. Negative Concord
 7. Quantification
 1. Universal
 2. Existential
 3. Non-Standard
 8. Comparatives
 9. Temporals
4. Common Sense/Knowledge

2. Annotation Guidelines

Annotation consists of two phases: generation and validation.

During generation, an annotator comes up with a number of novel examples, i.e. sentence pairs and a suggested set of inference labels and linguistic tags.

During validation, an annotator inspects a number of sentence pairs written by another annotator, and assigns them any inference label and linguistic tag they find appropriate.

2.1 Generation

Each annotator should prepare a minimum of 100 examples (the more, the better!).

The aim is to achieve a balanced inclusion of all labels and tags.

Given that the multiple label annotation is less natural, we propose the following percentages:

~30% for a singleton **Entailment**, **Contradiction** and **Unknown** label, and 5-10% for a combination of any two labels.

For tags, we propose including an element of each of the major categories (**Lexical Semantics**, **Predicate-Argument Structure**, **Logic** and **Common Sense/Knowledge**) in at least 25% of the overall samples.

2.2 Validation

Each annotator will be handed approximately 400 sentence pairs.

For each pair, they will carefully read the two sentences, and then assign the pair a set of possible inference labels and appropriate linguistic tags.

2.3 Tips for Good Annotation

When annotating try to adhere to the following:

- Avoid overusing the same lexical items and/or syntactic constructions. Variation is gold!
- After having written an example, inspect it anew and try to see if any more labels/tags are suitable.
- For certain examples, flipping the sentence pair around might lead to potentially interesting examples. In such cases, avoid lexical or syntactic variation (i.e. keep the sentences unchanged). Make sure to double check the labels and tags assigned – it's not necessarily the case that they will carry through from the original!
- Θέλουμε να ηρεμήσετε!

2.4 Annotation Format

When writing down an example or annotating a sentence pair, make sure to adhere to the following format:

- The first line contains the premise sentence.
- The second line contains the hypothesis sentence.
- The third line contains all possible inference labels, separated by a comma and a whitespace (the order doesn't matter). The first letter of each label is capitalized.
- The following lines contain the set of fitting linguistic tags, one tag per line (the order doesn't matter).
Each tag is specified by the most refined entry level available (e.g. we write just **Hyponymy** for **Lexical Semantics:Lexical Entailment:Hyponymy**).
- A blank empty line separates an example from the next one.

Inspect [annotation_example.txt](#) for an example.

3. Examples

Example 1

Premise

"Δεν ξαναπατάω εκεί, χτεσινό φαγητό μας σέρβιραν!", παραπονέθηκε ο μπαμπάς της Φανής.

Hypothesis

Ο μπαμπάς της Φανής δήλωσε ότι το φαγητό ήταν φρεσκομαγειρεμένο.

Labels

- Contradiction

Tags

- Lexical Semantics:Lexical Entailment:Antonymy

Explanation

The lexical semantics of the antonyms χτεσινό/φρεσκομαγειρεμένο are responsible for the label

Example 2

Premise

Προσπαθώντας να κάτσει, ο Τάκης έσπασε το πόδι της καρέκλας κι έπεσε στο πάτωμα.

Hypothesis

Η καρέκλα έσπασε.

Labels

- Entailment

Tags

- Lexical Semantics:Lexical Entailment:Meronymy
- Predicate-Argument Structure:Core Arguments

Explanation

The lexical semantics of meronyms (πόδι καρέκλας/καρέκλα) are responsible for the label. The Predicate-Argument Structure:Core Arguments label in this case denotes the change from a transitive to an intransitive meaning for the verb "έσπασε"

Example 3

Premise

Κατά τη διάρκεια των χειμερινών μηνών, η θάλασσα παραείναι κρύα για το μέσο λουόμενο.

Hypothesis

Ο μέσος λουόμενος βρίσκει τη θάλασσα υπερβολικά κρύα τον χειμώνα.

Labels

- Entailment

Tags

- Lexical Semantics:Morphological Modification

Explanation

Morphological modification with the morpheme παρά

Example 4

Premise

Τα παραδοσιακά ανάλατα τυριά είναι κατά βάση τα μαλακά λευκά τυριά.

Hypothesis

Τα μαλακά λευκά τυριά συνήθως δεν έχουν αλάτι.

Labels

- Entailment

Tags

- Lexical Semantics:Morphological Modification

Explanation

Morphological negation with the "α" morpheme

Example 5

Premise

Οι παρευρισκόμενοι είδαν τον μπάτσο να χτυπάει βάνουσα ένα παιδί.

Hypothesis

Ο αστυνομικός χτύπησε ένα παιδί.

Labels

- Entailment

Tags

- Lexical Semantics:Factivity:Factive
- Lexical Semantics:Lexical Entailment:Synonymy
- Lexical Semantics:Redundancy

Explanation

A factive verb, synonymy and redundancy (dropping βάνουσα will not affect the truth conditions of the hypothesis) play a role in this example.

Example 6

Premise

Οι παρευρισκόμενοι ανέφεραν ότι είδαν το όργανο της τάξης να χτυπάει βάνουσα ένα παιδί.

Hypothesis

Ο αστυνομικός χτύπησε ένα παιδί.

Labels

- Unknown

Tags

- Lexical Semantics:Factivity:Non-Factive
- Lexical Semantics:Lexical Entailment:Synonymy
- Lexical Semantics:Redundancy

Explanation

A non-factive verb, synonymy and redundancy (dropping βάνουσα will not affect the truth conditions of the hypothesis) play a role in this example.

Example 7

Premise

Η Αρετή εικάζει ότι η γη είναι επίπεδη.

Hypothesis

Η γη είναι επίπεδη.

Labels

- Unknown

Tags

- Lexical Semantics:Factivity:Non-Factive

Explanation

Non-factive verb "εικάζει"

Example 8

Premise

Η Νιόβη παντρεύτηκε τη Βιβή.

Hypothesis

Η Βιβή παντρεύτηκε τη Νιόβη.

Labels

- Entailment

Tags

- Lexical Semantics:Symmetry

Explanation

A verb expressing a symmetrical relation, a married b implies that b married a.

Example 9

Premise

Ο Αλέξης μάλωσε με τον Γιάνη.

Hypothesis

Ο Αλέξης και ο Γιάνης μάλωσαν.

Labels

- Entailment

Tags

- Lexical Semantics:Collectivity

Explanation

A transitive verb expressing a relation between two arguments that is turned into an intransitive verb taking the result of the conjunction of the two arguments as argument

Example 10

Premise

Ο Πέτρος άνοιξε την κονσέρβα.

Hypothesis

Ο Πέτρος άνοιξε την κονσέρβα με τα δόντια.

Labels

- Unknown

Tags

- Lexical Semantics:Redundancy

Explanation

The information "με τα δόντια" is the information referred as "redundant", but appears in the hypothesis.

Example 11

Premise

Ο Θωμάς και ο Αδάμ ψιθύριζαν δυνατά κατά τη διάρκεια της παράστασης.

Hypothesis

Ο Θωμάς και ο Αδάμ ψιθύριζαν κατά τη διάρκεια της παράστασης.

Labels

- Entailment

Tags

- Lexical Semantics:Redundancy

Explanation

Dropping "δυνατά" does not affect the truthfulness of the sentence.

Example 12

Premise

Χορεύοντας σάμπα, ο Περικλής κατάφερε να γοητέψει τη Θάλια.

Hypothesis

Ο Περικλής γοήτεψε μόνο τη Θάλια με το χορό του.

Labels

- Unknown

Tags

- Lexical Semantics:FAO

Explanation

FAO "μόνο" in the hypothesis. The premise does not imply that Pericles managed to charm Thalia only.

Example 13

Premise

Η Περσεφόνη πεινούσε τόσο πολύ που έφαγε και τα ψίχουλα.

Hypothesis

Η Περσεφόνη έφαγε μόνο τα ψίχουλα.

Labels

- Contradiction

Tags

- Lexical Semantics:FAO

Explanation

The semantics of FAO "μόνο" are responsible for the contradiction here.

Example 14

Premise

Κάθε Έλληνας ντράμερ θέλει να γίνει διάσημος.

Hypothesis

Κάποιος Έλληνας ντράμερ θέλει να γίνει διάσημος.

Labels

- Entailment

Tags

- Logic:Quantification:Universal
- Logic:Quantification:Existential

Explanation

Both quantifiers, universal and existential, play a role in deciding the label.

Example 15

Premise

Η Καλλιόπη είδε την πάπια με τα κυάλια.

Hypothesis

Η πάπια είχε κυάλια.

Labels

- Entailment
- Unknown

Tags

- Predicate-Argument Structure:Syntactic Ambiguity

Explanation

Multiple labels due to different possibilities of PP attachment.

Example 16

Premise

Η Καλλιόπη είδε την πάπια με τα κυάλια.

Hypothesis

Η Καλλιόπη είδε την πάπια.

Labels

- Entailment

Tags

- Lexical Semantics:Redundancy

Explanation

Dropping the PP "με τα κυάλια" does not affect truthfulness.

Example 17

Premise

Όσο ήμουν στην τουαλέτα, ο Γιάννης μου έφαγε το μήλο!

Hypothesis

Το μήλο τελικά φαγώθηκε απο μένα.

Labels

- Contradiction

Tags

- Predicate-Argument Structure:Alternation

Explanation

Passive alternation (plus ethical dative clitic vs by-phrase).

Example 18

Premise

Ο Μητσοτάκης δήλωσε με στόμφο "Κύριε Χατζηκυριακίδη λάμψατε διά της απουσίας σας", λαμβάνοντας την καίρια απάντηση "Κι εσείς!".

Hypothesis

Ο Μητσοτάκης έλαμψε δια της απουσίας του.

Labels

- Entailment

Tags

- Predicate-Argument Structure:Ellipsis

Explanation

VP ellipsis and FAO additive conjunction "και"

Example 19

Premise

Ο Γιάννης δήλωσε στη Μαρία ότι του αρέσει το γάλα αμυγδάλου.

Hypothesis

Στη Μαρία αρέσει το γάλα αμυγδάλου.

Labels

- Unknown

Tags

- Predicate-Argument Structure:Anaphora/Coreference

Explanation

Clitic pronoun "του" is coreferent with "Γιάννης".

Example 20

Premise

Ο Λούο είναι Κινέζος θεωρητικός τύπων.

Hypothesis

Ο Λούο είναι τυποθεωρητικός.

Labels

- Entailment

Tags

- Predicate-Argument Structure:Intersectivity:Intersective

Explanation

Intersective adjective "Κινέζος" does not alter the truthfulness of the premise.

Example 21

Premise

Ο Τρύφων καυχέται πως είναι επιδέξιος χορευτής.

Hypothesis

Ο Τρύφων καυχείται ότι είναι επιδέξιος.

Labels

- Unknown

Tags

- Predicate-Argument Structure:Intersectivity:Non-Intersective

Explanation

Non-intersective adjective "επιδέξιος" – an "επιδέξιος χορευτής" is not necessarily "επιδέξιος".

Example 22

Premise

Οι θαλάσσιοι ίπποι που ζυγίζουν έως και δύο τόνους απειλούνται άμεσα απο το λιώσιμο των πάγων.

Hypothesis

Οι θαλάσσιοι ίπποι απειλούνται άμεσα απο το λιώσιμο των πάγων.

Labels

- Unknown

Tags

- Predicate-Argument Structure:Restrictivity:Restrictive

Explanation

Restrictive relative clause; only walruses weighting up to 2 tons are endangered.

Example 23

Premise

Οι θαλάσσιοι ίπποι, που ζυγίζουν έως και δύο τόνους, απειλούνται άμεσα απο το λιώσιμο των πάγων.

Hypothesis

Οι θαλάσσιοι ίπποι απειλούνται άμεσα απο το λιώσιμο των πάγων.

Labels

- Entailment

Tags

- Predicate-Argument Structure:Restrictivity:Non-Restrictive

Explanation

Non-restrictive relative clause; all walruses are endangered, and they can coincidentally weight up to 2 tons.

Example 24

Premise

Ο Φώτης φοβάται μην δεν έρθει η Μαρία.

Hypothesis

Ο Φώτης φοβάται μήπως έρθει η Μαρία.

Labels

- Contradiction

Tags

- Logic:Single Negation

Explanation

The apparent double negation "μην δεν" is spurious, and Fotis can either worry that Maria comes, or that she doesn't.

Example 25

Premise

Ή θα φας το φαί σου ή θα'χουμε άλλα.

Hypothesis

Θα'χουμε άλλα.

Labels

- Unknown

Tags

- Logic:Disjunction

Explanation

Cannot infer the truthfulness of any single coordinate of a true disjunction.

Example 26

Premise

Θα'χουμε άλλα.

Hypothesis

Ή θα φας το φαί σου ή θα'χουμε άλλα.

Labels

- Entailment

Tags

- Logic:Disjunction

Explanation

From a true statement, one can weaken it by a disjunction with an arbitrary statement.

Example 27

Premise

Κανένα δεν ήθελε να δώσει κάτι.

Hypothesis

Κάποιο ήθελε να δώσει κάτι.

Labels

- Contradiction

Tags

- Logic:Propositional Structure:Negative Concord

Explanation

"Κανένα" and "δεν" are mutually affirming in negative concord.

Example 28

Premise

Κανένα δεν ήθελε να μη δε δώσει τίποτα.

Hypothesis

Όλα ήθελαν να δώσουν κάτι.

Labels

- Entailment

Tags

- Logic:Propositional Stucture:Multiple Negation
- Logic:Propositional Stucture:Negative Concord

Explanation

The double negation "μη δε" is spurious (or emphatic), negates "τίποτα" and affirms "κανένα".

Example 29

Premise

Κανένα δεν ήθελε να μη δώσει τίποτα.

Hypothesis

Όλα ήθελαν να δώσουν κάτι.

Labels

- Entailment

Tags

- Logic:Propositional Stucture:Multiple Negation
- Logic:Propositional Stucture:Negative Concord

Explanation

Ditto (see above).

Example 30

Premise

Ο Φωκίων ήθελε και την πίτα ολόκληρη και το σκύλο χορτάτο.

Hypothesis

Ο Φωκίων ήθελε το σκύλο χορτάτο.

Labels

- Entailment

Tags

- Logic:Conjunction

Explanation

A true conjunction is true on all coordinates.

Example 31

Premise

Αν η γιαγιά μου είχε ρόδες θα ήταν πατίνι.

Hypothesis

Η γιαγιά μου έχει ρόδες.

Labels

- Unknown

Tags

- Logic:Conditional

Explanation

A conditional statement allows no inference on its condition.

Example 32

Premise

Το συνέδριο ξεκίνησε στις 4 Ιουλίου και διήρκεσε δύο ημέρες.

Hypothesis

Το συνέδριο έληξε στις 6 Ιουλίου.

Labels

- Entailment

Tags

- Logic:Temporal

Explanation

$4+2=6$ (μόνο μαθηματικά και πράξεις)

Example 33

Premise

Η Δανάη είναι ψηλότερη από το Λυκούργο, και αυτός από την Πηνελόπη.

Hypothesis

Η Δανάη είναι ψηλότερη από την Πηνελόπη.

Labels

- Entailment

Tags

- Predicate-Argument Structure:Ellipsis
- Predicate-Argument Structure:Anaphora/Coreference
- Logic:Comparatives

Explanation

Transitivity of comparatives, part of the sentence is ellided and "αυτός" (1-0) is coreferent with Lycurgus.

Example 34

Premise

Ο ταξιτζής δε μιλούσε καμία γλώσσα.

Hypothesis

Ο ταξιτζής δε μιλούσε.

Labels

- Entailment

Tags

- Common Sense/Knowledge

Explanation

Inability to speak any language implies inability to speak in general.

Example 35

Premise

Ο ταξιτζής δε μιλούσε.

Hypothesis

Ο ταξιτζής δε μιλούσε καμία γλώσσα.

Labels

- Unknown

Tags

- Common Sense/Knowledge

Explanation

Not speaking does not imply inability to speak any language.

Example 36

Premise

Χιλιάδες φίλων συγκεντρώθηκαν για να αποδώσουν φόρο τιμής και να πουν το τελευταίο αντίο στον διάσημο κλόουν.

Hypothesis

Ο διάσημος κλόουν απεβίωσε.

Labels

- Entailment

Tags

- Common Sense/Knowledge

Explanation

Final goodbyes are said to departed entities.
