

The Co-Constructed Logic Framework for Understanding Children's Acts and Their Intentions

Hongya Fan

School of Foreign Languages, Shanxi University, Shanxi, China

Zeshan Ren

School of Foreign Languages, Shanxi University, Shanxi, China

Abstract—With the characteristics of the nonmonotonic logic and defeasible inference, abductive reasoning has been formalized in the field of artificial intelligence, dealing with the local pragmatics (e.g., the resolution of coreference, resolving syntactic and lexical ambiguity and interpreting metonymy and metaphor), recognizing discourse structure and even the speaker's plan and other issues for natural language understanding. However, Hobbs' analysis of abduction in recognizing the speaker's plan was conducted only from the point of view of the verbal information processing that the listener does. To demonstrate the collaborative way that conversational partners working together to understand the logic of human acts and their intentions, this article analyzes the two conversations about the parents questioning their children's intention for their acts with an abductive reasoning method. The results show that children and parents co-construct segments of discourse with coherence relations across several conversational turns, by that way they build together a simplified framework for understanding the logic of human acts and their intention. For example, when the father and his children co-constructed coherent segments of discourse with the result relation between them, they completed the particular intention understanding at the same time. This research helps in enriching the logic structure of artificial intelligence applications such as visual question answering models and enhancing their reasoning abilities.

Index Terms—abduction, natural language understanding, coherence relation, visual question answering

I. INTRODUCTION

In investigating the way how we use our knowledge of the world to understand discourse, a proof procedure based on the inference rule of abduction may help us. Treating abduction as one variety of nonmonotonic logic, Hobbs defined it as “from an observable Q and a general principle $P \supset Q$, we conclude that P must be the underlying reason that Q is true. We assume P because it explains Q ” (Hobbs, 2008, p.727). Different from the logic of mathematics which is monotonic and the truth value of a statement is so stable that other knowledge we learn later cannot change it, commonsense knowledge other than mathematics is uncertain or defeasible. “Whatever general principles we have are usually only true most of the time or true with high probability or true unless we discover evidence to the contrary” (Hobbs, 2008, p.726). When we gain more information, we may have to change what we believed to be the truth value of statement. Turning nonmonotonic logic to its advantage, artificial intelligence (AI) is able to work with simplified framework and develop with its database expanding.

Hobbs has described the Interpretation as Abduction (IA) framework as how to interpret a sentence step by step:

“Prove the logical form of the sentence,
together with the selectional constraints that predicates impose on
their arguments,
allowing for coercions,

Merging redundancies where possible,

Making assumptions where necessary.

By the first line we mean ‘prove, or derive in the logical sense, from the predicate calculus axioms in the knowledge base, the logical form that has been produced by syntactic analysis and semantic translation of the sentence’” (Hobbs, 2008, p.732).

All the participants in a discourse situation have their own sets of private opinions, and there is a large overlapping opinion area which is extended to include some private opinions of the speaker's. It is anchored referentially in mutual belief, and “when we succeed in proving the logical form and the constraints, we are recognizing this referential anchor” (Hobbs, 2008, p.733). The overlapping opinions that participants have are the given information which is definite and presupposed. As the new information which is indefinite and asserted, the speaker's private beliefs could be used by participants to make assumptions. As regards merging redundancies, it is “a way of getting a minimal, and

hence a best, interpretation. Merging redundancies and minimizing the assumptions result naturally from the method of weighted abduction” (Hobbs, 2008, p.733).

With the caveat of reasoning must be defeasible and the IA framework, Hobbs has explained how abduction provides a framework for addressing a broad range of natural language understanding problems, including reference resolution, recognizing the syntactic structures of sentences, and recognizing discourse structures and the speaker’s plans.

In fact, we found that when we use abductive reasoning to analysis discourse structures, we also realized the process of understanding human behavioral logic. Participants in the conversation collaborate for evidence that they have mutually understood what the speaker means (Wilkes-Gibbs, 1986). Besides, they have known where the conversation is going. In naturally performing communicative acts, people do spontaneous completions, for example, one person helps to finish another's suspended act. In this paper, we focus on inferring the ways how parents work with their young children to understand their behavioral logics and their intention of doing them. It is worth noting that in the two-way process of behavioral logic understanding, children are also learning rules of adults from the interactions with their parents.

II. THE OCCASION RELATION BETWEEN CHILDREN’S ACTS AND THEIR CONSEQUENCES

While connecting different ideas together, coherence relations play a role in discourse understanding. “Discourses are not simply arbitrary collections of utterances. A felicitous discourse must instead meet a rather strong criterion, that of being coherent” (Kehler, 2008, p.241). If parents recognize the coherence relation in understanding discourse about their children’s acts, they can make out the logic of those acts and the children’s intention of doing them.

Several strands of researchers in the language sciences have turned their attention to discourse coherence. Kehler summarized a sample of them from the perspectives of theoretical linguistics, computational linguistics and psycholinguistics respectively. Motivated by the need of interpreting and producing discourse by computational models, computational linguists have tried to characterize the set of coherence relations which could be used to connect clauses. The definitions for a set of relations that Hobbs provided in 1979 and 1990 to interpret discourse have been rooted in the operations of a computational inference system. In 1993 Hobbs claimed in subsequent work a proof procedure that can be used to identify coherence in texts based on abductive reasoning.

In regard to the discourse production, researchers have focused on how to make computers generate coherent text automatically. As a result, they developed the Rhetorical Structure Theory (RST) to build “natural language generation systems, since its relation definitions can be cast as operators in a text planning system that associates speaker intentions with the manner in which they can be achieved” (Kehler, 2008, p.245). In particular, a complex communicative goal can be achieved through an RST relation by splitting it into central and peripheral subgoals iteratively until the level is reached at which these constraints can be met by generating single clauses.

The previous research on coherence relations has been mainly carried out within the boundaries of the field of text coherence itself and coherence relation theories have not been used into accounts of particular linguistic forms. In his book, *Inquiry Concerning Human Understanding*, David Hume found its curiosity that there was a lack of attempting to enumerate or class all the principles of association, and he summarized “three principles of connection among ideas, namely Resemblance, Contiguity, in time of place, and Cause or Effect” (Kehler, 2008, p.246). Kehler argued that coherence relations can be seen as the canonical instances of three general classes of “connection among ideas,” and he presented categorization of a set of coherence relations and then summarized linguistic analyses based on the categorization. He showed that to solve distinct linguistic problems, in fact, an analysis of coherence is necessary.

These categories differ systematically in two respects: in the type of arguments over which the coherence relations constraints are applied, and in the central type of inference process underlying this application. Although Kehler’s categorizations have some details different from the classification of Sander et al, they found the common ground on that the relations are composites of more primitive, cognitively inspired features. The classes of relations also show “considerable overlap with the three categories that were common to the classifications of Halliday and Hasan, Longacre, and Martin” (Kehler, 2008, p.247).

In the neo-Humean classification of coherence relations (Result, Parallel, and Occasion), Kehler quoted Hobbs’ definition of the first kind of Occasion as follow:

“**Occasion(i)**: Infer a change of state for a system of entities from S_1 , inferring the final state for this system from S_2 .” (Kehler, 2008, p.250)

Now let us consider an example. When the parents (referred to by “Mom” and “Dad” here) were helping their daughter RyRy with her toilet training, their son Tydus (TyTy) found two dolls in the toilet. Then they had a talk as follow:

- 01 Dad: Tydus did you put those in there?
 02 Tydus: No:: ?
 03 Dad: RyRy did you put those in there?
 04 RyRy: No↓ (0.2) No ↓ (0.2) TyTy did
 05 (1.2)
 06 Dad: TyTy? (hhenhh)

- 07 Tydus: No=I didn't
 08 (1.2)
 09 Dad: Did you
 10 Mom: No::↓
 11 Dad: RyRy did you
 12 RyRy: No=you did
 13 Dad: °I'm° Pretty sure I didn't
 14 (1.5)Did you?
 15 Tydus: No::↓
 16 Dad: RyRy=did you put those in there?
 17 RyRy: Ye:::ah
 18 Dad: You did? (0.8) Why↓
 19 RyRy: (0.5)(the-the re=poi)
 20 Dad: What?
 21 Tydus: They are going potty↓
 22 Dad: They are going potty↑
 23 RyRy: =Yeah
 24 Dad: Ahh=hoho::
 25 Mom: Oh::
 26 Dad: How do you get mad at that
 27 Mom: It's true↑
 28 Dad: Oh baby↓ (.) that's really cute(0.5) but honey we don't put toys in the potty=Ok
 29 RyRy: Ok↓

After the parents affirmed that it was RyRy who had put the dolls in the toilet, the father used the utterance “Why (did you put those in there)” to ask her the reason why her did that. As RyRy's expression was not clear enough, the father took her brother's utterance “They are going potty” as the translation of her one. With this information, the father understood the logic of RyRy's act and her intention to providing the condition for the dolls' going potty. Then we use abductive reasoning to infer how the father achieved it.

(1) Predicate, Eventuality, and Segment

At the sentence level, the structure and meaning of strings are inseparable from syntactic rules. Hobbs proposed a simplified version of the predicate *Syn*:

“*Syn* (*w*, *e*, *f*, *x*, *a*, *y*, *b*)

w is a string of words. *e* is the eventuality described by this string. *f* is the category of the head of the phrase *w*. If the string *w* contains the logical subject of the head, then the arguments *x* and *a* are the empty symbol ‘-’. Otherwise, *x* is a variable referring to the logical subject and *a* is its category. Similarly, *y* is either the empty symbol or a variable referring to the logical object and *b* is either the empty symbol or the category of the logical object” (Hobbs, 2008, p.737).

Hobbs noted “*p* is true of *x*” as *p*(*x*), while “*e* is the eventuality that *p*'s being of *x*” as *p*'(*e*, *x*). At the same time, Hobbs pointed out that “(∀ *w*, *e*) *Syn* (*w*, *e*, *v*, -, -, -, -) ⊃ *Segment* (*w*, *e*). That is, a grammatical sentence conveying *e* is a coherent segment of discourse conveying *e*” (Hobbs, 2008, p.740).

Then with the knowledge of deixis we can get two different segments from the two utterances “Why (did you put those in there)” and “(the-the re=poi) (They are going potty)”:

Segment (“Why did RyRy put the dolls in the potty?”, *e*₁)

Segment (“The dolls are going potty.”, *e*₂)

And *put*'(*e*₁, *r*, *d*) means that there is a putting event *e*₁ by RyRy *r* of the dolls *d* in the potty, and *go*'(*e*₂, *d*, *p*) means that there is a going potty event *e*₂ of the dolls *d*.

“When two segments of discourse are adjacent, that very adjacency conveys information... Part of what it is to understand a discourse is to discover what that relation is.” (Hobbs, 2008, p.739)

Hobbs use *CoherenceRel* (*e*₁, *e*₂, *e*) to express that there is a coherence relation between eventuality *e*₁ and *e*₂, and the combination of the two conveys *e*, and then he proposed the following rule:

“(∀ *w*₁, *w*₂, *e*₁, *e*₂, *e*) *Segment*(*w*₁, *e*₁) ∧ *Segment*(*w*₂, *e*₂) ∧ *CoherenceRel*(*e*₁, *e*₂, *e*) ⊃ *Segment*(*w*₁*w*₂, *e*)” (Hobbs, 2008, p.740)

With Hobbs' definition of Occasion(i), we can express it as:

(∀ *e*₁, *e*₂, *x*) *change*' (*e*₁, *x*) ∧ *state*' (*e*₂, *x*) ⊃ *occasion*(*e*₁, *e*₂)

And also:

(∀ *e*₁, *e*₂) *occasion* (*e*₁, *e*₂) ⊃ *CoherenceRel* (*e*₁, *e*₂, *e*)

The coherence relation is of important because it can combine two segments into one, and through this way the whole discourse is constructed into a tree-like structure. Thus, to interpret the discourse is to prove abductively the expression:

(∃ *e*) *Segment* (“Why did RyRy put the dolls in the potty? The dolls are going potty.”, *e*).

After back-chaining above, now we need to prove the expression:

$(\exists e_1, r, d, p, e_2) RyRy(r) \wedge put'(e_1, r, d) \wedge dolls(d) \wedge in(d, p) \wedge potty(p) \wedge occasion(e_1, e_2) \wedge go'(e_2, d, p)$

That is, there is a putting event e_1 by RyRy r of the dolls d in the potty p . There is a going potty event e_2 by the dolls d . And putting event e_1 is an occasion for the going potty event e_2 . In this expression, there are only simplified logical forms of the original two utterances, and the hypothesized occasion relation between them. The tenses and some other complexities are ignored.

To prove the logic form we need to use predicate calculus axioms in the knowledge base. Suppose we have in the knowledge base the following axioms:

$put'(e_1, r, d) \supset move'(e_1, r, d)$

That is, if e_1 is a putting event by r of d , then that e_1 is also a moving event by r of d ;

$move'(e_1, r, d) \supset change'(e_1, d)$

That is, if e_1 is a moving event by r of d , then that will change the state of d .

$go'(e_2, d, p) \supset in(e_2, d, p)$

That is, if e_2 is a going potty event by d , then that e_2 is done by d in p ;

$in(e_2, d, p) \supset state(e_2, d)$

That is, if d is in p , then it is a state of d ;

So far, we have proved all the above logical forms as is illustrated in Figure 1.

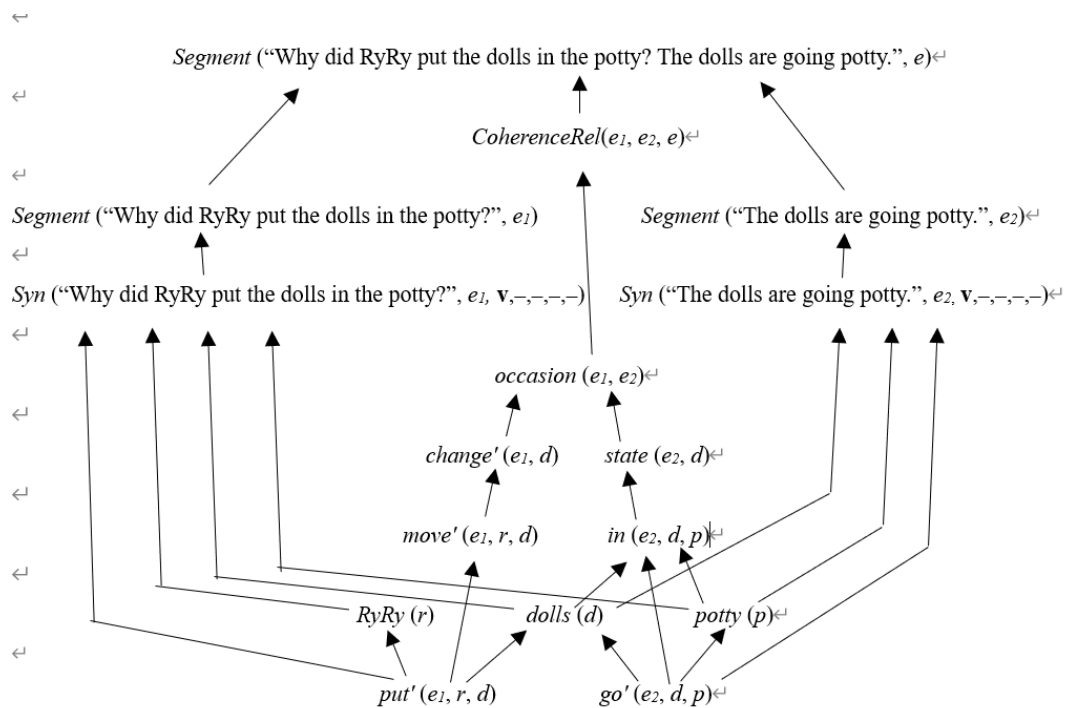


Figure 1: Interpretation of "Why did RyRy put the dolls in the potty? The dolls are going potty."

In brief, with the occasion relation between $Segment$ ("Why did RyRy put the dolls in the potty?", e_1) and $Segment$ ("The dolls are going potty", e_2), the father and his children co-constructed a coherent segment of discourse.

On the other hand, the father's intention questioning utterance and TyTy's spontaneous completion connected by the occasion relation constitute the process of intention understanding. It is the combination of participants' goals for understanding and also their memory for what they talked about that affect the degree of their explicit collaboration, not any one person's.

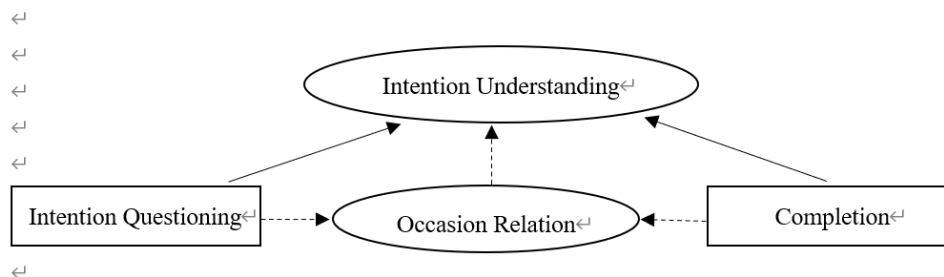


Figure 2: The collaboration on Understanding Intentions of Providing the Occasion

III. THE RESULT RELATION BETWEEN CHILDREN'S ACTS AND THEIR CONSEQUENCES

Result relation, another kind of coherence relation, also plays a crucial part in understanding human behavioral logics. Kehler defined Result relation as follow:

“**Result:** Infer P from the assertion of S_1 and Q from the assertion of S_2 , where normally $P \rightarrow Q$ ” (Kehler, 2008, p.247).

Let us consider an example. When the parents and their children are having the rehearsal of a wedding ceremony, the flower girl RyRy chased after her brother Tydus (TyTy) who held the ring to a room. Then they had a talk as follow:

- 30 Tydus: ((crying))
 31 Mom: oh=m-
 32 Dad: what happened↑
 33 Mom: what happened(hhenhh)
 34 Tydus: RyRy hit me:::
 35 Dad: RyRy (0.5) why did you hit him=what are you doing
 36 (1.2)
 37 Buddy (0.8) RyRy=why did you hit TyTy
 38 (2.2) ((RyRy raised the ring in her hand))
 39 Dad: Ahh:: She wanted the ring?
 40 Mom: Yeah
 41 Dad: Sweetheart=you are the flower↑ girl (0.2) Honey? (0.8) And you
 42 don't hit=If you want something(0.5)You come and tell mommy and
 43 daddy =°o°k?
 44 RyRy: = °o°k
 45 Dad: Can you please say sorry to TyTy↑
 46 RyRy: Sorry::↓
 47 (1.9)
 48 Mom: They are fine
 49 Dad: Ok=buddy (0.8) RyRy=don't hit
 50 (2.1)
 51 Hmmm:
 52 RyRy: °Here it is°
 53 Mom: Ahh:: (.) She's giving that back (.) buddy
 54 RyRy: Ehh?
 55 Dad: Can you [give it back to him?
 56 Mom: [You °wanna° give it back to TyTy
 57 (0.8)
 58 Dad: No
 59 ((Dad, Mom and Tydus laugh))
 60 Dad: I don't think you learned your lesson yet(hhenhh) (0.2) RyRy (0.8)
 61 That's Tydus's job=You did all the flowers (0.2) Very well (.) But
 62 this's TyTy's job= °o°k? (.) RyRy=can you give the ring back to TyTy?
 63 (2.1)
 64 Mom: Ahh
 65 ((Dad, Mom and Tydus laugh))
 66 Dad: Hold on (0.2) Let me work on the ring now(hhenhh) (1.2) RyRy (0.2)
 67 can you please give the ring back to TyTy now?
 68 (2.2)
 69 Mom: Come on
 70 (5.5)
 71 Dad: That a girl
 72 Mom: =Nice job=baby

In this conversation, when the father asked RyRy why her hit Tydus, he did not get a verbal response directly. However, when RyRy raised the ring, her father recognized her intention of getting the ring.

Suppose, plausibly enough, we can get *Segment* (“RyRy raised the ring in her hand.”, e_2) from the multimodal discourse of “RyRy raised the ring in her hand”. Together with *Segment* (“Why did RyRy hit TyTy?”, e_1) and the result relation between them, we can get a coherent *Segment* (“Why did RyRy hit TyTy? RyRy raised the ring in her hand.”, e). And we are faced with proving this expression:

$$(\exists e_1, r, t, e_2, i) RyRy(r) \wedge hit'(e_1, r, t) \wedge TyTy(t) \wedge result(e_1, e_2) \wedge raise'(e_2, r, i) \wedge the\ ring(i)$$

That is, there is a hitting event e_1 by RyRy r of TyTy t . There is a raising event e_2 by RyRy r of the ring i . The hitting event e_1 resulted in the raising event e_2 .

The required axioms in knowledge base are as follows:

$TyTy(t) \supset the\ ring\ holder(t)$

That is, $TyTy\ t$ is the original ring holder;

$the\ ring\ holder(t) \wedge hit'(e_1, r, t) \supset hit'(e_1, r, h)$

That is, if e_1 is a hitting event by $RyRy\ r$ of $TyTy\ t$ and $TyTy\ t$ is the original ring holder h , then that e_1 is also a hitting event by $RyRy\ r$ of the original ring holder h ;

$hit'(e_1, r, h) \supset force'(e_1, r, h)$

that is, if e_1 is a hitting event by $RyRy\ r$ of the original ring holder h , then that e_1 is also a forcing event by $RyRy\ r$ of the original ring holder h ;

$force'(e_1, r, h) \supset result(e_1, e_2) \wedge change'(e_2, h)$

that is, if e_1 is a forcing event by $RyRy\ r$ of the original ring holder h , then that will result in a ring holder changing event e_2 ;

$raise'(e_2, r, i) \supset hold'(e_2, r, i)$

that is, if e_2 is a raising event by $RyRy\ r$ of the ring i , then that e_2 is also a holding event by $RyRy\ r$ of the ring i ;

$hold'(e_2, r, i) \supset the\ ring\ holder(r)$

that is, if e_2 is a raising event by $RyRy\ r$ of the ring i , then that $RyRy\ r$ is the new ring holder;

$the\ ring\ holder(t) \wedge the\ ring\ holder(r) \supset change'(e_2, h)$

that is, if $TyTy\ t$ is the original ring holder and $RyRy\ r$ is the new ring holder, then there is the ring holder changing event e_2 ;

So far, we have proved all the above logical forms as is illustrated in Figure 3.

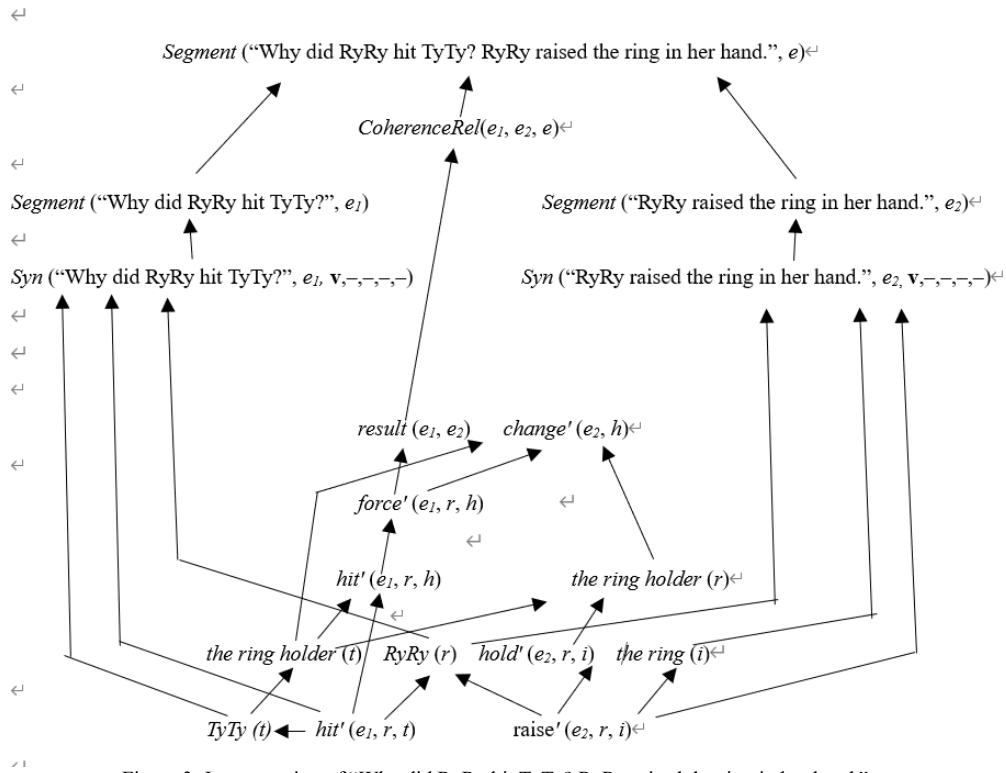


Figure 3: Interpretation of “Why did RyRy hit TyTy? RyRy raised the ring in her hand.”

As was the way discussed in the previous section, the father and his children co-constructed a coherent segment of discourse from *Segment* (“Why did RyRy hit TyTy?”, e_1) and *Segment* (“RyRy raised the ring in her hand”, e_2) and the result relation between them. On the other hand, the father’s intention questioning utterance and RyRy’s completion connected by the result relation constitute the process of intention understanding.

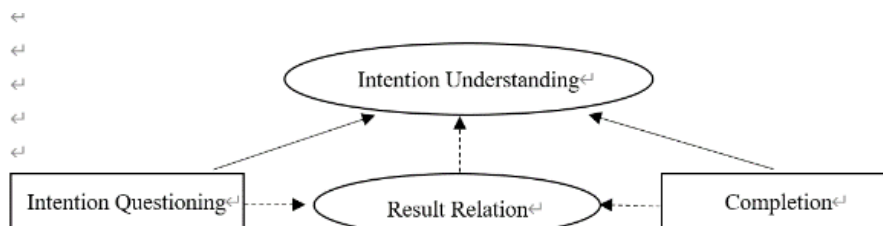


Figure 4: The collaboration on Understanding Intentions of Resulting in Something

IV. CONCLUSION

Hobbs and other researchers have explored series of approaches to establish coherent text based on making inferences, which has been applied predominantly to monologues. Combining work in speech act theory and plan recognition, we found that in dialogues the utterances determine coherence in the overall plan underlying the discourse production of the speaker's. In other words, the hearer understands utterances as speech acts and infers the speaker's intentions of doing them until the coherence is established. Besides the large set of relations in informational analyses, the satisfaction of the purpose of one discourse segment may be used to provide part of the satisfaction of the purpose of another segment, or to satisfy the purpose of another as a prerequisite. Coherence establishment is "not only a fundamental aspect of discourse interpretation, but that it needs to be accounted for in analyses of a variety of linguistic phenomena that operate across clauses" (Kehler, 2008, p.262).

Different from the intention recognition in man-machine interaction implemented by a single individual, the intention understanding here is a kind of speech act completed by the participants together in the conversation, based on the coherent segment of discourse co-constructed by each segment with the coherence relation between them. The structures of the visual question answering models currently studied are lack of reasoning abilities, resulting in their relatively simple content and form of the answers (Yang, Liu, Shi, & Li, 2019). Through establishing full coherent relation frameworks, the computer is able to output various kinds of answers to corresponding questions input.

REFERENCES

- [1] Davis Ernest. (2005). Knowledge and Communication: A first-order theory. *Artificial Intelligence*, 166 (1): 81-139.
- [2] Gordon Andrew S. (2004). The representation of planning strategies. *Artificial Intelligence*, 153(1):287-305.
- [3] Hobbs, Jerry R. (2008). "Abduction in natural language understanding." In L. R. Horn and Gregory Ward, *The Handbook of Pragmatics*. Blackwell Publishing Ltd., 724-741.
- [4] Hobbs Jerry R, Sagae Alicia, and Wertheim Suzanne. (2012). Toward a commonsense theory of microsociology: Interpersonal relationships. In *Proceeding of the Seventh International Conference (FOIS 2012)*, Graz, Austria. Amsterdam: IOS Press. 249-262
- [5] Hobbs Jerry R. and Mulker-Mehta Rutu. (2013). Toward as formal theory of information structure. In Bernd-Olaf Kuppers, Udo Hahn, and Stefan Artmann (eds.), *Evolution of Semantic Systems*. Berlin: Springer-Verlag. 101-126
- [6] Hobbs, Jerry R and Gordon Andrew S. (2014) Axiomatizing complex concepts from fundamentals (invited paper). In *Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2014)*, Kathmandu, Nepal. Heidelberg, Germany: Springer.341-365
- [7] Hobbs, Jerry R. (2017). *A Formal Theory of Commonsense Psychology—how people think people think*. Cambridge: Cambridge University Press.
- [8] Kehler Andrew. (2008). "Discourse Coherence." In L. R. Horn and Gregory Ward, *The Handbook of Pragmatics*. Blackwell Publishing Ltd., 241-265.
- [9] Kehler Andrew. (2002). Coherence, Reference, and Theory of Grammar. Stanford, CA: CSLI.
- [10] Ronen I. Brafman and Carmel Domshlak. (2009). Preference handling—an introductory tutorial. *AI Magazine*, 30 (1): 58—86.
- [11] Scherer Klaus R. (2010). Emotion and emotional competence: Conceptual and theoretical issues for modeling agents. In *A Blueprint for Affective Computing: A Sourcebook and Manual, Affective Sciences*. Oxford: Oxford University Press.
- [12] Wilkes-Gibbs, Deanna. (1986). "Collaborative Processes of Language Use in Conversation." PhD diss., Department of Psychology, Stanford University.
- [13] Wiebe Janyce, Wilson Theresa, and Cardie Claire. (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*. 39 (2-3): 165-210.
- [14] William Croft and D.A. Cruse. (2004). *Cognitive Linguistics (Cambridge Textbooks in Linguistics)*. Cambridge: Cambridge University Press.
- [15] Yang Rui, Liu Ruijun, Shi Yuqian, Li Shanxi. (2019). "A Survey of Research on Visual Question Answering Oriented to Intelligent Interaction." *Journal of Electronic Measurement and Instrumentation*. 33, no. 2(Feb):117-124 <https://doi.org/10.13382/j.jemi.B1801476>.

Hongya Fan was born in Shanxi, China in 1974. She received her PH.D. degree from Nankai University, China in 2012.

She is currently an associate professor in the School of Foreign Languages, Shanxi University, Shanxi, China. Her research interests include pragmatics, discourse analysis and identity construction.

Zeshan Ren was born in Shanxi, China in 1997. He is currently a postgraduate student in Shanxi University, majoring in Foreign linguistics and Applied Linguistics. His research interests include pragmatics and identity construction.