# Information-Theoretic Account of Dialogue Utterances

Yan Qu
Computational Linguistics Program
Carnegie Mellon University
yqu@cs.cmu.edu

## 1 Thesis Statement

Agents engaged in cooperative problem solving face the problem of what information to communicate from among the wealth of information of each agent. Only recently have researchers begun to study what information to include in a dialogue, when to include it, and the efficiency consideration for achieving the communication effect of the information ([6, 7, 8, 4, 3, 5]). However, previous researches fail to deliver an objective model of selecting information for efficient communicative dialogue design. In this paper, I describe an objective measure of measuring information of statements and the communicative effectiveness of statements. The theoretical framework is Bar-Hillel and Carnap's fundamental work on Semantic Information [1]. I first introduce the semantic information theory. Then I apply this theory to measure the information of two kinds of dialogue utterances, *positive constraint background statements* and *negative constraint background statements.* These utterances are taken from a corpus of naturally occuring dialogues, in which two people try to schedule a time to meet. Lastly, I show the objective measure of information in dialogue utterances is helpful for determining the communicative effectiveness of these utterances in multi-agent communication.

## 2 Semantic Information Theory

Bar-Hillel and Carnap's semantic information theory provides a definition of content of statements as well as an objective way of measuring content and information. *Content* of a statement $s$ is defined as the set of all statements logically excluded by $s$. We expect that the more precise a statement is, the less probable it is. As a statement grows in length and precision, the content, or set of excluded statements, grows. Thus the size of the content of a message is inversely related to the message's probability and proportional to the information in the statement. A content-based information measure is defined as

$$I(s) = -logPr(s) \tag{1}$$

and

$$Pr(s) \propto sizeof[Cont(s)] \tag{2}$$

where $s$ is a statement, and $Pr(s)$ is the probability of statement $s$. From (1) and (2), the information in statement $s$ can be calculated by:

$$I_s = log_2 \frac{1}{1 - sizeof(s)} \tag{3}$$

Bar-Hillel and Carnap's semantic model provides a numeric measure of the amount of information equal to that provided by Shannon's measure. However, their measure is derived from the principles of logic rather than from probability theory, and thus allows for an interpretation of information based on the notion of an informative *statement* in a language.

# 3 Information-Theoretic Account of Dialogue Utterances

## 3.1 Information measure of atomic positive and negative constraint background statements

Suppose in the scheduling domain, there are $n$ atomic time slots $t1, t2, ...tn$ for negotiation. To simplify the problem, suppose there is only one one-argument predicate *free* in this domain. The attribute *busy* is represented as the negation of *free*. Thus the whole language is $L_n^1$, with a conceptual space of size $2^n$.

Now consider the case of positive constraint background statements. The content of statement $free(t_i)$ is the set of all statements logically excluded by the statement. For example, it logically excludes all statement with literal

$-free(t_i)$, denoted by $Cont(free(t_i))$. The proportion of the space occupied by $Cont(free(t_i))$, as opposed to the entire conceptual space, is $\frac{1}{2}$. The information is $I_s = log_2 \frac{1}{1-\frac{1}{2}} = 1$.

Along the same line of calculations, we see negative constraint background statements are not very different from positive constraint background statements. The content of statement $-free(t_i)$ is the set of all statements logically excluded by the statement. In this case, it logically excludes all statement with literal $free(t_i)$, denoted by $Cont(-free(t_i))$. The proportion of the space occupied by $Cont(-free(t_i))$, as opposed to the entire conceptual space is again $\frac{1}{2}$. The information is $I_s = log_2 \frac{1}{1-\frac{1}{2}} = 1$.

A negative constraint background statement is a negated factual statements of a positive constraint background statement. Though the contents of the two types are different, the information carried by both types is actually the same.

## 3.2 Information measure of multiple positive and negative constraint background statements

The above discussion deals with statements with individual elements at the atomic clause level. Atomic statements can be combined together to form more complex statements. This section illustrates how to calculate information for conjunction of statements.

Intuitively, conjunctions should carry more information than the individual statements that constitute the conjunction. The intuition can bear out graphically by looking at the set of statements logically excluded by the conjunction and its individual components.

Figure **??** gives the graphical illustration for statement $free(t_1)$ and statement $free(t_1) \wedge free(t_2)$. By stating statement $free(t_1)$, the whole conceptual space is cut by half as shown in the left diagram in **??**. The set of statements excluded by saying statement $free(t_1)$ is represented by the shaded area. By stating $free(t_1)$ and $free(t_2)$, the conceptual space is first cut by half by statement $free(t_1)$, and then the statement $free(t_2)$ also cuts the conceptual space in half. The total set of statements excluded by the conjunction of statements is the total of the shaded areas excluded by both statements.

The content of the conjunction $free(t_1) \wedge free(t_2)$ is therefore:

$$\begin{aligned} cont(free(t_1) \wedge free(t_2)) &= cont(free(t_1)) + cont(free(t_2)) \\ &\quad - cont(free(t_1) \vee free(t_2)) \\ &= sizeof(free(t_1)) + sizeof(free(t_2)) \\ &\quad - sizeof(free(t_1) \vee free(t_2)) \\ &= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

The information carried by the individual statement and the conjunction of both statements are:
$I_{free(t_1)} = I_{free(t_2)} = log_2 \frac{1}{1-\frac{1}{2}} = 1$
$I_{free(t_1) \wedge free(t_2)} = log_2 \frac{1}{1-\frac{3}{4}} = 2$
Note that the set of statements consistent with $free(t_1) \wedge free(t_2)$ contains the set of statements consistent with $free(t_1)$ or $free(t_2)$. That is, the set of statements excluded by $free(t_1) \wedge free(t_2)$ contains and implies the set of statements excluded by $free(t_1)$ or $free(t_2)$. The result confirms our intuition that conjunctions are more informative than the individual statements that constitute the conjunction.

In general, the content of a conjunction of $n$ atomic clauses is $cont = (\frac{1}{2})^1 + (\frac{1}{2})^2 + ... + (\frac{1}{2})^n$. Information of the conjunction is $I = log_2 \frac{1}{1-(1-cont)} = n$.

# 4 Information Transfer Effect

In the previous section, I provided the information measure for statements in term of a *single* agent's belief space. Next, I show how information is calculated in multi-agent interaction. Clark et all [2, 9] propose that agents engaged in a conversation need to bring a certain amount of common ground to a conversation in order to understand each other. The process of adding to this common ground is *grounding*. Grounding can be seen as a process of updating the mutual beliefs of the agents.

I focus on a particular mutual belief space – the conceptual problem solving space, and on how this shared problem solving space is updated by information transfer. First, I describe the shared conceptual problem solving space, and then illustrate how positive and negative constraint background statements update the problem-solving space.

## 4.1 Mutual beliefs

A general view of mutual beliefs is to assume that each agent has a mutual belief space containing all beliefs he thinks of sharing with another agents. Mutual belief is generally defined in term of belief following [10]. The connection between belief and mutual belief is defined by the fix-point axiom, which captures the circularity of mutual belief [11]:
$SH_{xy}p \equiv BEL_x(p \wedge SH_{yx}p)$
where $SH_{xy}p$ means that agents $x$ and $y$ mutually share the belief that $p$.

## 4.2 Negotiation as search

Problem solving can be characterized as a task of searching through a large, predefined space of hypotheses to find the hypothesis that is the goal of the problem. In the scheduling domain, though agents are assumed to have complete knowledge of their respective calendars, they have incomplete (or no) knowledge as to the other's calendar. Thus the scheduling process is to search through a hypothesis space to find out the goal for both agents. While a variety of representations are possible to represent a hypothesis, I choose a simple representation in which each hypothesis is described by a conjunction of individuals with constraints on the values of one single attribute. For example, the entire hypothesis space for negotiation over $n$ time slots $t_1, t_2, ..., t_n$ with one attribute $free$ will be represented as:

$<?, ?, ..., ? >$

in which each question mark corresponds to a clause which shows whether an individual time slot $t_i$ is $free$ or not. Any instantiation of the question mark with a positive clause is considered the goal of the problem solving. The process of negotiation is to find the positive instantiation of any of individuals.

## 4.3 Update the mutual belief space

The design of multi-agent interaction systems needs making many assumptions about the agents and the task. I make the following assumptions as to the scheduling agents:

1. **Assumption 1**: Agents have no prior beliefs of the other agent's calendar.
2. **Assumption 2**: Agents have perfect understanding of each other during negotiation.
3. **Assumption 3**: If a time $t$ is bad for one agent, then negotiation is closed for this time $t$.
4. **Assumption 4**: If a time $t$ is good for one agent, then the other agent can either accept the time $t$, in which case the negotiation ends successfully, or reject the time $t$, in which case time $t$ is closed for negotiation. I postulate that when alternatives are involved, the partner must respond in order to give a status for time $t$.

It follows from assumption 1 and 2 that the agents start the negotiation with a shared problem solving space with all possible hypotheses. Since perfect communication is assumed, at each stage of negotiation, each agent updates mutual beliefs. However, the mutual problem solving space may or may not be updated depending on what information is conveyed.

**Definition. Communication Effectiveness (CE):** CE measures the effectiveness of a statement on the reduction of mutual search space. The formula for calculating $CE$ is:

$$CE = \frac{sizeof(s)}{N} \tag{4}$$

where $sizeof(s)$ denotes the proportion of the set of statements excluded by the statement $s$, and $N$ is the minimal total number of statements that are required in order to finalize the reduction of mutual search space.

It follows from assumptions 2 and 3 that a negative constraint background statement $i$ has $CE = sizeof(i)$. In particular, suppose that agent A utters statement $-free(t)$. By assumption 2, $Bel_B Bel_A(-free(t))$ holds. By assumption 3, since the time slot $t$ is closed for negotiation, the mutual search space is updated. Since no additional statement is required to finalize the mutual search space reduction, the total numbers of statements required is just one, i.e. $i$ itself.

Now consider the case of positive constraint background statement $i$. In particular, suppose agent A utters $free(t)$, agent B will update his belief model as $Bel_B Bel_A(free(t))$. In this case, however, the mutual belief space can not be updated since the status of time slot $t$ for agent B is not known to agent A. By assumption 4, agent B has to respond with either acceptance or rejection to inform agent A of the status of the time slot $t$. In the case of B rejecting $t$, the mutual problem solving space is reduced. If B accepts $t$, the mutual problem solving space is also reduced. At the same time, the agents reach a common goal state and negotiation is over. In either case, B's response is needed to finalize the search space reduction implied by A's proposal. Thus the minimal total number of sentences required to finalize the search space reduction is 2. Consequently $CE = \frac{sizeof(i)}{2}$.

# 5 Conclusions

In this paper, I apply the semantic information theory to two types of dialogue utterances: the positive constraint background statements and the negative constraint background statements. The objective measure of the information and communicative effectiveness of the dialogue utterances is useful for developing efficient human-computer or computer-computer interaction systems.

# References


[1] Yehoshua Bar-Hillel and Rudolf Carnap. Semantic Information. *The British Journal for the Philosophy of Science*, 4(13):147–157, 1953.

[2] Herbert H. Clark and Edward F. Schaefer. Contributing to discourse. *Cognitive Science*, 13, 1989.

[3] C. I. Guinn. A computational model of collaborative discourse. In *Workshop Proceedings from AI-ED 93 World Conference on Artificial Intelligence in Education*, 1993.

[4] Curry I. Guinn. *Meta-Dialogue Behaviors: Improving the Efficiency of Human-Machine Dialogue.* PhD thesis, Duke University, 1995.

[5] Elise H. Turner. Selecting information to communicate. In *AAAI Workshop on Planning for Inter-Agent Communication*, 1995.

[6] Marilyn Walker. *Information Redundancy and Resource Bounds in Dialogue*. PhD thesis, University of Pennsylvania, 1993.

[7] Marilyn Walker and Owen Rambow. The role of cognitive modeling in achieving communicative intentions. In *Proceedings of the Seventh International Workshop on Natural Language Generation*, 1994.

[8] Marilyn Walker and S. Whittaker. Mixed-initiative in dialogue: An investigation into discourse segmentation. In *Proceedings of the 28th Annual meeting of the Association for Computational Linguistics*, 1994.

[9] Deanna Wilks-Gibbs and Herbert H. Clark. Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2):183–194, 1992.

[10] S. R. Schiffer. Meaning. *Oxford University Press*, 1972.

[11] G. Harman. Review of "Language Behavior" by Jonathan Bennett. *Language*, 53:417–424, 1977.