Fueling a Cycle for Continuous Improvement in Discussion Based Learning

Carolyn Penstein Rosé
cprose@cs.cmu.edu
http://www.cs.cmu.edu/~cprose
http://dance.cs.cmu.edu
Language Technologies Institute
and Human-Computer Interaction Institute
Carnegie Mellon University

Carnegie Mellon University

Outline

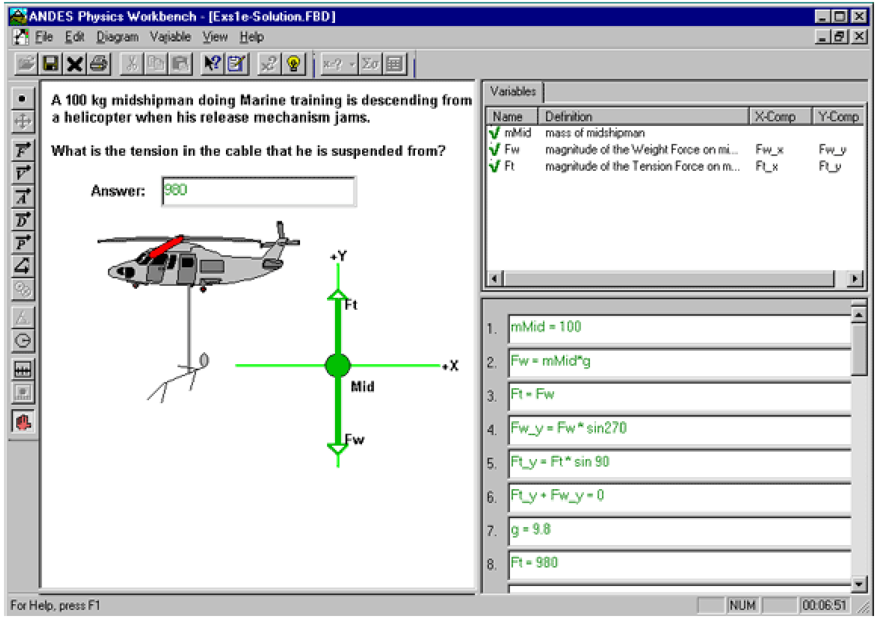
- Historical Overview: 18 years of work
- Technology Approach: Highlighting Chinese Discourse Analysis tools
- Recent Work on Advanced Discourse Analysis Techniques
- Conclusion

Outline

- **Historical Overview: 18 years of work**
- Technology Approach: Highlighting Chinese Discourse Analysis tools
- Recent Work on Advanced Discourse Analysis Techniques
- Conclusion

Why Computer-Supported Collaborative Learning?

Students learn significantly more when interactive directed lines of reasoning replace hints.
(Rosé et al., 2001)



Tutor: Let's think about the difference between speed and velocity. A closely related distinction is that of the difference between distance traveled and displacement from the origin. Take as an example a bee flying from point A to point B by means of a curvy path. If you draw a vector from point A to point B, you will have drawn the bee's displacement vector. What does that vector represent?

Student: The bee's distance?

Tutor: The displacement vector only represents the net change in position. So it does not tell you how far the bee traveled. If you want to know how far the bee traveled, what do you need to measure?

Student: the bee's path?

Tutor: Right. Now, the bee's displacement is a vector. Is the distance a vector or a scalar?

Student: Scalar.

Tutor: So then distance traveled is the scalar counterpart to displacement. Now thinking about our analogy, what is the difference between speed and velocity?:

Student: Speed is a scalar, and velocity is a vector.

Insights towards an ongoing research focus…

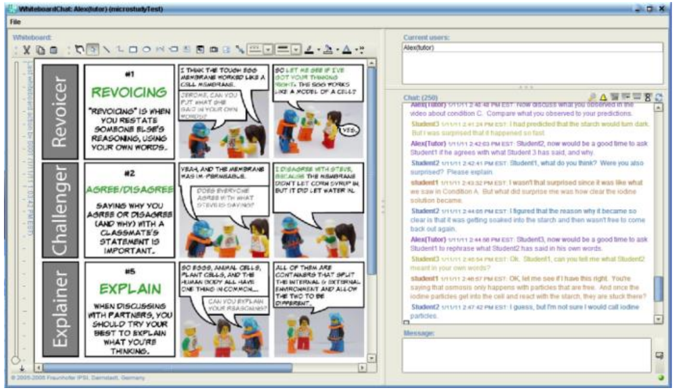
- Human tutoring not always better than non-interactive support (VanLehn et al., 2007)
  - Focus shift to capturing what it is about interaction that is effective for instruction
- Effective human tutors guide students towards opportunities for reflection (Rosé & Torrey, 2004)
- Student interaction with dialogue agents lacks evidence of reflection (Rosé et al., 2003)
- Students expect to behave differently with agents than with humans (Rosé & Torrey, 2004)

SOCIAL Interaction as a Learning Resource

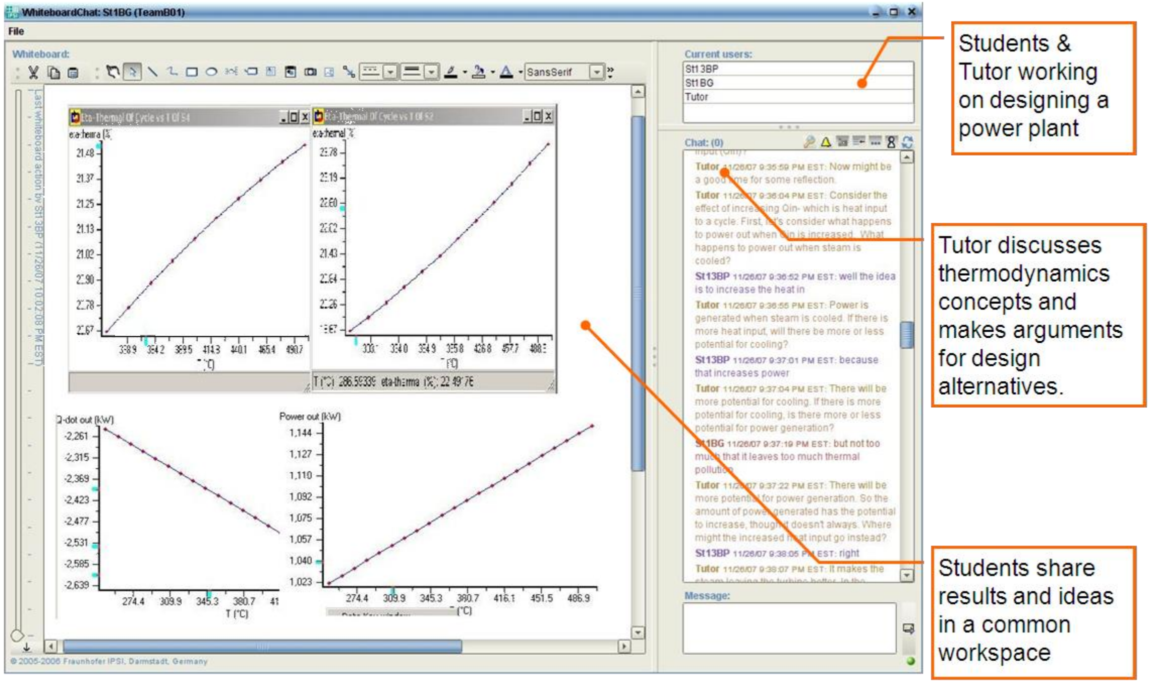
- Students can benefit from working with another student, even in the absence of scaffolding (Gweon et al., 2006; Kumar et al., 2007)

- Students gain as much from a human partner as from a carefully crafted tutor agent (*Kumar et al., 2007*)

- Context sensitive support for collaboration is more effective than static support (*Kumar et al., 2007*)



Conversational Agent Based Support in Computer Supported Collaborative Learning



Students learn 1.24 s.d. more when working with a partner and automated support than students working alone (Kumar et al., 2007)

Effective in Multiple Learning Contexts

- A decade and a half of successful classroom studies
  - Middle school, High school, College level
  - Urban school districts
  - Top tier and second tier universities
  - Math, Science, Engineering, Social Sciences

- Massive Open Online Courses (MOOCs)
  - Demonstrates that success generalizes to massive scale

Empirical Support for Design Principles

- *Personalized agents* increase supportiveness and help exchange between students (*Kumar et al., 2007*)
- Agents are more effective when ***students have control*** over timing of the interaction (*Chaudhuri et al., 2008; Chaudhuri et al., 2009*)
- Agents that employ *Balesian* ***social strategies*** are more effective than those that do not (*Kumar et al., 2010; Ai et al., 2010*)
- Students are sensitive to agent *rhetorical strategies* such as displayed bias (*Ai et al., 2010*), **displayed openness** to alternative perspectives (Kumar et al., 2011), and targeted elicitation (*Howley et al., 2012*)
- **Accountable talk agents** (*Dyke et al., 2013; Adamson et al., 2014*)

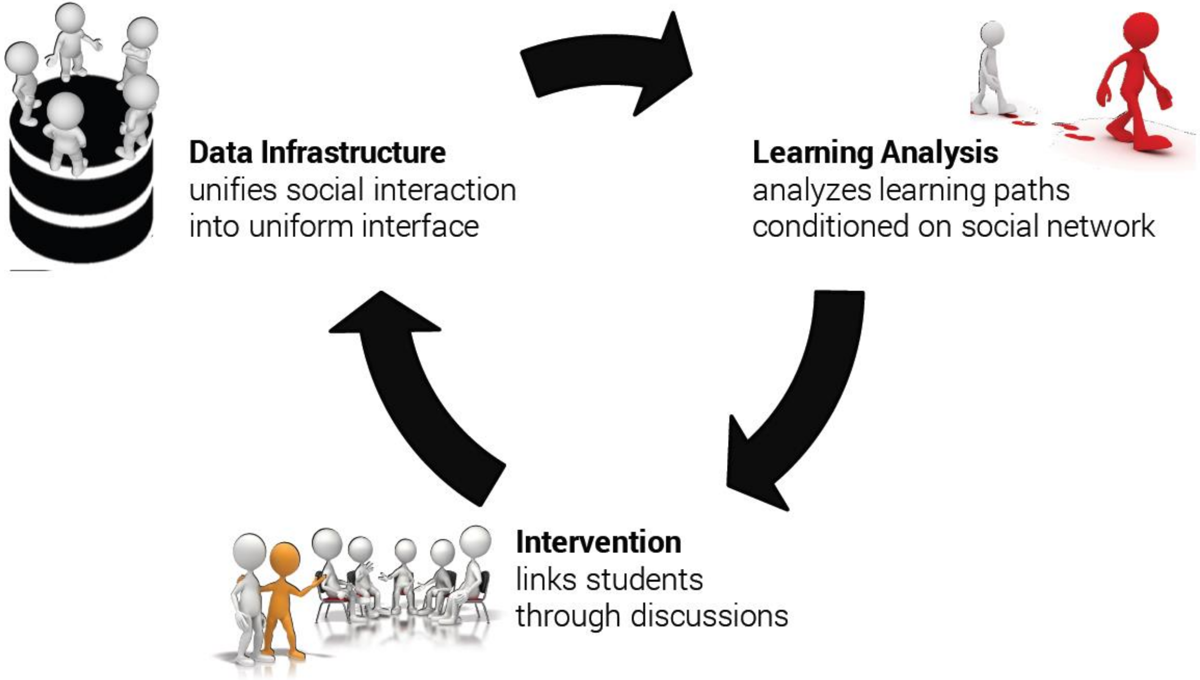
Effective
Collaborative Learning
is rare
without support

Technology Support for Collaborative Learning

Outline

- Historical Overview: 18 years of work
- **Technology Approach: Highlighting Chinese Discourse Analysis tools**
- Recent Work on Advanced Discourse Analysis Techniques
- Conclusion

From Data to Design



New Partnership CMU-BNU

未|来|教|育|高|精|尖|创|新|中|心
Advanced Innovation Center for Future Education
AICFE

AICFE BNU

Advanced Innovation Center for Future Education Beijing Normal University

北京 海淀区 | 学术研究

目前就职	Beijing Advanced Innovation Center for Future Education
教育背景	Beijing Normal University
网站链接	公司网站

learnsphere.org

LearnSphere

Tigris Explore Leadership About Help Workshops

Current Resources

Tigris is a workflow authoring tool which is part of the community software infrastructure being built for the LearnSphere project. The platform will provide a way to create custom analyses and interact with new as well as existing data formats and repositories, such as DataShop, MOOCdb, DiscourseDB and DataStage.

Tigris@CMU

Tigris@Memphis

Tigris@Stanford

Tigris@TutorGen

DATASHOP

MOOCDB

DATASTAGE

DISCOURSEDB

TIGRIS

Secure | https://discoursedb.github.io

View on GitHub

DiscourseDB

As part of a broader effort to provide tools for enabling research and practice in the space of collaborative and Discussion based learning, DiscourseDB is an NSF funded data infrustructure project designed to bridge data sources from multiple platforms for hosting those learning experiences. Our vision is to provide a common data model designed to accommodate data from diverse sources including but not limited to **Chat, Threaded Discussions, Blogs, Twitter, Wikis, and Text messaging**.

We will make available analytics components related to constructs including **role taking, help exchange, collaborative knowledge construction, showing openness, taking an authoritative stance, attitudes, confusion, alliance and opposition**. In enabling application of such metrics across datasets from multiple platforms, research questions related to the mediating and moderating effect of these process and state measures on information transfer, learning, and attrition can be conducted, building on the earlier research of our team.

Current Capabilities

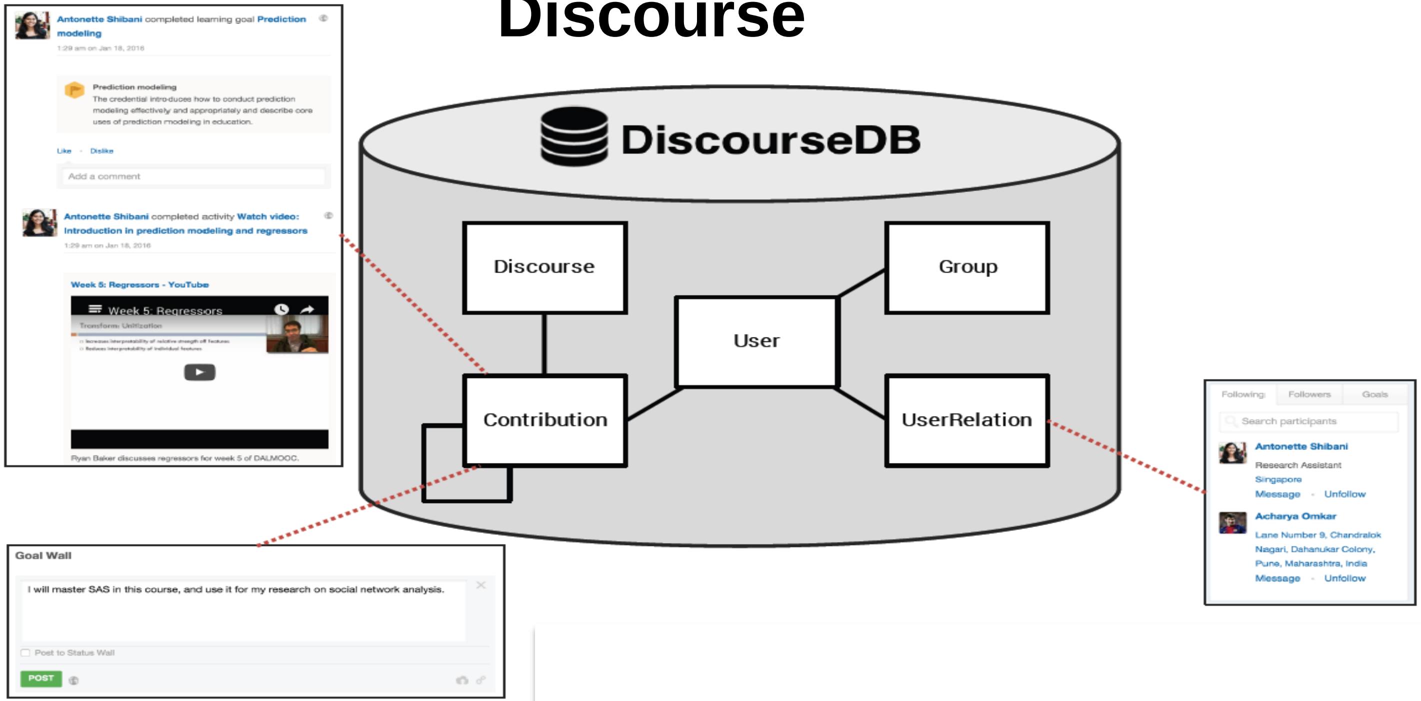
We have one publically available dataset, consisting of online discussion of bugs and features in a set of related open source software projects, OpenFL. Other datasets are available to researchers by request, subject to IRB approval.

Vision

Facilitate analysis
of discussion data
across multiple research sites
and multiple platforms

(e.g., Knowledge Forum and Idea Thread Mapper,
or Knowledge Forum and a Wiki)

A Common Representation of Discourse



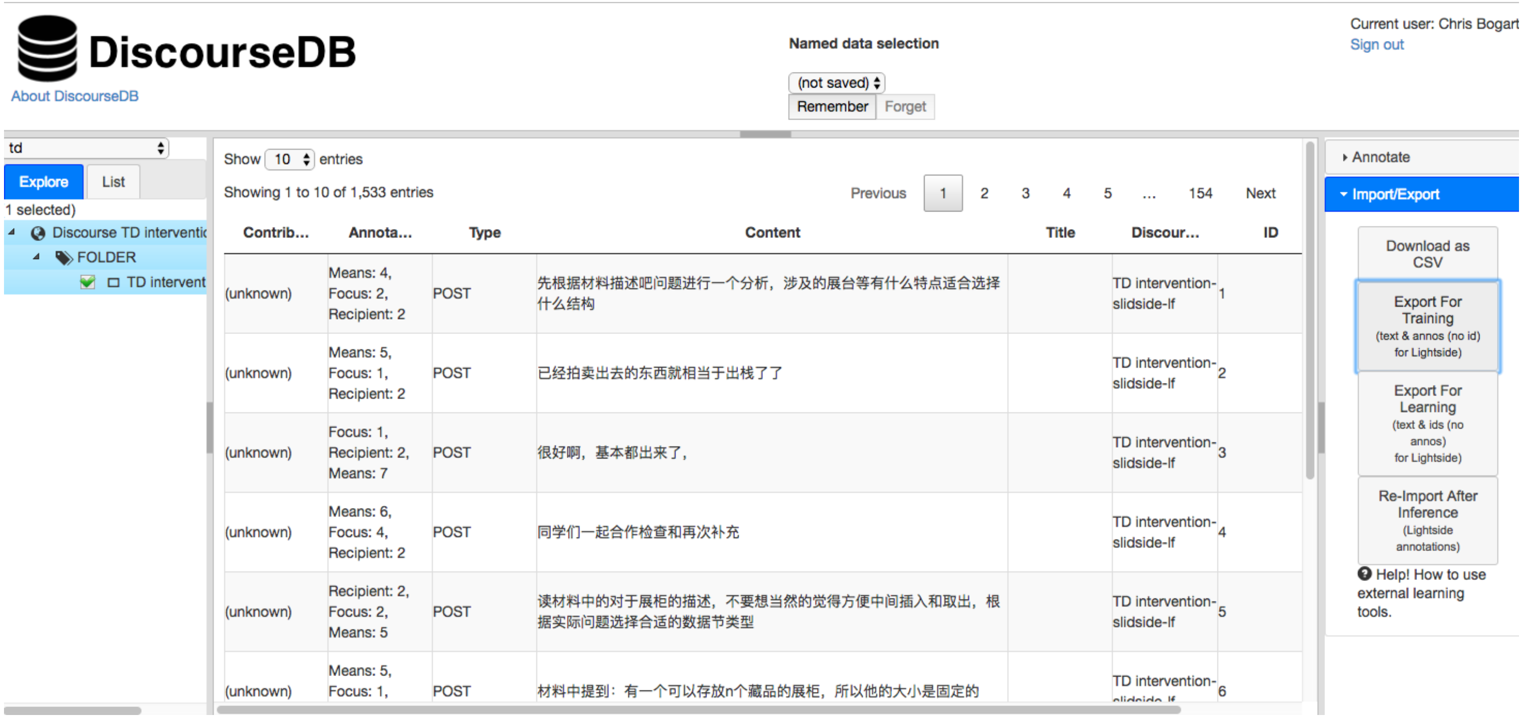
Data Analytics Pipeline

- Browse data in DiscourseDB
- Import/Export data

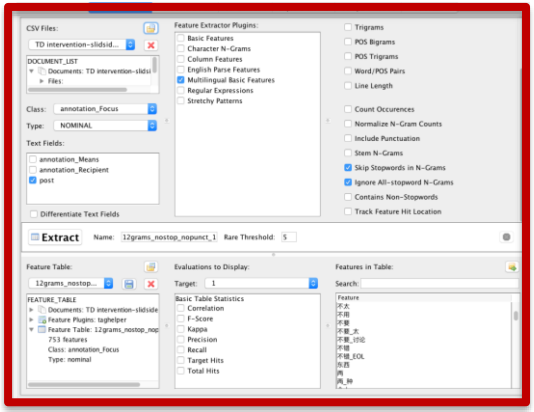
- View, manipulate, create Annotations
- http://brat.nlplab.org

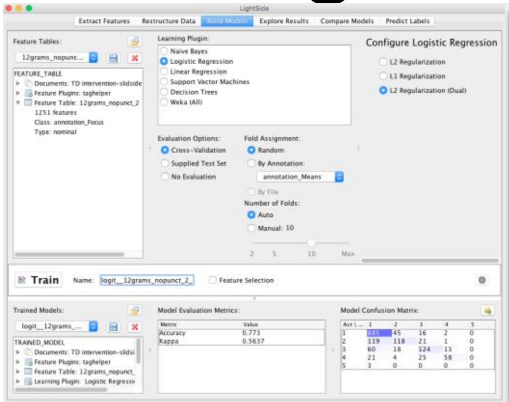
- Use annotations on DiscourseDB data to train models.
- Use models to annotate DiscourseDB data
- http://ankara.lti.cs.cmu.edu/side

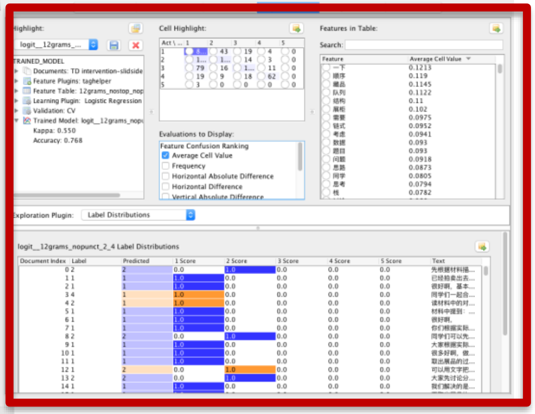
Chinese DiscourseDB

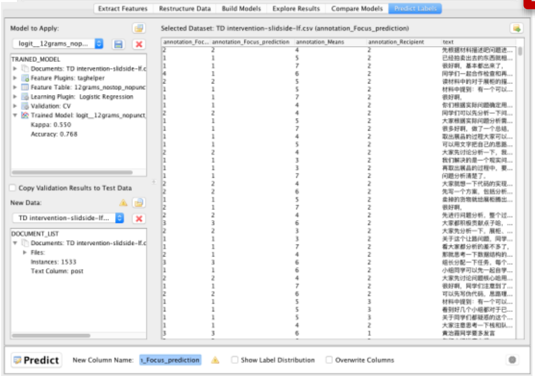


Chinese LightSIDE









Feature Engineering

Error Analysis

Error Analysis Process

High Level Overview

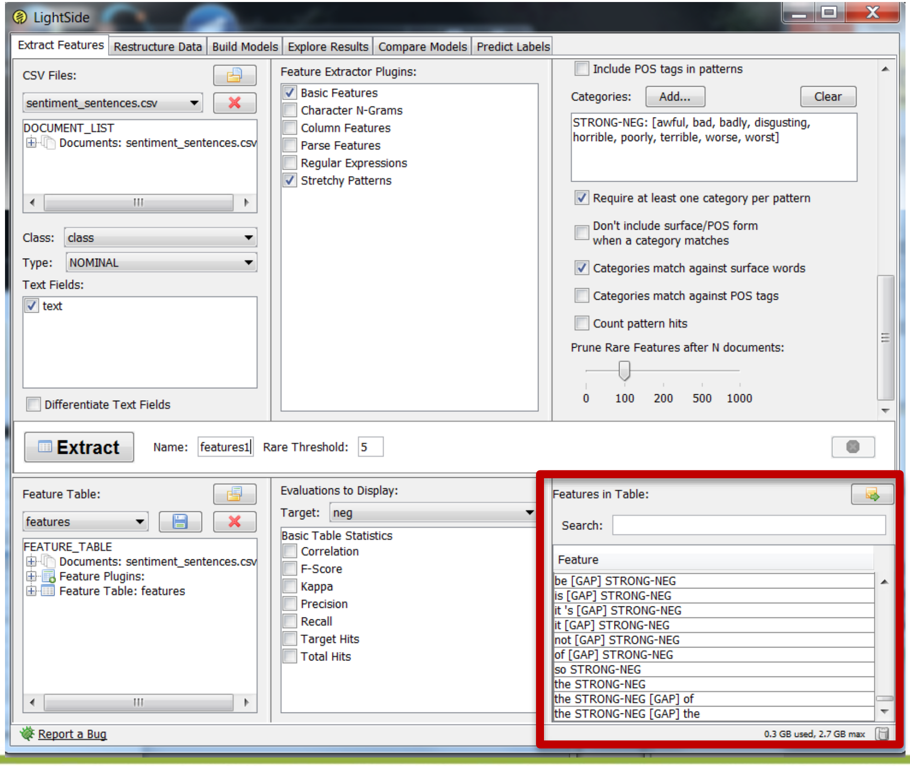
```
=== Confusion Matrix ===

  a   b   <-- classified as
 44  25 |
 10 149 |
```

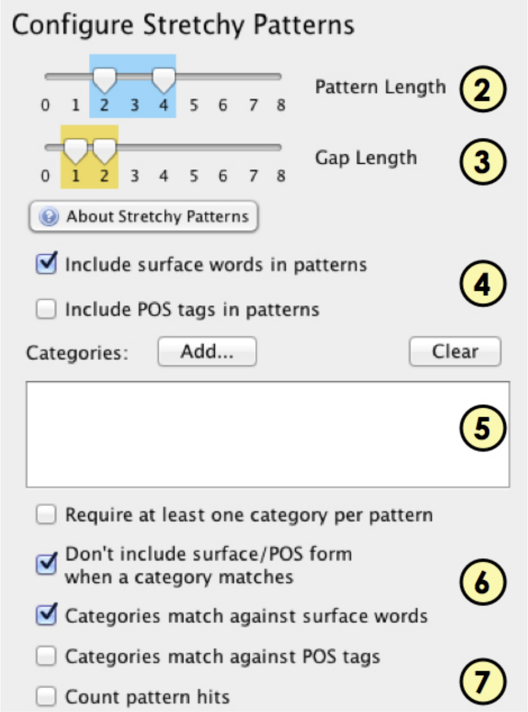
Goal: We want to discover how to re-represent the data so that instances with the same class value look more similar to one another and instances with different class values look more different

- Identify large error cells
- Make comparisons
  - Ask yourself how it is similar to the instances that were correctly classified with the same class (**vertical comparison**)
  - How it is different from those it was incorrectly not classified as (**horizontal comparison**)

Stretchy Patterns



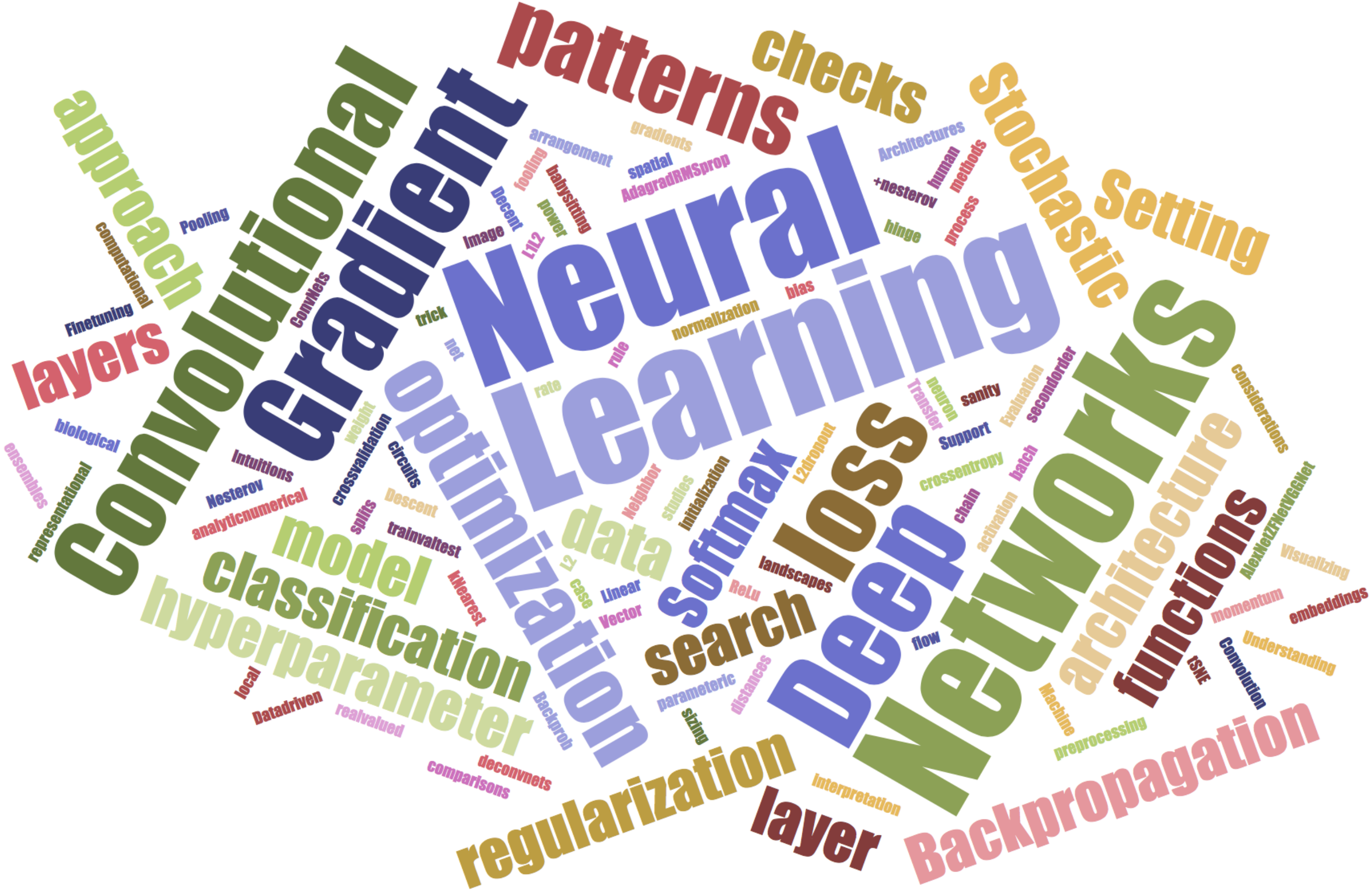
Configuring Stretchy Patterns

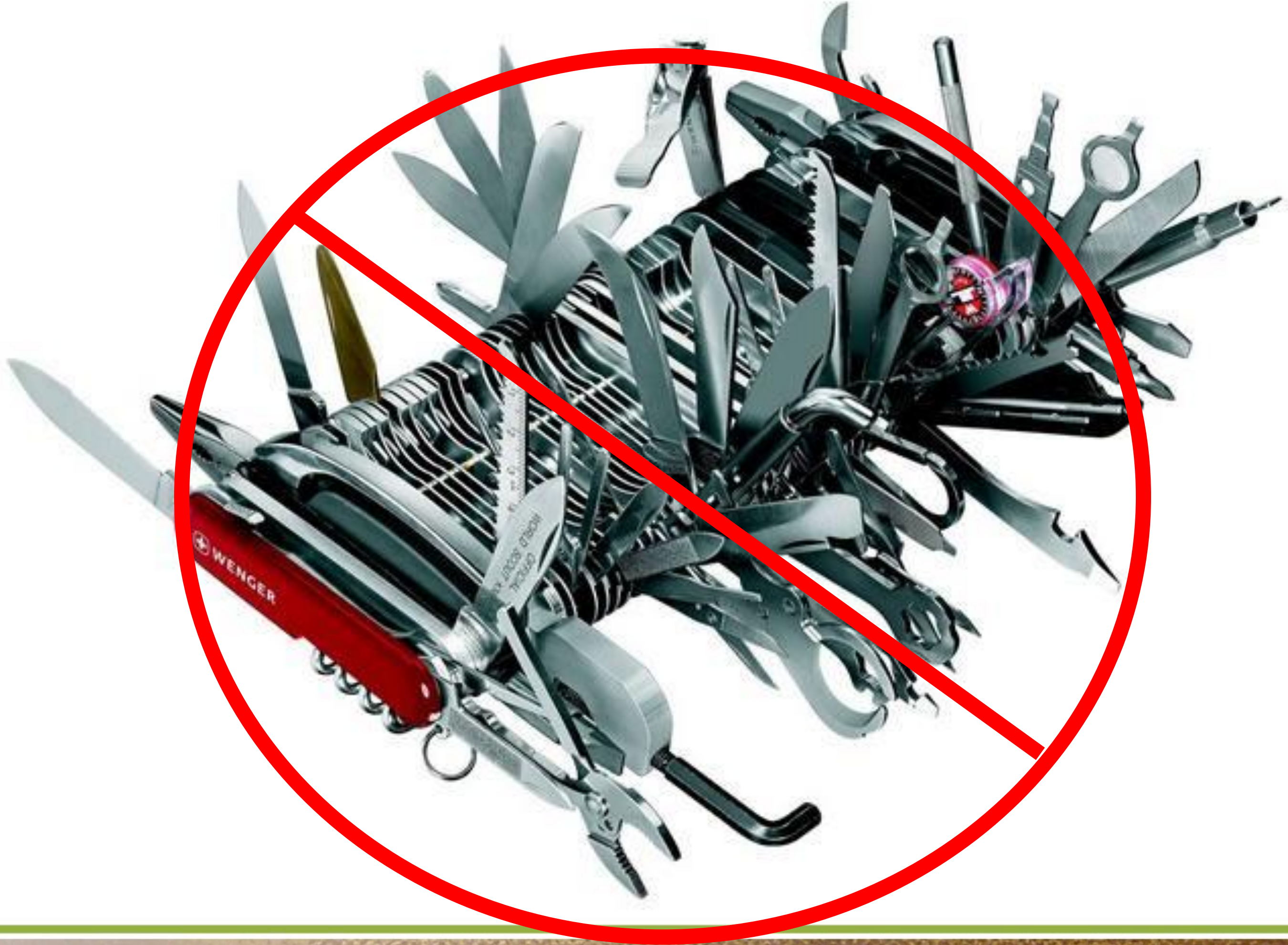


- Longer patterns and longer gaps lead to larger numbers of features
- Categories are useful both for abstraction and for anchoring the

Outline

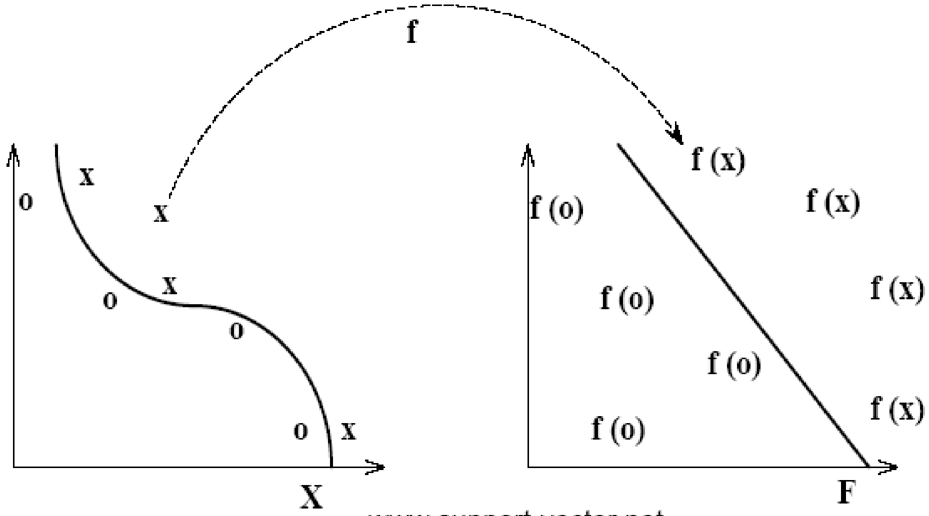
- Historical Overview: 18 years of work
- Technology Approach: Highlighting Chinese Discourse Analysis tools
- **Recent Work on Advanced Discourse Analysis Techniques**
- Conclusion


patterns
checks
Stochastic
Setting
Neural
Learning
Networks
Deep
loss
Gradient
Convolutional
approach
layers
optimization
classification
hyperparameter
model
data
Softmax
search
regularization
layer
Backpropagation
architecture
functions
gradients
arrangement
spatial
AdagradRMSprop
Architectures
human
methods
+nesterov
process
hinge
babysitting
fooling
Decent
power
image
L1L2
Pooling
computational
Finetuning
ConvNets
trick
normalization
bias
rule
net
rate
Transfer
neuron
sanity
Evaluation
secondorder
considerations
Support
crossentropy
batch
L2dropout
initialization
studies
Neighbor
weight
crossvalidation
circuits
Intuitions
Nesterov
analyticnumerical
Descent
splits
trainvaltest
biological
ensembles
representational
kNearest
L2
case
Linear
Vector
ReLu
landscapes
chain
activation
AlexNetZFNetVGGNet
Visualizing
momentum
embeddings
tSNE
Convolution
Understanding
Machine
preprocessing
flow
parametric
distances
sizing
Backprop
local
Datadriven
realvalued
deconvnets
comparisons
interpretation


WENGER

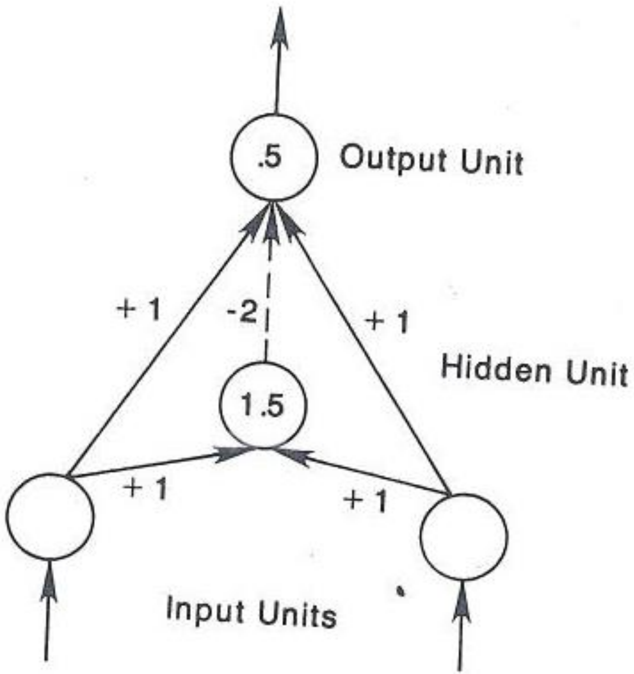
Building Blocks: Types of Nonlinear Transformation

- Map data into a feature space where they are linearly separable $x \rightarrow \phi(x)$



www.support-vector.net

XOR: Non Separable Function



Building Blocks: Types of Nonlinear Transformation

- Context and Co-occurrence
  - Common: PCA, LSA, LDA
  - Neural: Skip Gram models (Embeddings), Autoencoders
- Sequences
  - Common: HMMs, DBNs, Lag models, other time series models
  - Neural: Recurrent networks, LSTM, BiLSTM
- Filters
  - Common: SVM with nonlinear kernels, Filters and templates as feature extractors
  - Neural: Convolutional Networks, Additional fully-connected (hidden) layers

Recent work on Transactivity Detection

Building Reasoning Together

Transactivity

Transactivity

Transact: elaborate, build upon, question, or argue against the ideas presented by his/her partners [Berkowitz & Gibbs, 1983]

"I recommend nuclear energy for this city since it's very efficient."

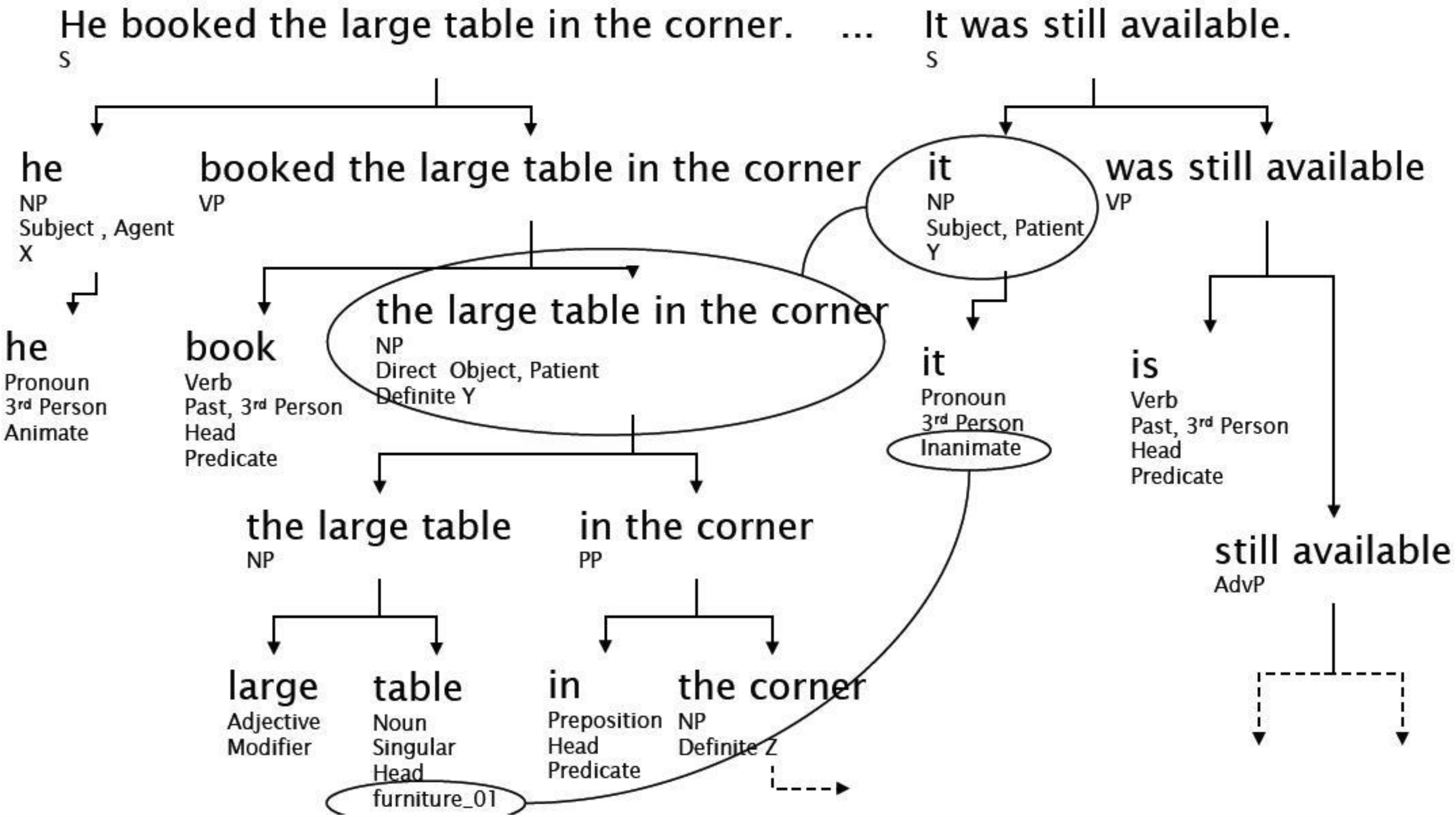
Transactive:

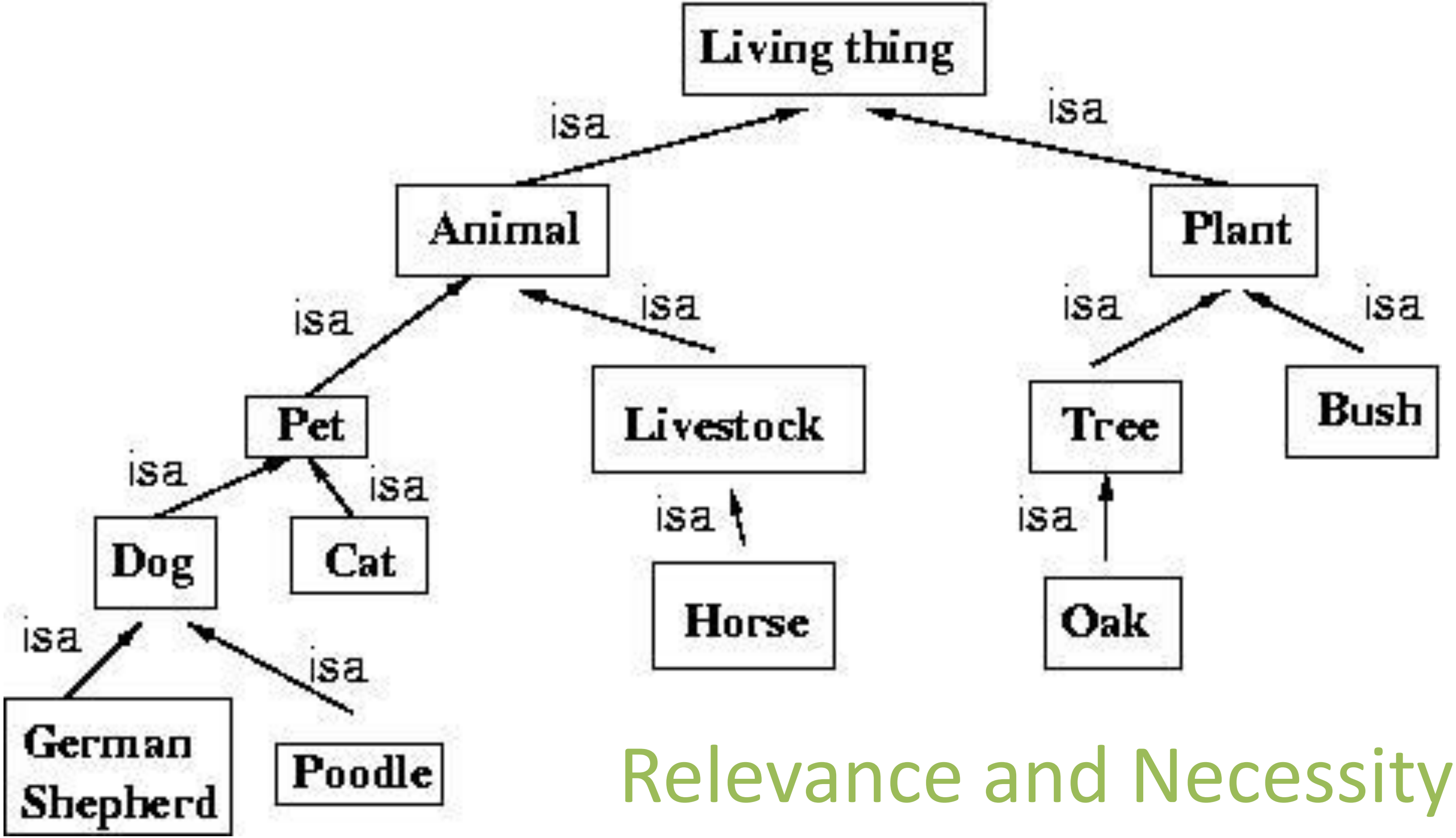
"Nuclear energy, as it is efficient, it is not sustainable. Also, think of the disaster probabilities."

Non-transactive:

"I agree that nuclear power would be the best solution."

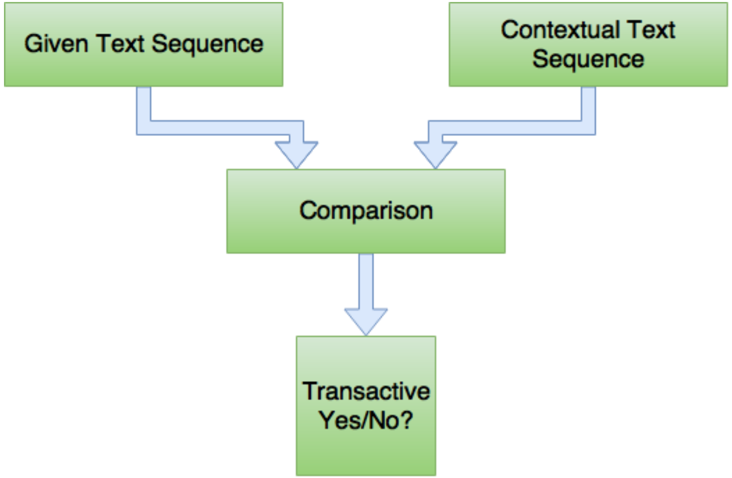

TEXT
PRAGMATICS
SEMANTICS
SYNTAX
MORPHOLOGY
PHONOLOGY
PHONETICS


He booked the large table in the corner. ... It was still available.
S
S
he
NP
Subject , Agent
X
booked the large table in the corner
VP
it
NP
Subject, Patient
Y
was still available
VP
he
Pronoun
3rd Person
Animate
book
Verb
Past, 3rd Person
Head
Predicate
the large table in the corner
NP
Direct Object, Patient
Definite Y
it
Pronoun
3rd Person
Inanimate
is
Verb
Past, 3rd Person
Head
Predicate
the large table
NP
in the corner
PP
still available
AdvP
large
Adjective
Modifier
table
Noun
Singular
Head
furniture_01
in
Preposition
Head
Predicate
the corner
NP
Definite Z



Relevance and Necessity

The Task Envisioned



Building Blocks: Types of Nonlinear Transformation

- Context and Co-occurrence
  - Common: PCA, LSA, LDA
  - Neural: Skip Gram models (Embeddings), Autoencoders
- Sequences
  - Common: HMMs, DBNs, Lag models, other time series models
  - Neural: Recurrent networks, LSTM, BiLSTM
- Filters
  - Common: SVM with nonlinear kernels, Filters and templates as feature extractors
  - Neural: Convolutional Networks, Additional fully-connected (hidden) layers

Task Challenges

- Limited quantity of annotated data for transactivity.
- Annotated data is in a single domain.
- Proposed solution – using entailment:
  - Use pre-trained semantic vector models from a large data set and, given a method of comparison between the vectors, determine if a text is transactive as compared to another text.

Entailment: Wikipedia Definition

- In semantics, entailments depend on the ″dictionary definition″ of the words in question.

Relevance and Necessity

- To judge whether an entailment is true, one can ask, ″Could it *ever* be the case that *B* isn't true while *A* is true?″

Entailment: Example

- Example from M. Lynne Murphy's *Lexical Meaning*
- "If it is a shoe, then it is made to be worn on a foot."

Leveraging Entailment as a Pretraining Task

Transactivity

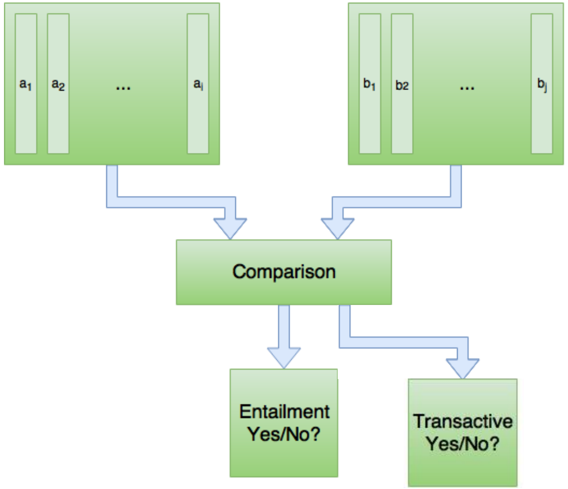
Entailment

Reframing the Task

- Learning inference (entailment, contradiction, neutral) with Stanford Natural Language Inference dataset.

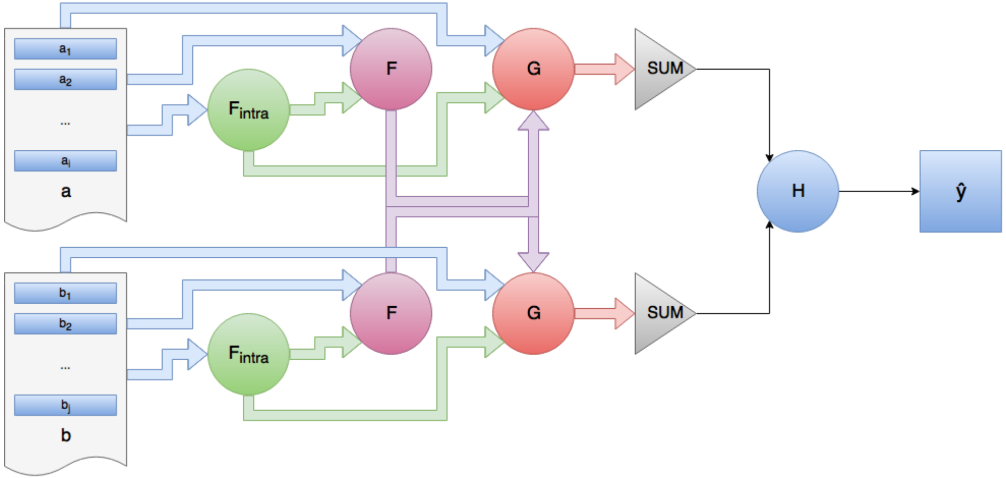
- Starting with word embeddings.
  - GloVe embeddings

- Classification

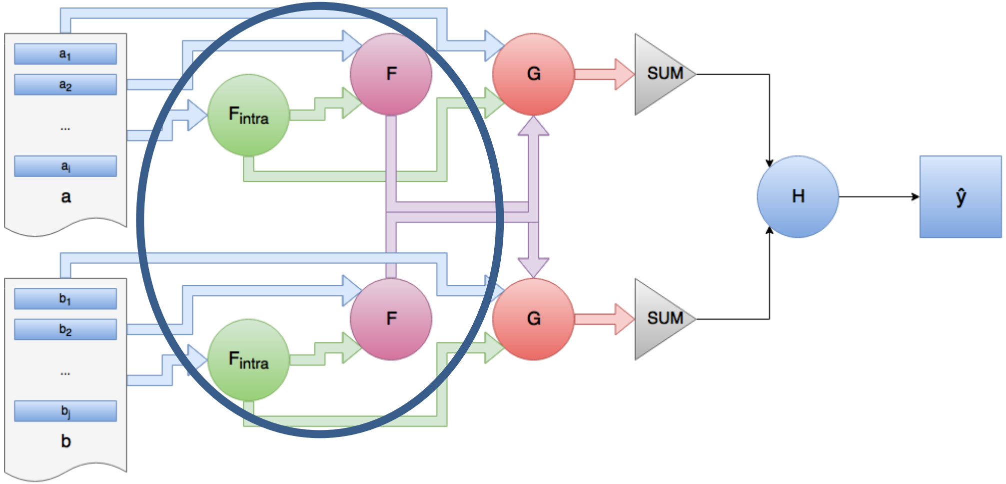


Foundation: Deep Learning for Entailment Detection

Parikh et al. 2016



Step 1: Attend



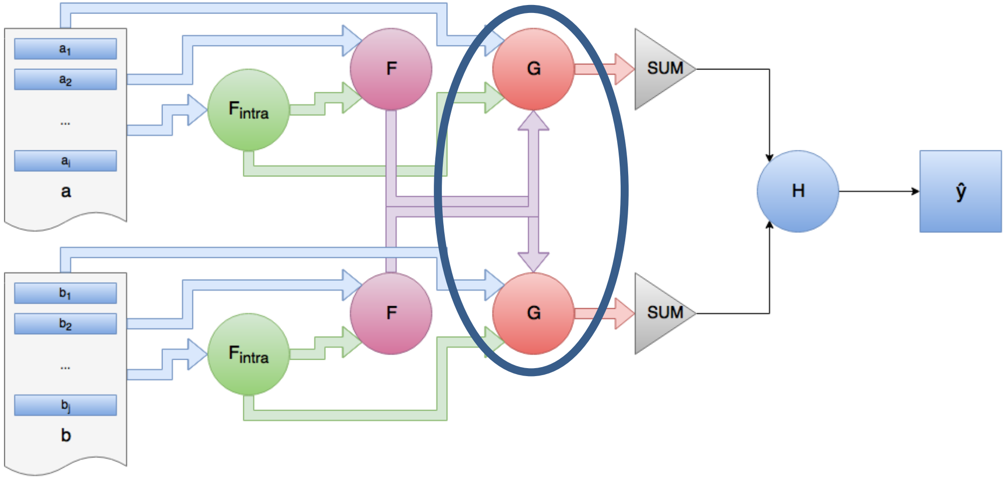
Step 1: Attend

- For each pair of words in the two posts, determine some attention score via 2 layer dense feed-forward neural network.

- For each word in each post, average all the attention scores with relation to the other post.

- What you get:
  - Information that indicates how important each word in a given post is with respect to the other post.

Step 2: Compare

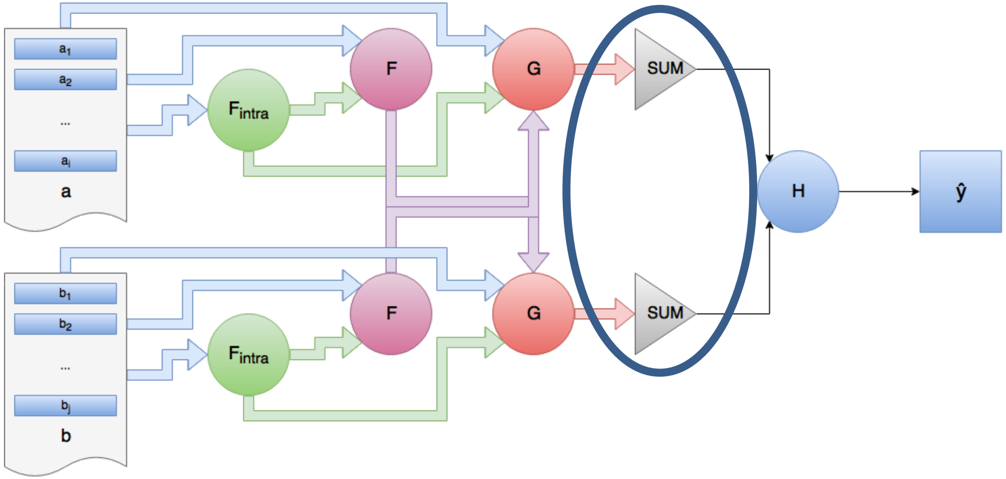


Step 2: Compare

- Using the representation from the attention step along with the corresponding vectorized input post, run though a 2 layer dense feedforward neural network.

- What you get:
  - Two sets of vectors for that contain information comparing the posts with respect to each other.

Step 3: Aggregate

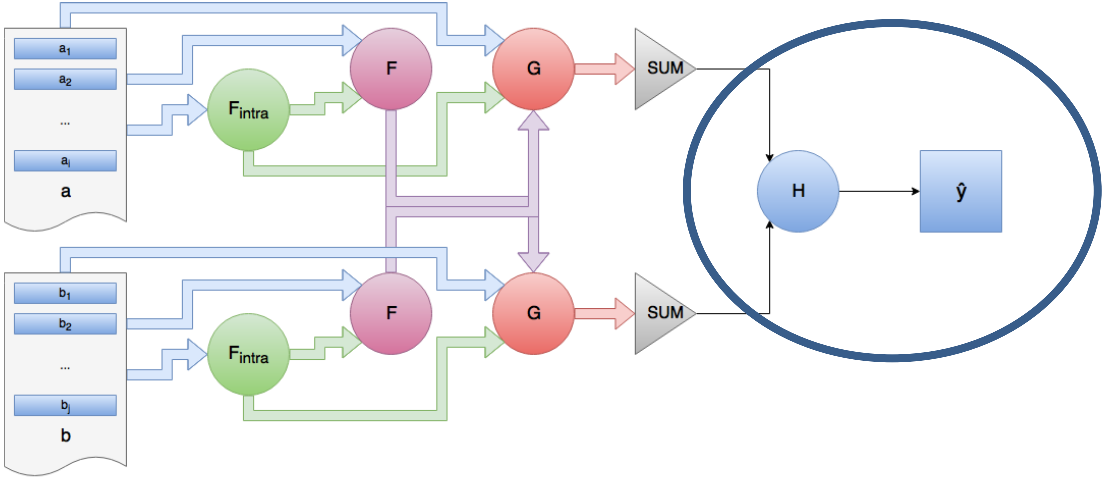


Step 3: Aggregate

- Sum each set of comparison vectors into two one dimensional vectors.

- Each of these vectors is a representation of a given post in relation to the other.

Step 4: Classify



Step 4: Classify

- With the resulting vectors from the aggregation step, we concatenate them and run them through another 2 layer dense feedforward neural network with cross-entropy loss to classify the data.

Experiment 1

1. Train Entailment task first
2. Use trained weights as initialization for Transactivity task
3. Train on Transactivity task

Entailment Dataset

- Stanford Natural Language Inference Corpus, Bowman et al. 2015.
- Collection of 570,000 English sentence pairs labeled for balanced classification of entailment, contradiction, and neutral.
- Examples were generated by humans in response to sentences describing pictures from Flickr

- Example:
  - Sentence1: “A soccer game with multiple males playing.”
  - Sentence2: “Some men are playing a sport.”

Transactivity Dataset

- Discussion data from online forum where students offered feedback to one another on their proposals for city power plans
- 476 human annotated posts.
- Example:
  - Sentence 1:
    "But if the energy is saving them some money it could go towards the batteries. Whats
    frustrating is that it doesn't really give us information regarding the costs of generating electricity currently."
  - Sentence 2:
    "But those batteries add even more cost, and for a city concerned with cost, that would be a
    problem. Plus, without the batteries, it's not very reliable, and that's also a problem for a touristry driven economy."

Results, Part 1

Model	Accuracy	Cohen's Kappa
Logistic Regression with unigrams	0.795	0.510
Logistic Regression with embeddings	0.626	0.182
Neural model	0.848	0.542

Experiment 2

- Transactivity prediction with in domain data vs. out of domain data
- Train the model as in experiment 1, however on each cross validation fold, evaluate the model on out of domain annotated transactivity data.
- Note that there is no point in which the model is trained on the out of domain data.

Out of Domain Transactivity Dataset

- 57 human annotated transacts from an Massive Open Online Course (MOOC) in which students were asked to design their own superheroes and provide feedback on other students’ designs.

Results, Part 2

Model	Accuracy (in | out)	Cohen's Kappa (in | out)
Logistic Regression with unigrams	0.795 | 0.667	0.510 | 0.376
Logistic Regression with embeddings	0.626 | 0.635	0.182 | 0.195
Neural model	0.848 | 0.824	0.542 | 0.586

Recent work on Rhetorical Structure Analysis

Rhetorical Structure in Student Writing

Dataset

- Research Writing Tutor Dataset (RWT) from Iowa State University
- 700-1000 documents for each Introduction, Methods, Results, and Discussion/Conclusion
- 120,000+ annotated sentences in a 90%/10% split

Prior Results by Elena Cotos

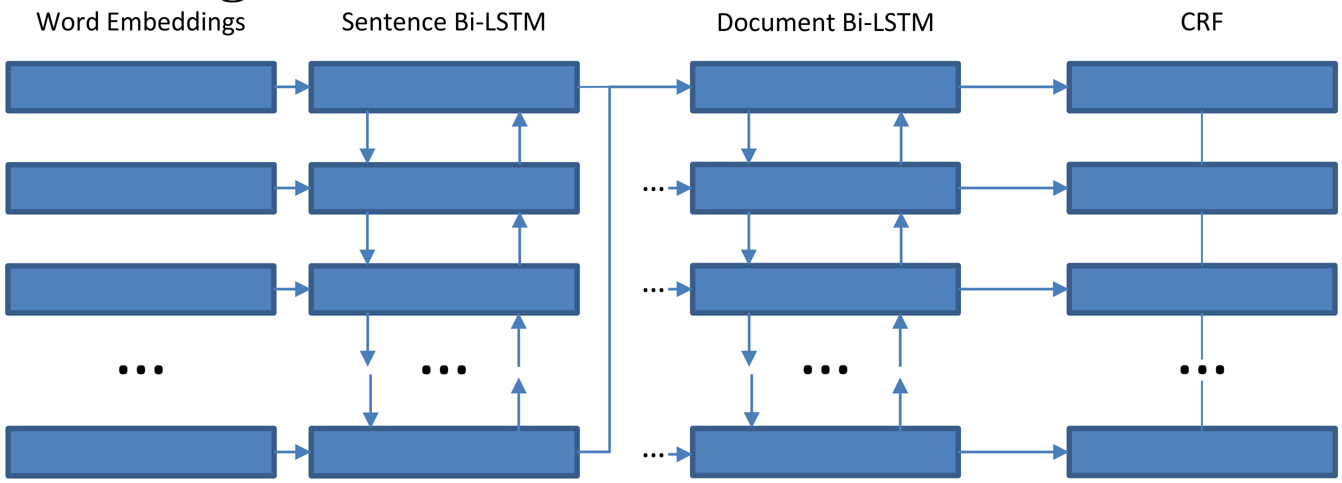
- Models: Naive Bayes, SVM, and MaxEnt
- N-Grams on stemmed text
- N-Grams on part-of-speech tags
- Unigram, bigram, and trigram for each stemmed text and part-of-speech

Kappa classification agreement for Step Classification

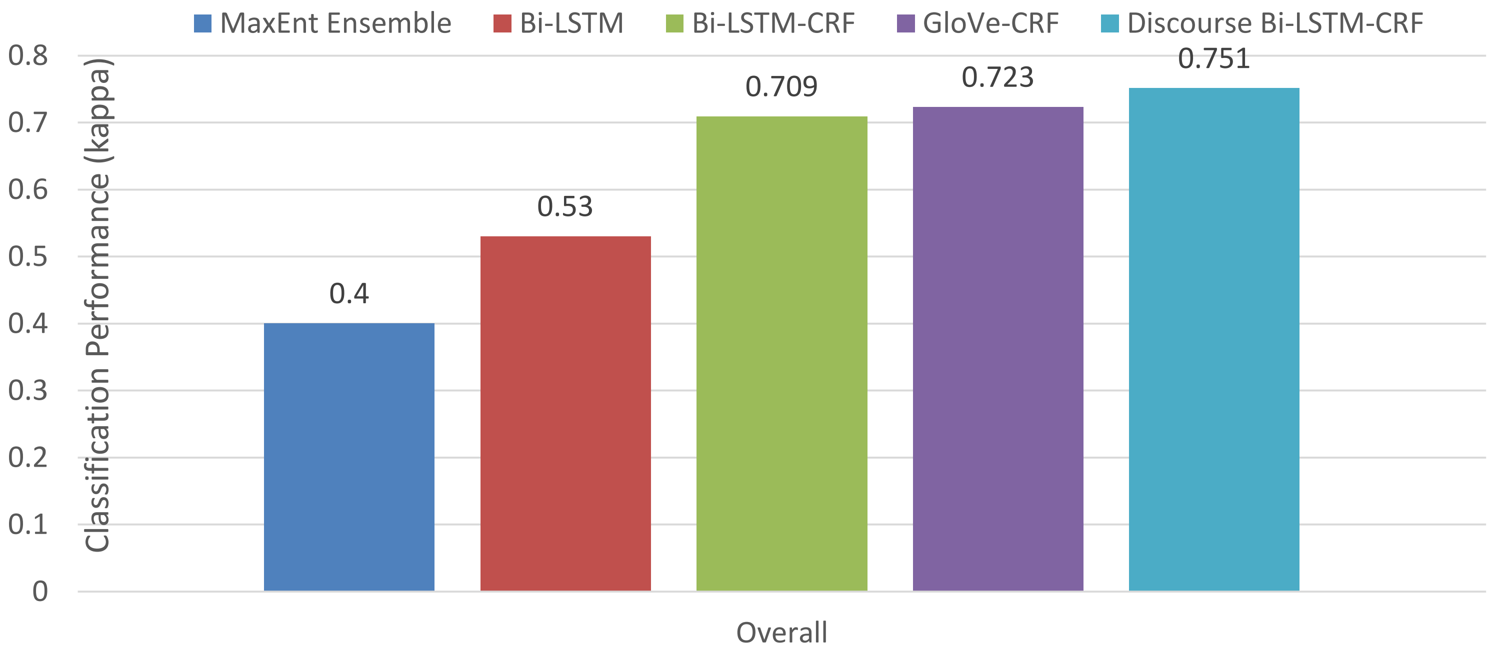
	Unigrams	Bigrams	Trigrams
Stemmed text	0.32	0.24	0.17
Part-of-speech	0.26	0.28	0.28

Sequence model

- Document-level Bi-LSTM-CRF (Huang et al. 2015; Song et al. 2017)



Results


MaxEnt Ensemble
Bi-LSTM
Bi-LSTM-CRF
GloVe-CRF
Discourse Bi-LSTM-CRF
Classification Performance (kappa)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.4
0.53
0.709
0.723
0.751
Overall

Outline

- Historical Overview: 18 years of work
- Technology Approach: Highlighting Chinese Discourse Analysis tools
- Recent Work on Advanced Discourse Analysis Techniques
- **Conclusion**

DANCE: Discussion Affordances for Natural Collaborative Exchange

JOIN | MAILING LIST | TALK SERIES | RESOURCES | DANCE SCHOLAR

About DANCE

Drawing from two decades of research in Computer Supported Collaborative Learning, we are working to design an extension of the edX platform to enhance instructionally beneficial discussion opportunities available to students. With this working group, we want to bring together people from academia and industry to build a common vision regarding what kinds of research would be valuable to the community once such a platform extension was in place to support it. Our work is initially focusing on the edX platform in particular, but in the long run we seek to provide these capabilities to Massive Open Online Courses and other online learning platforms more generally. In particular, this working group is partnering with edX as a satellite collaborative, seeking to involve researchers and developers from multiple universities, foundations, and industrial organizations.

Our foundational work is beginning with specific interventions designed to offer synchronous collaboration activities supported by intelligent conversational agents and enhancements to threaded discussions to support more intensive help exchange by leveraging social recommendation technology. However, our goals are much broader than this, seeking to leverage insights and methodologies from the field of Human-Computer Interaction and encompassing both synchronous and asynchronous communication very broadly. Our vision includes text, speech, and video based interactions, instrumented with all sorts of intelligent support powered by state-of-the-art analytics and leveraging language technologies and artificial intelligence more broadly in order to offer contextually appropriate support. We will coordinate this effort with regular online meetings and occasional in-person workshops.

Director: Carolyn Penstein Rosé
Carnegie Mellon University
cprose@cs.cmu.edu

Co-Director: Sreecharan Sankaranarayanan
Carnegie Mellon University
sreechas@cs.cmu.edu

DANCE is a community of practice with many open source resources

Thousands of visitors

Hundreds of return visitors each month

Resources

Resources

One goal of the DANCE initiative is to provide a community platform to help researchers in the MOOC/CSCL space interested in contributing to the OpenEdX platform. This page aggregates resources the came out of the DANCE effort as well as related third party tools and artifacts. It is our hope that this can inspire contributions by others and foster discussion among the community of CSCL researchers about possibilities for collaborative tools that can be deployed and tested on platforms such as OpenEdX.

DANCE XBlock Development

The DANCE discussion forum XBlock represents a first step towards improving scripted support for collaboration in MOOCs. The XBlock provides all basic features expected of a forum while augmenting the experience with social recommendation (that as an example recommendation application matches help seekers with help givers). It provides light social awareness and semi-synchronous interaction through a Personal Messaging capability. It also provides the forum data in a source agnostic data infrastructure model using DiscourseDB (see below) that will allow the contextualization and comparison of discourse data from across platforms.

- The source code and documentation can be found on GitHub.
- A walkthrough of its functionality can be found here.
- The features currently on offer from the XBlock as well as planned additions can be found here.
- The design document detailing the overarching goals and guiding principles behind the development of the XBlock can be found here.

DANCE Home

DANCE Discussion Forum is *compatible with Open edX*

Includes hooks for interventions like Social Recommendation and Discussion Scaffolding

Resources

LightSide

The open-source LightSide platform, including the machine-learning and feature-extraction core as well as the researcher's workbench UI, has been and continues to be funded in part through Carnegie Mellon University, in particular by grants from the National Science Foundation and the Office of Naval Research. See the full acknowledgements and grant details below!

We make the LightSide research platform freely available for research and education. In exchange, we ask that you provide us with basic information about who you are and how you're making use of LightSide's capabilities.

- You can download the current LightSide binaries and the user manual here
- The source code is freely available on GitHub.

Social Recommendation

Massive Open Online Courses have experienced a recent boom in interest. Problems students struggle with in the discussion forums, such as difficultly in finding interesting discussion opportunities or attracting helpers to address posted problems, provide new opportunities for recommender systems.

We developed a social recommendation technology to support help seekers in MOOC discussion forums implemented using a context-aware Matrix Factorization model to predict students' preferences for answering a given question. This recommendation framework allows for this two-way recommendation. The source code is freely available on GitHub.

References:

- Yang, Diyi, Piergallini, Mario, Howley, Iris and Carolyn Penstein Rose. "Forum thread recommendation for massive open online courses." Educational Data Mining 2014. 2014.
- Yang, Diyi, David Adamson, and Carolyn Penstein Rose. "Question recommendation with constraints for massive open online courses." Proceedings of the 8th ACM Conference on Recommender systems. ACM, 2014.

LightSIDE

Text mining tool bench

Over 10,000 users have downloaded LightSIDE

Automated collaborative process analysis

Automated writing assessment/feedback generation

Social Recommendation

deployed so far in one MOOC to support help exchange

Resources

DiscourseDB

DiscourseDB is an NSF funded data infrustructure project designed to bridge data sources from multiple platforms for hosting those learning experiences. Our vision is to provide a common data model designed to accommodate data from diverse sources including but not limited to Chat, Threaded Discussions, Blogs, Twitter, Wikis, and Text messaging.

We will make available analytics components related to constructs including role taking, help exchange, collaborative knowledge construction, showing openness, taking an authoritative stance, attitudes, confusion, alliance and opposition. In enabling application of such metrics across datasets from multiple platforms, research questions related to the mediating and moderating effect of these process and state measures on information transfer, learning, and attrition can be conducted, building on the earlier research of our team.

- Source code and additional documentation on GitHub
- Instructional video explaining the foundations of DiscourseDB
- Slides of the 1/2016 DiscourseDB HackDay

Bazaar

Bazaar is a publicly available architecture for orchestrating conversational agent based support for group learning. It is a powerful tool for facilitating research in collaborative learning. It hosts a library of reusable behavioral components that each trigger a simple form of support. More complex supportive interventions are constructed by orchestrating multiple simple behaviors. Its flexibility and simplicity mean it can be used to very rapidly develop platforms for investigating a wide range of important questions within the design space of dynamic support for collaborative learning.

- The source code is freely available on GitHub.
- Slides and video recordings of a Bazaar tutorial (May 2016) can be found here
- You can find additional documentation and links to relevant research papers here.

DiscourseDB:
Data infrastructure to offer discourse data and analytic tools through LearnSphere

Bazaar:
Tutorial dialogue architecture
Dialogue agents for individual or collaborative learning

Conclusion

- Language Technologies like Text Classification and Dialogue Agents help make collaboration effective
- Some resources have already been ported to Chinese
- **Join us**: We are happy to extend our work to collaborate with you
  - Let me know if you would like to collaborate