Summary

This dataset consists of conversational implicatures of utterances. Conversational implicatures are the meanings of an utterance more than what is literally stated by the utterance.

The data consist of 1001 utterances that come as responses in a specific context and their implicatures. These written representations of the utterances are collected manually by scraping and transcribing from relevant sources from August, 2019 to August, 2020. The source of dialogues in the data include TOEFL listening comprehension short conversations, movie dialogues from IMSDb and websites explaining idioms, similes, metaphors and hyperboles. The implicatures are annotated manually.

This data is collected as part of the research work done at Language Technologies Research Centre (LTRC), International Institute of Information Technology, Hyderabad by Elizabeth jasmi George under the supervision of Dr. Radhika Mamidi for the MS by research in Computer Science and Engineering course.

Citation

To acknowledge use of the dataset in publications, please cite the following paper:

Format and Content of the File

Formatting

The dataset file is written as comma-separated values file. Columns that contain commas (,) are escaped using double-quotes ("). The dataset is also available as an excel sheet.

Content

Each entry in the dataset consists of a context utterance, a response utterance and an Implicature.

Context Utterance

The written representation of an utterance which serves as the context in which the response utterance can implicate a meaning different from its literal meaning.

Response Utterance

The written representation of an utterance which has a different meaning than the meaning of the sentences used in it.

Implicature

The implicated meaning of the response utterance.

Statistics

The data majorly consist of polar questions coming as context and their nonpolar answers coming as responses. A direct 'yes' or 'no' with an optional explanation statement is given as the conversational implicature. The major genres of response utterances present in this dataset and their statistics are given below.

Kind of response utterance	Number of Snippets
Simile as a response utterance	35
PopeQ as a response utterance	10
Metaphor as a response utterance	41
Tautology as a response utterance	9

Hyperbole as a response utterance	23
Idiom as a response utterance	96

Table 1. Number of dialogue snippets per kind of response utterance

Source of data	Number of Snippets
Movie dialogues from IMSDb	492
TOEFL listening comprehension questions	246
Websites	263

Table 2. Number of dialogue snippets per source of data

Polarity of implicature	Number of snippets
Yes	371
No	340

others	291

Table 3. Number of dialogue snippets per specific polarity.

Kind of context utterance	Number of Snippets
Questions	857
Statements	144
Others.	0

Table 4. Number of dialogue snippets per kind of context utterance

Examples of Conversational Implicatures

Type	Context utterance	Response utterance	Implicature
Scalar	Who made these donuts?	I made some of these donuts.	I did not make all of these donuts.
Generalized	Did you call John and Benjamin?	I called Benjamin.	I did not call John.
Particularized	Did you drink the milk I kept on the table?	The cat seems to be happy.	No. I did not drink milk. The cat might have drunk the milk.
Relevance	How about going for a walk?	Isn't it raining out?	No. I am not coming for a walk now.

PopeQ as a response	Are you sure you can take care of yourself this weekend?	Can a duck swim, mother?	Yes. I am sure I can take care of myself.
Metaphor as a response	Do you like her?	She is like cream in my coffee.	Yes. I like her a lot.
Tautology as a response	Do you want to taste my Hamburger?	Hamburger is hamburger.	No.Hamburgers are not great to taste.
Hyperbole as a response	Are you hungry?	I could eat a horse.	Yes. I am extremely hungry.
Idiom as a response	I could have been more careful.	It is useless to cry over spilled milk.	It is useless to be sad about what had already happened.

Table 5. Types of conversational implicatures with examples.