From Belief Change to Language Usage in Social Networks

Fenrong Liu
Tsinghua University

In this talk I will first present my joint work with Girard and Seligman on belief revision in social networks, and explain how an agent changes her belief/preference with consideration of her friends' attitude. To connect to themes of this workshop, I will discuss how this relates to some interesting phenomena in language usage in social networks. In particular, an agent prefers using some expressions that the other members of a community are using. I will end with some discussions and open research issues.

Reference

Fenrong Liu, Jeremy Seligman, and Patrick Girard. 2014. Logical Dynamics of Belief Change in the Community. *Synthese*. 191(11): 2403-2431.

The effect of the speaker’s communicative style and the listener’s pragmatic ability on irony comprehension: Evidence from ERPs

Xiaolu Wang, Binyao Huang
Zhejiang University

This study aims to assess an important issue in irony comprehension concerning when and how listeners integrate extra-linguistic information to compute the speakers' intended meaning in the dynamic process, from both the perspectives of speakers (communicative style) and listeners (pragmatic ability). To reach this end, the listener's pragmatic ability was measured by the indicator created by Niewland, Ditman & Kuperber (2010), and ERPs were recorded as the participants read the short passages that ended either with a literal or an ironic statement made by one of two speakers, who would not appear in the same discourse. The experiment was carried out in two sessions in which each speaker's use of irony was manipulated to a certain extent to create a different communicative style. In session one, 80% of the ironic statements were made by Speaker A who was supposed to be more ironic turn of mind, while Speaker B who was supposed to be less ironic turn of mind made only 20% of them. For Speaker B, an increased P600 was observed relative to literal utterances compared with ironic utterances. By contrast, both ironic and literal statements made by Speaker A elicited similar P600 amplitudes. In Session 2, both speakers' use of irony was balanced (i.e. 50% ironic, 50% literal). ERPs for Session 2 showed an irony-related P600 for Speaker A but not for Speaker B. Moreover, P200 amplitude was larger for sentences congruent with each speaker's communicative style (i.e. for irony made by the more ironic speaker, and for literal statements made by the less ironic speaker). The experimental results indicate that the listener can obtain the implied pragmatic knowledge relevant to the speaker's communicative style without definite clues, that the speaker's communicative style has impact on the listener's irony comprehension in both phases of early identification (200-300ms) and late integration (600-700ms), and that the listener’s pragmatic ability affects irony comprehension through the perception of the differences between speakers with communicative styles at the 200-300ms and 300-400ms time window.

Keywords: irony; communicative style; pragmatic ability; ERP

When Topic and Focus overlap with Contrast: Evidence from processing Chinese OSV and SOV sentences

Luming Wang[1], Xuping Li[2]

[1]School of Foreign Languages, Zhejiang University of Technology

[2]Center for the Study of Language and Cognition, Zhejiang University

The human brain is highly sensitive to mapping syntax and information structure of a sentence in order to achieve an efficient communication. Previous electrophysiological studies from Chinese suggest that the parser prefers the noun phrase at the clause-initial position (NP1) to be the topic [1]. However, there seems to be no dedicated syntactic position for contrast, and it can be overlapped with a topic (contrastive topic) or with a focus (contrastive focus) [2]. Chinese OSV and SOV sentences provide a good testing ground for examining the interface between syntactic position and information structure (including the factor of contrast). Whereas the clause-initial object in OSV is undisputably topical, the preverbal object in SOV is either identified as a secondary topic following the primary topic S, or as a focus that usually requires a second clause with the focal objects in contrast [3]. We recruited 30 native speakers of Chinese to participate our acceptability study with question-answer pairs as shown in Example (1).

Three factors were manipulated: Context (1: topicalize S vs. 2: topicalize O), Order (OSV vs. SOV) and Contrastive Position (contrastive S vs. contrastive O), resulting in 8 critical conditions. Contrastive NPs are capitalized in the literal translation of original Chinese sentences:

(1) Context 1: what about Xiaowang? Context 2: what about apple?

a SOV-S XIAOWANG apple ate, XIAOZHANG didn't eat.

b SOV-O Xiaowang APPLE ate, BANANA didn't eat.

c OSV-S Apple XIAOWANG ate, XIAOZHANG didn't eat.

d OSV-O APPLE Xiaowang ate, BANANA didn't eat.

+ filler sentences including SVO, SOV/OSV, SV, and ungrammatical OVS sentences

36 question-answer pairs per condition were constructed, which were assigned to 6 lists in a Latin Square design to avoid lexical repetition. For each list, there were 24 critical sentence interspersed by 36 filler sentences. Participants were asked to judge whether or not the question-answer

pairs sound natural based on a 4-point scale with 4 being the most acceptable point and 1 the least one.

Our results support a general preference for topic-continuity at the NP1, i.e. sentences starting with the topic NPs were accepted significantly higher than those not (e.g., 1a/1b > 1c/1d). Additionally, the acceptability difference in Context 2 mirrored with Context 1, but in a reversed direction, i.e. 1b > 1a > 1d > 1c while 2c > 2d > 2a > 2b. This further supports that topic-continuity also holds across clauses: when the contrastive NP served to maintain the same topic as in the first clause (e.g., 1b), the sentence received the highest acceptability. However, when it leads to topic-shift in the second clause (e.g.,1a), the acceptability reduced significantly. The consistent differences observed in both contexts thus lend behavioral evidence for topic-continuity being a general preference.

Finally, our results support a topic-focus order. When the NP1 was the topic, sentences with a focus in contrast were more acceptable than those with topic in contrast, especially in SOV sentence. However, to disentangle a context-induced contrast from a structural-induced contrast (SOV) requires further studies on online sentence processing.

Selected References

[1] Hung, Y.-C. & Schumacher P.B. (2012). *Neuroscience Letters*, 511, 59-64

[2] Neeleman & Vermeulen (2012). Walter de Gruyter.

[3] Wang & Schumacher P.B. (2013). *Frontiers in Psychology*, 4, 363.

A dynamic account of verb doubling clefts construction in Chinese

Xiaolong Yang, Yicheng Wu*

Center for the Study of Language and Cognition, Zhejiang University

Abstract: Cheng and Vicente (2013) claim that verb doubling in one specific construction may lead to a topic/focus distinction, which is called as verb doubling clefts construction (Henceforth VDCC), exemplified in (1):

(1) Q: *Ni chiguofan meiyou?*

you eat EXP food not. have

'Have you eaten food already?'

A:*Chi*, w*oshichi guole*, *buguo*….

eat I COP eat EXP LE but

'As for eating, I have eaten, but…'

According to Cheng and Vicente (2013), the first verb *chi*'eat' can be interpreted as a topic; the second verb *chi*'eat' should be construed as a focus. They further show that the two verbs stand in an A-bar movement relation:

In this paper, within the framework of Dynamic Syntax (Kempson et al. 2001; Cann et al. 2005), we demonstrate that VDCC is interpreted and produced in line with the principle of linearity. The DS paradigm seeks to develop a parsing-based grammar formalism for characterizing structural properties of language by modeling the dynamic process of semantic interpretation which is defined over the left–right sequence of words uttered in context. What is distinct about this model is that syntactic explanations can be grounded in the time-linear projection of the requisite predicate-argument structure. Syntactic mechanisms are defined as a set of licensing actions for inducing semantic content, incrementally, on a word-by-word basis.

Under the dynamic analysis, the first verb is treated as an elliptical topic which is dependent on the discourse. The parse of the first verb is actually a process of reusing the structure established previously representing the same event, as can be seen in fig. 1:

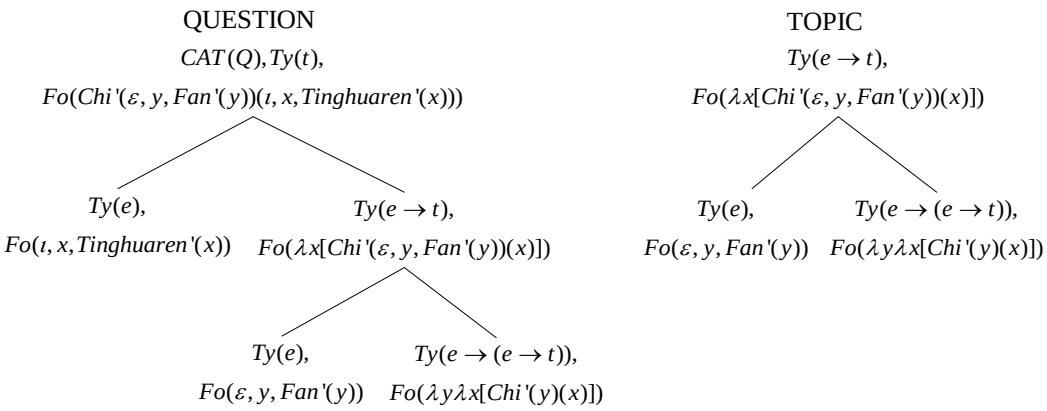


Fig. 1

Shi is analyzed as a predicate pro-form, projecting a one-place predicate node with an outstanding requirement, which will be satisfied by unifying the unfixed node constructed by the second verb. The second verb is an elliptical structure as well. The lexical entry of the second verb in VDCC creates an internal argument node, which projects a metavariable whose semantic value will be provided by the topic, which is in fact a process of recovering some content from the immediate context:

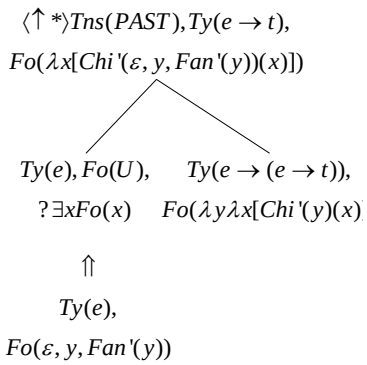


Fig. 2

This paper presents a dynamic account of verb doubling clefts construction in Standard Chinese, with a particular focus on the two verbs within it. The parse of the first verb is shown to be a process of reusing the structure established previously representing the same event. The second verb is also treated as an elliptical form recovering its content based on the topic, which further identifies with the claim that topic preferably precedes comment both in production and comprehension.

References:

Cann, R., Kempson, R., Marten, L., 2005. The Dynamics of Language. Elsevier, Oxford.

Cheng, L.-S.L., Vicente, L. 2013.Verb doubling in Mandarin Chinese. Journal of East Asian Lingusitics 22, 1-37.

Kempson, R., Meyer-Viol, W., Gabbay, D., 2001. Dynamic Syntax: The Flow of Language Understanding. Blackwell, Oxford.

Introduction to formal argumentation

Nir Oren
University of Aberdeen

Argumentation theory is a form of non-monotonic reasoning which borrows ideas from models of human discussion, and since Dung's seminal 1995 paper, argumentation has become increasingly popular. In this talk I will provide a short survey of the field, covering both abstract and instantiated models of argument. The former considers arguments as atomic entities, and seeks to identify coherent sets of arguments, while the latter delves into how arguments can be constructed from a logical formalism. I will then discuss extensions to these models, as well as some open questions in the field.

Arguing2Text

Federico Cerutti
University of Aberdeen

One of the main goals in the Natural Language Generation (NLG) community is to build "articulate machines" which communicate with people in the same way that other people do. To achieve this vision, machines have to be provided with the capability not just to express data, but also to argue on it. This is becoming nowadays also an applicative question given IBM's interest in actively investing in the development of so-called the Debating Technologies http://researcher.ibm.com/researcher/view_group.php?id=5443. Among other approaches, argumentation theory has emerged as one of the mainstream research field in artificial intelligence for providing support to complex decision making activities in part due to the close alignment between its semantics and human intuition. In past research, we assessed this claim by means of an experiment. Within the experiment, participants read a paragraph in natural language --- handcrafted to be natural and fluent --- depicting an indirect dialogue between fictitious actors: each actor plays a role by defending a specific position. The original knowledge base is formalised using the Prakken & Sartor 1997 approach, which was developed to support legal reasoning. We thus hypothesized that there is a correspondence between statement acceptability of the natural language interface of the knowledge base (as judged by humans) and justification status (according to the formal model of Prakken & Sartor). In particular, we expected that the majority of the participants agrees with the skeptically accepted arguments, but not with any of the credulously accepted ones.

Our results suggest such a correspondence, however, post-hoc analyses show that there are some significant deviations. Two elements should be considered to explain it. First of all, evidence suggests that people considered implicit, "collateral," knowledge. However, in this context we will focus more on understanding whether we conveyed the desired message in an effective way. Indeed, we chose to present several arguments, including arguments in favour and against preferences, as part of a short story involving fictitious characters. There are various alternatives to this one as well as variations. For instance, we could have used direct speech, different ordering in the presentation, or even a scientific presentation of evidences and reasoning rules. However, a different question also arises, regarding whether, instead of describing the knowledge base, a more effective way would be to describe the result of the reasoning process, and only indirectly the original knowledge base. Finally, although partially borderline with the topic of this symposium, connections with other means of human-computer interaction --- such as graphical representation, automatic video generation, and a mixture of them ---will be briefly discussed.

Understanding Complex Systems, From Models, to Arguments, to Language

Nir Oren
The University of Aberdeen

Complex computational systems are built using technologies such as automated readers and planners, as well as decision and game theoretic components. While such systems can often act in an optimal (or near-optimal) manner, and achieve even better results than trained experts, they suffer from several important shortcomings. Of these, perhaps the most important is the lack of scrutability: since such systems must interact with humans — not only in the context of humans performing the actions specified by the system, but also with humans debugging the system and feeding information into it — human understanding of the reasons for the systems outputs is necessary.

The goal of the Scrutable Autonomous Systems (SAsSy)[1] project has been to make complex systems (exemplified by automated planners) scrutable. To this end, the project has utilised several technologies including Natural Language Generation (NLG), Argumentation and Diagrammatic Reasoning. We view the scrutable portions of the system as a mixture of facts and (potentially defeasible) rules. Such facts and rules can then be combined to form arguments, with the output of the system being an argument's conclusion [2]. In turn, these arguments can be encoded in natural language via NLG. Previous work [4] has shown that argumentation, and this approach, holds promise for explanation.

Standard argumentation theory would allow a user to identify all consistent, valid or justified arguments (which effectively represent the reasoning used by the system to compute its outputs)[5], and NLG would be used to provide an understandable explanation of these arguments. However, doing so raises the problem of information overload — the human is exposed to too many arguments simultaneously, and may therefore struggle to follow the explanation. Additionally, such a solution suffers from redundancy, exposing information the human is already aware of.

An alternative approach involves the dialogical presentation of arguments, via so-called proof dialogues [3]. Here, the system and human engage in dialogue with each other, in order to incrementally explain the justification status of an argument. While previous work has examined the use of such dialogues to efficiently determine whether an argument exists within the grounded extension, or within a preferred or stable extension, such work suffers from two notable shortcomings that we wish to address. First, argument (rather than meta-argument [6]) based proof dialogues ignore the skeptical preferred semantics, identifying whether an argument exists within a single extension, rather than whether it exists in all extensions. The preferred skeptical semantics are of potential importance in areas such as practical reasoning, and should therefore be considered. Second, these dialogues attempt to determine whether an argument appears within the extension, but do not differentiate between arguments which do not appear due to being explicitly unjustified, or which have an undetermined status. Differentiating between these two classes of excluded arguments is important in explanation, as a user may wish to know why an argument was excluded from consideration (e.g., by asking a question of the form "why was this plan not adopted").

Having argued for the importance of scrutability, and outlined the knowledge to argument to NLG pipeline, we will present existing and ongoing work on argumentation proof dialogues.

References

Martin Caminada, Roman Kutlak, Nir Oren, and Wamberto Weber Vasconcelos. 2014. Scrutable plan enactment via argumentation and natural language generation. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 1625-1626.

Martin Caminada, Sanjay Modgil and Nir Oren. 2014. Preferences and Unrestricted Rebut. In *Proceedings of the Fifth International Conference on Computational Models of Argument*.

Martin Caminada and Sanjay Modgil. 2009. Proof Theories and Algorithms for Abstract Argumentation Frameworks. In *Argumentation in Artificial Intelligence*. Editted by Iyad Rahwan and Guillermo R. Simari. Springer. 105–129.

Federico Cerutti, Nava Tintarev and Nir Oren. 2014. Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In *Proceedings of the 2014 European Conference on Artificial Intelligence*. 207—212.

Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*. 77 (2): 321-357.

S. Doutre and J. Mengin. 2004. On Sceptical vs Credulous Acceptance for Abstract Argument Systems. In *Ninth European Conference on Logics in Artificial Intelligence*. 462–473.

Kees van Deemter, University of Aberdeen, Scotland, UK

Algorithmic models of language production: generating complex referring expressions

Psycho-linguists have a long tradition of constructing models of language production (e.g. Levelt 1989). In recent years, researchers have started to grow dissatisfied with these models, because

- they are not detailed enough to make concrete predictions as to what utterance might be generated in a given situation,
- they do not take differences between speakers into account, and
- they do not have great explanatory value, for example because they do not assess the effectiveness of different possible utterances.

In this talk, which focusses specifically on the production of *referring expressions*, I will discuss a new line of work that seeks to eliminate these weaknesses. Researchers in this line of work construct computational algorithms that convert a stimulus (e.g., a scene observed, and an intended referent within the scene) into an utterance (e.g., a referring Noun Phrase) or a probability distribution over utterances. Much attention is devoted to empirical testing of the algorithms constructed. In most cases, testing proceeds by comparing the output of an algorithm to the utterances produced by human speakers (under strictly controlled circumstances), and measuring the degree to which the two classes of utterances resemble each other.

The second part of my talk is likely to zoom in on a less-often studied problem area within this domain, which links the computational modelling of language production with modern Knowledge Representation (e.g., Description Logic).

Concretely, we ask how the above research program can address the generation of referring expressions whose logical structure is not a simple conjunction of atoms, as in "the man who loves *all* cars", "the program that uses lexicons of two different languages" etc.. If time allows, we discuss how the proposed approach can also model the production of *attributive* descriptions, as in Donnellan's example of "the murderer of Smith", uttered in a situation in which the identity of the murderer is not known (so it means "whoever murdered Smith"). For example, the production algorithm could start from a Knowledge Base that says:

knife-in-chest(Smith)
(Smith has a knife in his chest)
knife-in-chest $\subseteq$ 1.HasMurderer.Person
(Having a knife in one's chest implies having one murderer)

deducing that

Cardinality (Person $\cap$ $\exists$.Inverse(HasMurderer).{Smith}) = 1
(The set of persons who murder Smith has a cardinality of 1)

Given this deduction, it is not difficult to let the program produce a Noun Phrase like “the murderer of Smith”.

We discuss the pro’s and con’s of extending algorithmic models of reference production in this way.

References

Kees van Deemter (in press) Computational Models of Reference. To appear with MIT Press, April 2016.

Kees van Deemter, Albert Gatt, Roger P. van Gompel and Emiel Krahmer (2012) Towards a Computational Psycholinguistics of Reference Production. *Topics in Cognitive Science* **4** (2).

Emiel Krahmer and Kees van Deemter (2012) Computational Generation of Referring Expressions: a Survey. *Computational Linguistics* **38** (1).

Levelt (1989). Speaking: from Intention to Articulation. MIT Press.

Yuan Ren, Kees van Deemter, and Jeff Pan (2010) Charting the potential of Description Logic for the Generation of Referring Expressions. Proceedings of the 6th International Conference on Natural Language Generation.

A Brief Study on Surface Realization for Chinese NLG

Dr. Xiwu Han, 28th Aug 2015, University of Aberdeen

Natural Language Generation (NLG) is the natural language processing task of generating natural language from a machine readable representation such as a knowledge base or a logical form. Especially for a pipeline system, there are generally four stages for NLG, i.e. Data Analysis, Content Selection, Document Structuring, and Surface Realization. Surface Realization is the most fundamental NLG stage of creating the actual text, which should be correct according to the rules of syntax, morphology, and orthography. SimpleNLG, originally developed at the University of Aberdeen's Department of Computing Science, is intended to function as a "realization engine" for Natural Language Generation architectures, and has been used successfully in a number of projects, both academic and commercial. We made a brief study about some important points of Chinese morphology and syntax in potential vision of a SimpleNLG-like Chinese NLG system. On the basis of these points, we tried to put forward some concrete rules or graphs for the surface realization of Chinese NLG.

I. Briefly about SimpleNLG

SimpleNLG (https://github.com/simplenlg/simplenlg) is a simple Java API designed to facilitate the generation of Natural Language. It handles the following issues:

1. Lexicon/morphology system: The default lexicon computes inflected forms (morphological realization). We believe this has fair coverage. Better coverage can be obtained by using the NIH Specialist Lexicon (which is supported by simpleNLG).
2. Realizer: Generates texts from a syntactic form. Grammatical coverage is limited compared to tools such as KPML and FUF/SURGE, but we believe it is adequate for many NLG tasks.
3. Microplanning: Currently just simple aggregation, hopefully will grow over time.

We hereby give a simple example for SimpleNLG.

Input:

leave(boy, house)

Java Coding:

```
Phrase s1 = new SPhraseSpec('leave');
s1.setTense(PAST);
s1.setObject(new NPPhraseSpec('the', 'house'));
Phrase s2 = new StringPhraseSpec('the boys');
s1.setSubject(s2);
```

Output:

The boys left the house.

II. Some Morphological and Syntactic Points

Chinese is not so inflective as English such that a clear separation between morphology and syntax seems difficult to draw. We therefore grouped these important points in an easier way for references and relevance in further programming, while roughly considering morphological and syntactic classification. Our morphological study involved Chinese verbs, nouns, adjectives, adverbs, numbers, quantifiers, etc. Our syntax study covered Chinese constituent order, question format, negation format, usage of tendency verbs, usage of pronoun and connection words, usage of prepositions and postpositional words, usage of punctuations, idioms, and auxiliary nouns.

III. Rules for Surface Realization

We considered phrase structures, sentence structures, and relevant pragmatic meanings beneath these structures. Some basic principles we complied with include:

1. Try to keep a balance between over-generation and under-generation;
2. Try to keep a balance between syntactic conciseness and semantic distinguishability;
3. Try to keep enough details for possible meaning subtleness;
4. Hierarchically clustered according to constituents from left to right;
5. Try to minimize overlap.

IV. Some Existing Resources

There exist some lexical, syntactic, and pragmatic resources, which may well be helpful for developing Chinese NLG systems.

1. HowNet Knowledge Database (http://www.keenage.com/);
2. HIT-CIR Chinese Dependency Treebank (http://ir.hit.edu.cn/demo/ltp/Sharing_Plan.htm);
3. Tsinghua Chinese Treebank (http://cslt.riit.tsinghua.edu.cn/~qzhou/eng/Resources.htm);
4. Penn Chinese Treebank (www.cis.upenn.edu/~chinese/);
5. Other Chinese annotated or raw corpus.

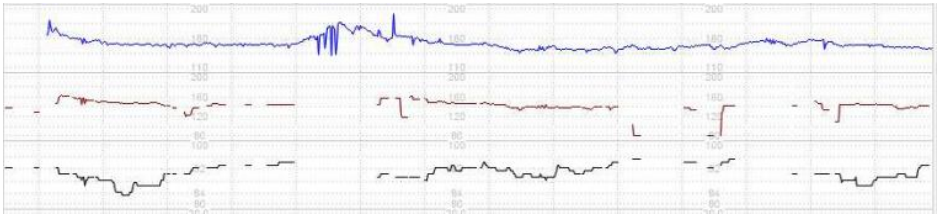
HUMAN-LIKE NATURAL LANGUAGE PRODUCTION

Epsilon Lee, epsilonlee.green@mails.ccnu.edu.cn
Machine Intelligence and Translation Lab, Harbin Institute of Technology.

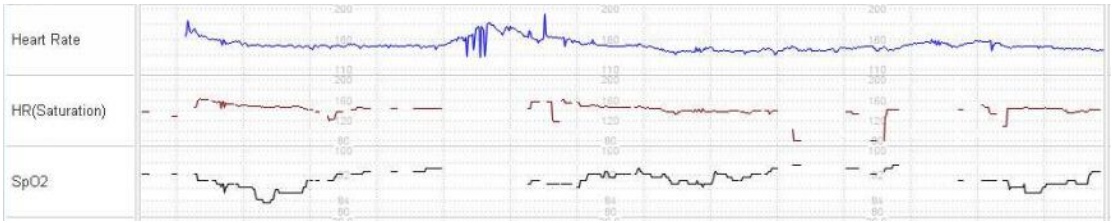
1. On natural language generation

Nowadays, researches on NLG embrace a data-driven approach, which means that NLG as a language competence is ignored within CL(computational Linguistics) community. We always refer to this approach as DTT(Data-to-Text). In this beginning section, I would like to propose my enthusiastic love on the cognitive aspect of NLG and argue that there is a paradox emerging between these two approaches.

Data, which is unreadable and everywhere anytime in daily life as well as in every aspect of various academic disciplines, means a lot to us, especially for those scientists in statistics and machine learning. However, data doesn't really make sense without semantic annotations, even for experts in respective areas (Refer to Figure 1 for explanation) , let alone ordinary people. But with a language coat, data does mean something even to lay people. And that is why the trend nowadays will continue as far as I could see.



(a) Without semantic annotations



(b) With semantic annotations

Figure 1: When you see the figure (a) alone, you could not figure out what it is really talking about, so it is non-sense to human without the semantics in (b).[1]

As a result, NLG Systems at present almost all have a clear and computational specification of the input data. And in turn, those data lying in specific domain corresponds to specific application requirements. On the one hand, it is a golden law in computer science that a computer systems(software-based) must have a input, output specification and a software with none of these is weird and of no use at all. On the other, if one is going to research along cognitive approach, which aims at making computers simulate the mechanism of language production, where and how does he get the input data. It is hard to say that what the representation of the data input to a human brain for language production is and it's a research issue among neurosicentists. And that's the paradox.

In terms of a set-theoretic point of view, the linguistic mechanism of language production has

a bunch of extensions, namely, the ability to describe objective things perceived by eyes, ears, etc. from which is mostly not in memory but from real-time environment, the ability to argument upon a specific topic or a set of topics from which the data is stored in memory. Whatever kind of extensions, data is not as the same as in DTT mode. As a matter of fact, the explorations within cognition-based Natural Language Production(NLP) should be of somewhat inter-cognition. That means to test or evaluate a model of human NLP, some other kind of cognitive mechanisms should be on the line as well.

2. A qualitative model of NLP

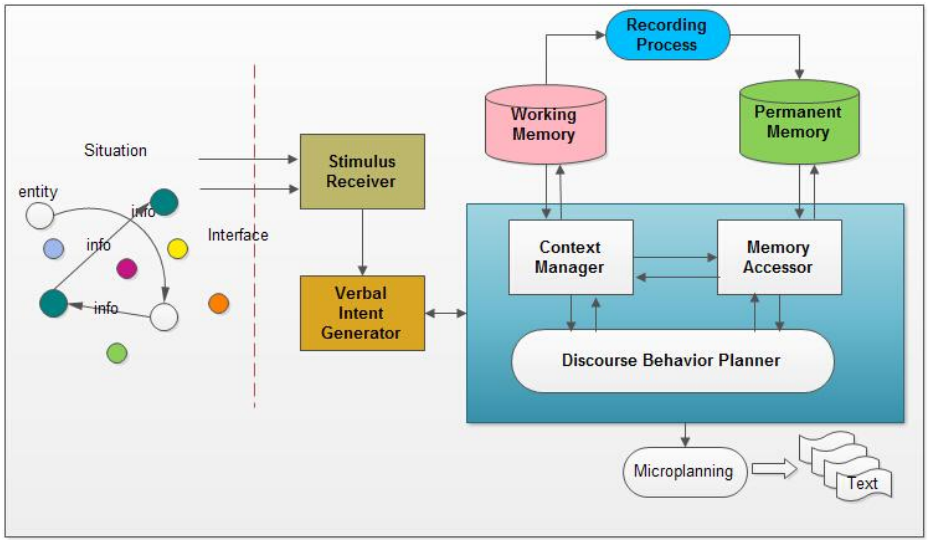


Figure 2: A qualitative model of Natural Language Production

Although it left out many detailed considerations, Figure 2 illustrates a model of the whole process of NLP. What I want to emphasize is that it is a model of the overall mechanism of NLP from raw data, the situation outside the model, to the Stimulus Receiver which filters the noises, captures the point of interest for the 'brain' and passes it through to the verbal intent generator. Then the generator take all the information in the stimulus and context into consideration to produce a set of verbal intents that are eventually realized as a response in grammatically correct strings.

Because of the word limit(although I have exceeded a lot), I won't discuss here about the model in great detail, it is from my Bachelor's dissertation in June this year and has a lot of flaws seen at present. In the last section, I would like to propose three research ideas which may seek to make us a good understanding of human-like NLP.

3. Some Research Issues

In the second section, I talked about an model as a simple approximation of human-like NLP, which has many points worth digging into.

3.1Putting clauses together

To really make a computer produce natural language like human, knowing the pure syntax of a language is not enough, though nowadays no computers really grasp a simplest grammar like how human does. More than that, using language is to map from the meaning one wants

to convey to the carrier of that meaning, a surface structure. In cognitive linguistics, there is a grammar named construction grammar, which explores the human cognitive manipulation of the linguistic conventional <meaning, from> pairs in language. As their working assumption, a certain meaning is accompanied with a construction which is a surface form of a clause /sentence without specific value of subject, object, etc. Assume that there have a distribution of all those constructions in use, if we've got the semantic representation we could use Bayesian theory to get the surface structure. The understanding of the assumed distribution could be a social and linguistic convention. For instance, if one is going to introduce a person to anther person, there are conventional way of start the talk.

3.2'Orator' on general topics

To build a system that can argument on general controversial topics, such as 'what do you think of peace?', 'What is your feeling about DINK family?'. Initially, I conceived the process of doing these as a basic issue named computational coherence which has been greatly promoted by RST and the debates in 1990s but still remains unsolved. From a rationalistic point of view, my original assumption is that there exists a discourse behavior logic that interlocutors conforms to. You can see in Figure 2, there is a discourse behavior planner and that's where the logic takes place. The logic could be based on dynamic epistemic logic or maybe in the future on glue logic[2]. And it is used to select content in a knowledge base to construct discourse structure according to manually designed operators. The future work would be how to automatically learn the operators using machine learning techniques according to specific genre and style of the topic.

3.3Usage-based language acquisition

If it is possible to teach a computer piece-by-piece of a language like children do through their preschool period, the computer should recognize speech, do some basic language understanding or even recognize visual situations. The core question concerns what is the minimal innate capability we should program on it, according to Chomsky Minimalist Programme and Universal Grammar. How can it embody the ability it learned?

References:

[1] A Gatt, F Portet, E Reiter, J Hunter, S Mahamood, W Moncur, S Sripada (2009). From Data to Text in the Neonatal Intensive Care Unit: Using NLG Technology for Decision Support and Information Management AI Communications 22:153-186.

[2] N Asher, A Lascarides (2003). Logics of Conversation, Cambridge University Press.

Build Up An Automated Chinese Essay Evaluation System For Ethnic Minority Writers

LI Lin[1], ZHAO Weina[1]

(1. The Computer College of Qinghai Normal University, Xining, Qinghai 810004)

Increasingly, ethnic minorities have left their hometown for better education or job opportunities. Their experiences have shown that poor Chinese writing skill is a major obstacle for them to integrate into the mainstream society. Hence we determine to develop a Chinese essay evaluation system (CEES) to assist them to improve writing skill.

The main goal of CEES includes helping the minorities understanding their Chinese writing level, pointing out strength and weakness, and encouraging them to practice more. CEES is designed to consist of two modules: automatic essay scoring (AES) and feedback report generating. To achieve our goal, we attempt to apply both NLP and NLG methods in CEES.

In the rest of this abstract, we discuss potential problems and solutions for developing CEES.

1. Automated Essay Scoring (AES)

   AES is defined as the computer technology that grades and scores essay, which is usually employed as a complement of manual scoring. Obviously, AES is more economical and effective than manual scoring.

   Early AES system grades an essay according to text features. Researchers have adopted both grammar and semantic features in ASE system such as the longevity of sentences, the variety of words, or the number of punctuations.

   In recent years, some statistical NLP technique and methods have been widely used in AES such as Naïve Bayesian algorithm, document classification and clustering technique. Statistical method could evaluate both the content and language quality of an essay, which is helpful to improve the scoring precision.

   We plan to build an integrated AES system based on the above two methods, therefore our research plan is as follow:

   (1) To start off with building a Chinese essay corpus with tagging including POS and manual score.

   (2) To build up a scoring model based on statistical methods. In this step, we will focus on three sub-tasks: to find a statistical method suitable to AES, to acquire effective features from corpus, to train and obtain a scoring model.

   (3) For improving the precision of our model, we consider to integrate text features into our model.

2. Automated Feedback Report Generating

We plan to build up a corpus that contains both Chinese essay and its feedback reports written by experts. We estimate that it is possible to collect enough essays for the corpus. But it could be very difficult to obtain the reports, so we decide to ask experts to write reports for every essay.

The report generating could be divided into the following steps:

(1) To acquire knowledge about content and structure of a feedback report from corpus and domain experts.

(2) Content determination

The content of a feedback report is determined by score and writing quality. For instance, whether an essay presents all information required should be presented in the report. Therefore, we will set up a rule bank for report generation according to not only AES result but also the quality of an essay.

(3) Micro-planning and realization by rule-based approach

Mapping numerical information to words: towards a fully statistical approach

Xiao Li[1], Chenghua Lin[2], Kees van Deemter[3]
[1]xiao.li@abdn.ac.uk; [2]chenghua.lin@abdn.ac.uk; [3]k.vdeemter@abdn.ac.uk

This work sketches the outlines of a new approach to the generation of textual summaries of numerical information (e.g., in weather forecasts).

Researchers in Natural Language Generation (NLG) often try to convert numerical data (e.g., projected weather data) into text (e.g., a written weather forecast). For example, time phrases such as 'midday', 'by afternoon', etc. are often used to refer to times, so the system has to choose which of these words to use in a specific case (e.g., Reiter et al. 1997, 2005). These words are generally selected from a corpus of textual summaries written by domain experts. Some systems model the use of these words by crisp thresholds, (e.g. 'midday' could be used to denote any time from 10:00 – 14:00). Other systems select words on the basis of frequencies in a data-text corpus (which couples texts with the data they describe), but the candidate words themselves are still given by experts. This limits the generality of the approach and makes it difficult to scale it up (e.g., by applying it to all the words in a long text).

Therefore, we aim to construct an algorithm which avoids any expert-based rules to give an NLG system the ability to automatically detect when a given word (or ultimately, each phrase) is used. The input of this algorithm is some numerical data in the data-text corpus (e.g., the temperature at various times on a given day); the output of the algorithm is a probability vector; each entry in the vector gives us, for a given word w (e.g., the word "hot"), the probability that w occurs in a summary in the corpus. We have started to experiment with this method using the Sumtime-Meteo corpus (REF).

This algorithm is different from Liang (2009), which simultaneously segments the text into utterances and maps each utterance to a numeric data field. In Liang's method, one piece of data can only be mapped to one continuous section of text (as the utterance). For this reason, a word always corresponds to data field. If a word (such as 'muggy') corresponds to more than one data field ('muggy' might correspond to both temperature and humidity), this is not taken into account. Our approach does not suffer from this limitation: it can relate the probability of a word's occurrence to more than one data field. Conversely, one data field can be linked to several words.

Once the algorithm is finished, some metrics will be applied to it to determine whether the occurrence of each word lies on the given data or not. On one hand, some words (content words) should lie on the given data, for example, 'midday' illustrates 10:00 – 14:00. On the other hand, other words (structure words) such as 'the', 'it', even 'today' in weather forecast are only used to complete sentence. The occurrence of these type of words will not change under the different give data. By analysing the metrics corresponding to these words, we can somehow evaluate this algorithm. If we

find contradictions that content words are indicated as structure words, we know there are errors involved in somewhere.

Reference

Liang, P., Jordan, M. I., & Klein, D. (2009, August). Learning semantic correspondences with less supervision. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1* (pp. 91-99). Association for Computational Linguistics.

Reiter, E., & Dale, R. (1997). Building applied natural language generation. *Journal of Natural Language Engineering*, *3*(1), 57-87.

Reiter, E., Sripada, S., Hunter, J., Yu, J., & Davy, I. (2005). Choosing words in computer-generated weather forecasts. *Artificial Intelligence*, *167*(1), 137-169.

Playing Games with Dynamic Epistemic Logic

Thomas Ågotnes
University of Bergen

Dynamic epistemic logic describes the possible information-changing actions available to individual agents, and their knowledge pre- and post-conditions. For example, public announcement logic describes actions in the form of public, truthful announcements. However, little research so far has considered describing and analysing rational choice between such actions, i.e., predicting what rational self-interested agents actually will or should do. Since the outcome of information exchange ultimately depends on the actions chosen by all the agents in the system, and assuming that agents have preferences over such outcomes, i.e., over multi-agent epistemic states, this is a game theoretic scenario. This is an interesting general research direction, combining logic and game theory in the study of rational information exchange. In the talk I will focus on one particular setting: the case where available actions are public announcements, and where each agent has a (typically epistemic) goal formula that she would like to become true. What will each agent announce? The truth of the goal formula also depends on the announcements made by other agents, thus we have a game-theoretic scenario. I discuss how such *public announcement games* can be analysed. I will also briefly discuss two other settings. First, consider coalition formation: if agents are allowed to form coalitions, which coalitions will form, i.e., which are coalitions are stable? We can answer such questions by studying the *coalitional* public announcement games inherent in Kripke models. Second, consider the setting where instead of choosing an announcement each player chooses a question the other player is obliged to truthfully answer. What are the best questions to ask? Again, this question can be discussed by analysing the resulting *question-answer games*. The talk is based on joint work with Hans van Ditmarsch, and parts also with Johan van Benthem and Stefan Minica.

Play a Game with a Metaphor

——A game-theoretic account of using metaphor

Cihua Xu
Zhejiang University

Although Relevance Theory explicates the inference process and the conditions constraining metaphor comprehension, its analysis is still essentially descriptive, or at most partially formalized. The use of concepts such as “mutual manifestness”, “non-demonstrative inference”, “relevance”, “cognitive effect”, “cognitive effort” and “cognitive contexts” poses a great challenge for the formalization of its analysis. The IBR model of Game-theoretic Pragmatics emerges as an effective model which can meet the challenge. Focusing on analyzing communicative contexts, the model covers the shared information, signal strategy, rational selection, utilities and probabilistic belief. Its solution concept takes an internal perspective in order to show how communicators achieve the equilibrium (the correct understanding of an expression). Therefore, the IBR model can provide a holistic method for formalizing the inference process and its constraint conditions explicated by RT. Within the IBR model, this article analyzes the process of interaction among different elements in metaphor usage, in an attempt to demonstrate how the process of using metaphors can be effectively formalized.

A Game-Theoretic Analysis on the Use of Indirect Speech Acts

Mengyuan Zhao
School of Social Sciences, University of Shanghai for Science and Technology

In our daily communication, instead of explicitly expressing our intentions, we often do so in an indirect manner. According to the speech act theory, which was introduced by Austin (1962) and developed by Searle (1969), this kind of pragmatic phenomenon is called *indirect speech act* (ISA).

Brown and Levinson (1987) suggest a reason for the use of ISA: it is a strategy in politeness. In their Politeness Theory, people would like to adopt some strategies to save each other's face when their communication involves face-threatening acts, such as criticism, insults, disagreement, suggestions, refusal, requests etc. For example, people would say, *Could you pass the salt?* rather than, *Pass me the salt*. Conventional ISA is often used to show politeness.

However, Pinker, Nowak and Lee (2008) point out that the Politeness Theory is not comprehensive enough to account for the use of ISA, for the theory presupposes pure cooperation in human communication, which is not always the case during instances of ISA. They list several cases involving a mixture of conflict and cooperation, such as sexual comes-ons, veiled threats and concealed bribes. According to Pinker *et al.*, the use of ISA in these circumstances is due to the fact that it allows for speaker's plausible deniability facing an uncooperative hearer. They also introduce a game-theoretic model for such cases to support their claims that ISA strategies guarantee the speaker a better payoff. It is noted that such cases involve non-conventional ISA, which is deniable.

On the other hand, non-conventional ISA, especially indirect requests, is considered to be risky speech. Sally (2003) studies a variety of linguistic experiments and argues that risky speech is more commonly used among people who are more sympathetic towards each other. His theory is based in a game-theoretic analysis: the strategy profile involving risky speech corresponds to a payoff dominant but risk dominated equilibrium, which may turn into both payoff and risk dominant equilibrium when the interlocutors are close enough. Van Rooij and Sevenster (2006) introduce *Super Conventional* signaling games to model Sally's work on risky speech. Unlike Pinker *et al.*, Sally's cases involve communication situations under certain cooperation.

In this paper we construct a uniform model to analyze the use of ISA. We argue that both conventional and non-conventional use of ISA can be explained in terms of a game theoretical logic, specifically through iterated best response (IBR) reasoning. The conventional use of ISA corresponds to the Lewis theory of convention (1969), which introduces signaling games to explain how meaning is assigned to language through its use. Non-conventional ISA involves two types of communication

situations: communication under certain cooperation, such as ironical requests and that under uncertain cooperation, such as bribes. To solve the signaling games of the above situations, we build a stronger version of IBR reasoning framework by introducing the concepts of *higher-order beliefs* and *strategy filters*, realizing the qualification of sympathy and deniability in our model. This reasoning results in the following predictions: the use of non-conventional ISA under certain cooperation relies on the sympathy between interlocutors, which blocks its evolution towards conventional ISA; in uncertain cooperation situations, people are more likely to use ISA, which helps its conventionalization.

Keywords: indirect speech acts; Game Theory; iterated best response reasoning; convention