Detecting Indirectness in Conversation

Pranav Goel, Yoichi Matsuyama, Michael Madaio and Justine Cassell

Indian Institute of Technology (BHU) Varanasi and Carnegie Mellon University

Example Scenario: A student tutoring their classmate finds them making a mistake...

Example Scenario: A student tutoring their classmate finds them making a mistake...

"*I think* it *might* help if we multiply, and not add."

instead of "Don't add, multiply!' - Why?

What is Indirectness?

Definition

The linguistic phenomena where the speaker intentionally does not communicate their intention in a straightforward or direct manner.

How?

Definition

The linguistic phenomena where the speaker intentionally does not communicate their intention in a straightforward or direct manner.

How?

- By distancing themselves from their intended meaning

Definition

The linguistic phenomena where the speaker intentionally does not communicate their intention in a straightforward or direct manner.

How?

- By distancing themselves from their intended meaning
- By introducing vagueness
- Reducing certainty or intensity

Definition

The linguistic phenomena where the speaker intentionally does not communicate their intention in a straightforward or direct manner.

How?

- By distancing themselves from their intended meaning
- By introducing vagueness
- Reducing certainty or intensity
- Making their statements appear more subjective (and more)

Importance: Where and Why?

Used in a variety of linguistic contexts

Importance: Where and Why?

Used in a variety of linguistic contexts

Business Negotiations

Counseling

Education

Health Care

Importance: Where and Why?

The stakes are often high! Speak indirectly for effective communication - why?


Profitable Business Deals


Life-Saving Medical Advice

Importance: Why?

‘Give me that!’ - Being overtly direct can threaten the interlocutors‘ desired self-image (***Face Threat***)

Importance: Why?

'Give me that!' - Being overtly direct can threaten the interlocutors' desired self-image (***Face Threat***)

Being indirect helps mitigate the face threat.

In Business Negotiation

Being indirect is used to

In Business Negotiation

Being indirect is used to



- Present tentative views
- Weaken one's commitment to a particular bid
- Build trust between negotiators

In Doctor-Patient Discourse

Softening the blow, long term relationship, and even the success of the conversation itself...

In Doctor-Patient Discourse

Softening the blow, long term relationship, and even the success of the conversation itself...

Example: A dire medical diagnosis

In Doctor-Patient Discourse

Softening the blow, long term relationship, and even the success of the conversation itself...

Example: A dire medical diagnosis

Example: An expectant mother asked to take an HIV test

Detecting the Strategic Use

Detecting the Strategic Use

Example: Bosses mitigating orders

Boss:
Less threatening,
more friendly way of speaking

Detecting the Strategic Use

Example: Bosses mitigating orders

Importance with respect to Spoken Dialogue Systems

For Spoken Dialogue Systems

- Computer-mediated discourse and the use of virtual assistants is increasing in many domains.

For Spoken Dialogue Systems

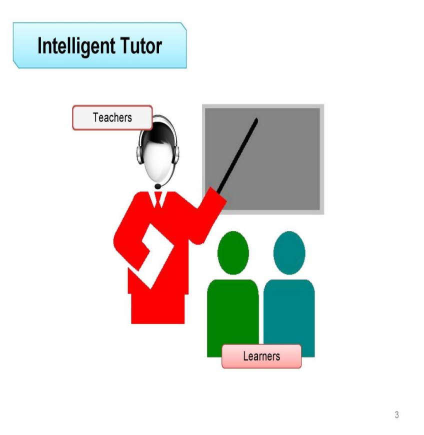
- Computer-mediated discourse and the use of virtual assistants is increasing in many domains.
- Knowing when a user is being indirect can help such systems *better understand, respond to, and build the user-agent relationship*.

Spotlight: Education

Education is a productive domain to target

Spotlight: Education

Education is a productive domain to target



Students use indirectness when proposing answers to their teachers.

'I guess the answer's fifty?'

Spotlight: Education

Students peer tutoring one another use indirectness to reduce the threat to their partner's self-image and self-esteem ('face').

'I would be subtracting that number first, but that's just the way I would do it.'

Spotlight: Education

Per Madaio et al. (2017) -

Observation: Peer tutors who had *greater self-efficacy* used *more indirectness*

Spotlight: Education

Per Madaio et al. (2017) -

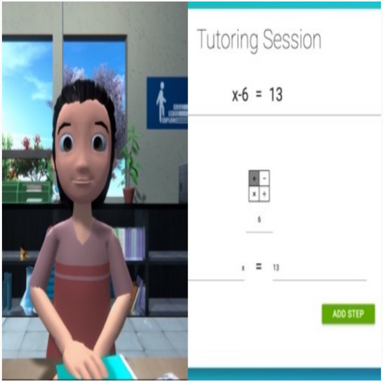
Suggestion: Peer tutors who had *greater self-efficacy* used *more indirectness* – it plays a **strategic face-saving** role – a **relationship-building** role.

Spotlight: Education

Per Madaio et al. (2017) -

Observation: Tutors with a stronger relationship with their partners were more direct

Spotlight: Education



Per Madaio et al. (2017) -

Suggestion: Tutors with a stronger relationship with their partners were more direct – In a **spoken dialogue system**, having the agent be continuously polite or indirect may in fact be perceived as *distancing* – may harm the rapport between the agent and user.

For Spoken Dialogue Systems

Automatically detecting indirectness in user utterances can thus help all kinds of spoken dialogue systems better estimate the state of the social relationship - an aid in designing an appropriate response - to more effectively communicate information.

Related Phenomena

Politeness

- Being indirect is *one of the ways* of being polite (in requests). One can be polite without being indirect.

Politeness

- Being indirect is *one of the ways* of being polite (in requests). One can be polite without being indirect.
- "**Nice work** so far on the rewrite". Indirect? ✗ Polite? ✓

Politeness

- Being indirect is *one of the ways* of being polite (in requests). One can be polite without being indirect.
- "**Nice work** so far on the rewrite". Indirect? ✗ Polite? ✓
- "Indirectness is not always interpreted as politeness and can even be associated with lack of politeness."

Politeness

- Being indirect is *one of the ways* of being polite (in requests). One can be polite without being indirect.
- "Nice work so far on the rewrite". Indirect? ✗ Polite? ✓
- "Indirectness is not always interpreted as politeness and can even be associated with lack of politeness."
- "can you please *just* stay with me and not doodle". Indirect? ✓ Polite? ✗

Uncertainty

Uncertainty

- Indirectness is often produced through the use of hedges - “single- or multi-word expressions used to indicate uncertainty about the propositional content of an utterance or to diminish its impact”. Thus, *uncertainty is just one of the ways in which indirectness can manifest in conversations*.

Uncertainty

- Indirectness is often produced through the use of hedges - “single- or multi-word expressions used to indicate uncertainty about the propositional content of an utterance or to diminish its impact”. Thus, *uncertainty is just one of the ways in which indirectness can manifest in conversations.*
- “er A equals twenty-four sorry”. Indirect? ✓ Polite? ✓ Uncertain? ✗

Uncertainty

- Indirectness is often produced through the use of hedges - "single- or multi-word expressions used to indicate uncertainty about the propositional content of an utterance or to diminish its impact". Thus, *uncertainty is just one of the ways in which indirectness can manifest in conversations*.
- "er A equals twenty-four sorry". Indirect? ✓ Polite? ✓ Uncertain? ✗
- A statement may also be uncertain (due to lack of exact data) without being indirect.

Uncertainty

- Indirectness is often produced through the use of hedges - "single- or multi-word expressions used to indicate uncertainty about the propositional content of an utterance or to diminish its impact". Thus, *uncertainty is just one of the ways in which indirectness can manifest in conversations*.
- "er A equals twenty-four **sorry**". Indirect? ✓ Polite? ✓ Uncertain? ✗
- A statement may also be uncertain (due to lack of exact data) without being indirect.
- "The club enjoyed **most of its success** in its early years". Indirect? ✗ Uncertain? ✓

Gaps in Prior Work

Focused on a Specific Aspect

- Focused only on a specific function (politeness), or some specific manifestation (uncertainty) of indirectness.

Focused on a Specific Aspect

- Focused only on a specific function (politeness), or some specific manifestation (uncertainty) of indirectness.
- Example: Politeness Classifier (Danescu-Niculescu-Mizil et al., 2013)

Focused on a Specific Aspect

- Focused only on a specific function (politeness), or some specific manifestation (uncertainty) of indirectness.
- Example: Politeness Classifier (Danescu-Niculescu-Mizil et al., 2013)
- Example: CoNLL 2010 Shared Task on Uncertainty Detection (Farkas et al.) - Automated separation of uncertain and factual statements. Two annotated datasets - a BioScope corpus (abstracts and articles from biomedical literature) and a Wikipedia corpus.

Used Single Modality

Used Single Modality

- Systems have been using a single modality - mostly, text.

Used Single Modality

- Systems have been using a single modality - mostly, text.
- In the spoken dialogue setting, past works have used prosody-based features to detect ‘certainness’ or uncertainty.

Used Single Modality

- Systems have been using a single modality - mostly, text.
- In the spoken dialogue setting, past works have used prosody-based features to detect 'certainness' or uncertainty.
- What about visual nonverbal behaviors (deemed important for face-threat mitigation)?

Used Single Modality

- Systems have been using a single modality - mostly, **text**.
- In the spoken dialogue setting, past works have used **prosody**-based features to detect ‘certainness’ or uncertainty.
- What about **visual** nonverbal behaviors (deemed important for face-threat mitigation)?
- Using multiple modalities together!

Prior Work vs Our Work: The Crucial Difference

- To summarize:
  Prior work has focused on related but different phenomena like politeness and uncertainty, and have only used a single modality to get features for their system. There is also no prior attempt to use visual features.

Data

Corpus Collection

- 12 American-English speaking pairs (or dyads), aged 12-15.

Corpus Collection

- 12 American-English speaking pairs (or dyads), aged 12-15.
- Tutored each other in basic linear algebra for a total of 60 hours.

Corpus Collection

- 12 American-English speaking pairs (or dyads), aged 12-15.
- Tutored each other in basic linear algebra for a total of 60 hours.

- Audio and video data were recorded, transcribed, and segmented for clause-level dialogue annotation.

Corpus Annotation

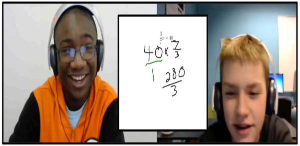
Code	Definition	Example
Apology	Apologies used to soften direct speech acts	Sorry, it‘s negative 2.
Qualifiers	Qualifying words for reducing intensity or certainty	You just add 5 to both sides.
Extenders	Indicating uncertainty by referring to vague categories	You have to multiply and stuff.
Subjectivizer	Making an utterance seem more subjective to reduce intensity	I think you divide by 3 here.

Table 1: Annotation of codes under the ‘indirect’ label (1113/23437)

Nonverbal Behaviors

Used the front-facing camera (Skype based interactions)

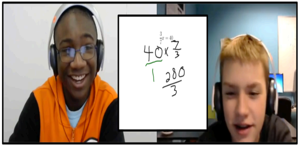
OpenFace software for analyzing the nonverbal visual behaviors



Nonverbal Behaviors

Used the front-facing camera (Skype based interactions)

OpenFace software for analyzing the nonverbal visual behaviors



The visual behaviors annotated -

- Eye Gaze (Gaze at Partner (gP), Gaze at the worksheet (gW), and Gaze elsewhere (gE))
- Smile
- Head Nod

Methodology

Non Neural Method

Non Neural Method

Step 1: Feature Representation

Step 1: Feature Representation

From the **text** of the utterances in our corpus

Non Neural Method

Step 1: Feature Representation

From the **text** of the utterances in our corpus

- Unigrams
- Pair-based (Bigrams, POS bigrams, Word-POS pairs)

Non Neural Method

Step 1: Feature Representation

From the **text** of the utterances in our corpus

- Unigrams
- Pair-based (Bigrams, POS bigrams, Word-POS pairs)
- Word2Vec model trained on our data (RPT Word2Vec)

Non Neural Method

Step 1: Feature Representation

Text, but external data (pre-trained models): We used word vectors or embeddings which were trained on different, larger datasets.

Non Neural Method

Step 1: Feature Representation

Text, but external data (pre-trained models): We used word vectors or embeddings which were trained on different, larger datasets.

- Twitter Word2Vec
- Wikipedia Word2Vec
- GloVe wiki
- GloVe Common Crawl

Non Neural Method

Step 1: Feature Representation

Text, but external data (pre-trained models): We used word vectors or embeddings which were trained on different, larger datasets.

- Twitter Word2Vec
- Wikipedia Word2Vec
- GloVe wiki
- GloVe Common Crawl

Visual features (10): Gaze at Partner, Gaze at Worksheet, Gaze Elsewhere, Smile and Head Nod for both tutor and tutee.

Step 2: Training

Non Neural Method

Step 2: Training

- 60 sessions of peer tutoring interaction
- Training-validation-test split (48-6-6).
- Validation set used to decide best approaches, tune hyperparameters.

Non Neural Method

Step 2: Training

- 60 sessions of peer tutoring interaction
- Training-validation-test split (48-6-6).
- Validation set used to decide best approaches, tune hyperparameters.
- For each feature combination, many supervised ML approached tried: Logistic Regression, Naive Bayers, Random Forests, and SVM.

Neural Method

Neural Method

- Tried various neural architectures that have worked well for past NLP classification based tasks.
- Checked the performance on the validation set to decide the best performing architectures.

Neural Method

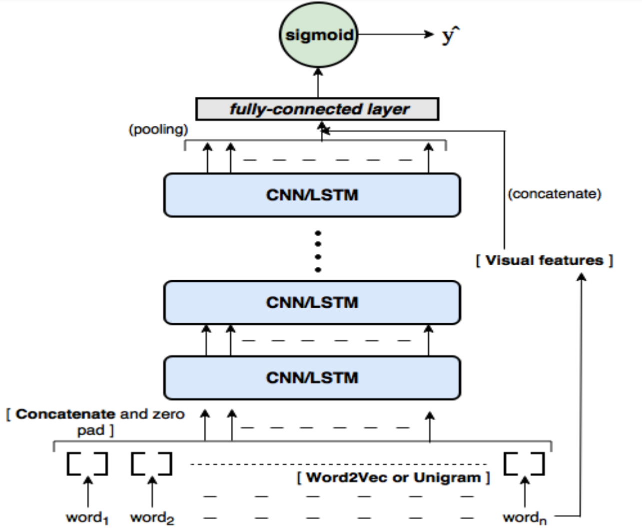


Figure 1: A general representation of the different neural architectures (and combinations) tried

Neural Method

The best two architectures are -

1. Stacked LSTMs
2. Attention based CNN (also previously gave state-of-the-art results for uncertainty detection in CoNLL 2010 shared task setting)

Evaluation

- We use F1 score as our evaluation metric.

Evaluation

- We use F1 score as our evaluation metric.
- The best performing approaches developed for uncertainty detection (ConLL 2010) were also tried for our task for comparison.
- Our neural approaches (like Stacked LSTMs) were also tried on uncertainty detection.

Results and Implications

Non-Neural Method (comparison)

	Logistic Reg.	SVM
Unigram	57.71	59.1
Unigram+Visual	57.74	**59.3**
Pair-based	57.09	58.28
Pair-based+Visual	55.89	58.41
Twitter Word2Vec	44.83	53.86
GloVe Wiki	37.91	45.25
GloVe Common Crawl	38.94	45.06
Wikipedia Word2Vec	44.56	49.54
RPT word2vec	44.95	39.36

Table 2: F1 score (%) on test set for various features and their combinations fed to non-neural ML algorithms for indirectness detection on reciprocal peer-tutoring dataset.

Non-Neural Method (comparison)

	Logistic Reg.	SVM
Unigram	57.71	59.1
Unigram+Visual	57.74	**59.3**
Pair-based	57.09	58.28
Pair-based+Visual	55.89	58.41
Twitter Word2Vec	44.83	53.86
GloVe Wiki	37.91	45.25
GloVe Common Crawl	38.94	45.06
Wikipedia Word2Vec	44.56	49.54
RPT word2vec	44.95	39.36

Table 2: F1 score (%) on test set for various features and their combinations fed to non-neural ML algorithms for indirectness detection on reciprocal peer-tutoring dataset.

Reliance of tasks like indirectness detection on the specific domain (hinted at by previous studies) - results using features learned from our peer tutoring data (unigrams or bag of words), versus word vectors trained on different, much larger datasets.

Non-Neural Method (comparison)

	Logistic Reg.	SVM
Unigram	57.71	59.1
Unigram+Visual	57.74	**59.3**
Pair-based	57.09	58.28
Pair-based+Visual	55.89	58.41
Twitter Word2Vec	44.83	53.86
GloVe Wiki	37.91	45.25
GloVe Common Crawl	38.94	45.06
Wikipedia Word2Vec	44.56	49.54
RPT word2vec	44.95	39.36

Table 3: F1 score (%) on test set for various features and their combinations fed to non-neural ML algorithms for indirectness detection on reciprocal peer-tutoring dataset.

The social media like nature of peer tutoring conversations - Amongst the pre-trained word vector based models, Twitter word2vec performs best. Many utterances in our corpora do share the short length and informal nature of Twitter tweets.

Non-Neural Method (comparison)

	Logistic Reg.	SVM
Unigram	57.71	59.1
Unigram+Visual	57.74	**59.3**
Pair-based	57.09	58.28
Pair-based+Visual	55.89	58.41
Twitter Word2Vec	44.83	53.86
GloVe Wiki	37.91	45.25
GloVe Common Crawl	38.94	45.06
Wikipedia Word2Vec	44.56	49.54
RPT word2vec	44.95	39.36

Table 4: F1 score (%) on test set for various features and their combinations fed to non-neural ML algorithms for indirectness detection on reciprocal peer-tutoring dataset.

Different visual behaviors, or different fusing technique needed.

Overall Results

	Reciprocal Peer-Tutoring Corpus (Indirectness Detection)	Wikipedia (Uncertainty Detection)	BioScope (Uncertainty Detection)
Att_Inp CNN	**62.03**	65.13*	84.99*
Att_Conv CNN	61.4	**66.49***	84.69*
Pre-trained W2V + Stacked LSTM	61.15	66.07	82.62
Pre-trained W2V + Stacked LSTM + Visual	61.35	-	-
Unigram + Stacked LSTM	56.5	43.71	73.03
Unigram + Stacked LSTM + Visual	57.11	-	-
SVM on Bag-of-Words	58.28	60.2*	**85.2***

Table 5: F1 score (%) for the various neural models compared with SVM approach for two different tasks on different datasets. Results marked with * have been taken from previous literature.

Overall Results

	Reciprocal Peer-Tutoring Corpus (Indirectness Detection)	Wikipedia (Uncertainty Detection)	BioScope (Uncertainty Detection)
Att_Inp CNN	**62.03**	65.13*	84.99*
Att_Conv CNN	61.4	**66.49***	84.69*
Pre-trained W2V + Stacked LSTM	61.15	66.07	82.62
Pre-trained W2V + Stacked LSTM + Visual	61.35	-	-
Unigram + Stacked LSTM	56.5	43.71	73.03
Unigram + Stacked LSTM + Visual	57.11	-	-
SVM on Bag-of-Words	58.28	60.2*	**85.2***

Table 5: F1 score (%) for the various neural models compared with SVM approach for two different tasks on different datasets. Results marked with * have been taken from previous literature.

The viability of neural networks

- 3-4% improvement over SVM.
- Effective in detecting both indirectness and uncertainty.

Overall Results

	Reciprocal Peer-Tutoring Corpus (Indirectness Detection)	Wikipedia (Uncertainty Detection)	BioScope (Uncertainty Detection)
Att_Inp CNN	**62.03**	65.13*	84.99*
Att_Conv CNN	61.4	**66.49***	84.69*
Pre-trained W2V + Stacked LSTM	61.15	66.07	82.62
Pre-trained W2V + Stacked LSTM + Visual	61.35	-	-
Unigram + Stacked LSTM	56.5	43.71	73.03
Unigram + Stacked LSTM + Visual	57.11	-	-
SVM on Bag-of-Words	58.28	60.2*	**85.2***

Table 5: F1 score (%) for the various neural models compared with SVM approach for two different tasks on different datasets. Results marked with * have been taken from previous literature.

The viability of neural networks

- 3-4% improvement over SVM.
- Effective in detecting both indirectness and uncertainty.
- Reinforces the importance of capturing the *context* well.

Conclusion: The Takeaways

Conclusion

Detecting indirectness can help virtual conversational agents and SDS respond to user in an appropriate manner.

Conclusion

Detecting indirectness can help virtual conversational agents and SDS respond to user in an appropriate manner.

- More effective in their task goals (business deals, medical advice or tutoring instructions)

Conclusion

Detecting indirectness can help virtual conversational agents and SDS respond to user in an appropriate manner.

- More effective in their task goals (business deals, medical advice or tutoring instructions)
- Managing interpersonal relationship with user (mitigating face threat and building trust)

Conclusion

Insights from our work can help, like

Conclusion

Insights from our work can help, like

- Twitter-like nature of collaborative educational discourse (Intelligent Tutoring Agents)

Conclusion

Insights from our work can help, like

- Twitter-like nature of collaborative educational discourse (Intelligent Tutoring Agents)
- The implied domain dependency, and the effectiveness of neural networks (SDS in general)

Future Work

Improving indirectness detection

Improving indirectness detection

- Leverage more visual behaviors (head tilts, laughter)

Improving indirectness detection

- Leverage more visual behaviors (head tilts, laughter)
- Use acoustic or paralinguistic features to create a fully multi-modal system

Incorporation into a larger system

Incorporation into a larger system

- Incorporate our findings into a general purpose socially-aware spoken dialogue system like SARA.

Figure 2: A Socially Aware Robot Assistant (SARA)

I guess I'm done.

Might as well say **Thank you.**

Should we have some **Questions** or something**?**