Avoiding Unwanted Conversational Implicatures in Text and Graphics

Joseph Marks and Ehud Reiter†

Aiken Computation Laboratory
Harvard University
Cambridge, Mass 02138
marks@harvard.harvard.edu
reiter@alpna.edinburgh.ac.uk

Abstract
We have developed two systems, FN and ANDD, that use natural language and graphic displays, respectively, to communicate facts about objects to humans. Both systems must deal with the fundamental problem of ensuring that their output does not carry unwanted and inappropriate conversational implicatures. We describe the types of conversational implicatures that FN and ANDD can avoid, and the computational strategies the two systems use to generate output that is free of unwanted implicatures.

Introduction
We have developed a natural-language generation system (FN) and an automatic graphic-design system (ANDD) that use their respective media to communicate facts about objects to users. Both systems address the surface content-determination problem: given as input a set of predicates about objects in the discourse domain that need to be communicated to the user, FN and ANDD are expected to produce output that either directly informs, or allows the user to infer, that even if the input set are true. One of the most important conceptual and computational problems that both systems face is generating output that is free of unwanted conversational implicatures (Grice 1975). Generating syntactically and semantically correct communications (utterances and diagrams) that convey the target predicates is relatively straightforward for both FN and ANDD; what is difficult is the pragmatic problem of ensuring that the utterances and diagrams do not mislead the user into making incorrect conversational implicatures (Hirschberg 1984).

More precisely, FN and ANDD produce (respectively) attributive descriptions of individuals and network diagrams. Attributive descriptions of individuals are natural-language object descriptions that are intended to inform the hearer that a particular object has certain properties: they differ from referring expressions, which are object descriptions that are intended to identify particular objects in the current discourse context (Donnellan 1966). Network diagrams (Bertin 1983) are graphical displays that are used to depict network models. A network model is an attributed graph, i.e., a graph in which the vertices and edges have nominal and quantitative attribute values associated with them. Network models are useful abstractions for many complex systems, including computer, communication, and command-and-control systems.

The FN system determines the words that will best communicate to the user that the object in question has the relevant properties; however, it does not decide which properties are important in the current discourse context and need to be communicated to the user. The ANDD system designs a network diagram that depicts a given network model; it does not decide what information the network model should contain, or how a real-world system is modeled as a network model. Both systems are intended to be used as components of a larger collaborative system, in which a discourse analysis subsystem that is based on the SharedPlan framework (Grosz & Sidner 1990; Lochbaum, Grosz, & Sidner 1990) will be used to solve the information content-determination problem, i.e., the problem of deciding what information is important in the current discourse context and needs to be communicated to the user. The full system may also use media coordination (Feiner & McKeown 1990) to tie FN and ANDD together, but such coordination has not been investigated to date.

Conversational Implicatures
Implicatures in Text
Suppose a speaker is given the communicative goal of informing a human hearer that a particular object is a computer network with the attributes \{data-rate:10Mbit/sec, circuit-type:packet-switched\}. Consider two possible descriptions that might be used to convey this information:

1a) "10Mbit/sec packet-switched computer network"
1b) "Ethernet"

One might think that (1a) and (1b) convey the same information, provided that the hearer knows that Ethernet is a computer network that have the attributes \{data-rate:10Mbit/sec, circuit-type:packet-switched\}. However, if the hearer does not have this domain knowledge, the use of utterance (1a) might lead her to draw the conversational implicature (i.e., to interpret the utterance as implicating) that the object being described is not an Ethernet — because if it were, the hearer would reason, then the speaker would have uttered (1b).
A similar phenomenon occurs in referring expressions: consider, for example, the difference between

2a) "Sit by the table"
2b) "Sit by the brown wooden table"

If there was only one visible table (which was brown and made of wood), then utterances (2a) and (2b) would both fulfill the referring goal, because a hearer who heard them would have no trouble picking out the object being referred to. However, a hearer who heard utterance (2b) would probably draw the additional conversational implicature that it was somehow important that the table was brown and made of wood, i.e., that the speaker was trying to do more than just identify the table. If the speaker did not have this intention, and only wanted to tell the hearer where to sit, then the hearer would have drawn an incorrect conversational implicature.

Implicatures in Graphics

Figure 1 shows a typical network diagram that depicts a network-model abstraction of a computer-disk system. The topology of the network model is communicated directly by the connectivity of the network diagram, and the nominal and quantitative attributes of the model are communicated directly by the graphical properties (e.g., shape, pen width) of symbols, by text labels, and by diacritical symbols (e.g., enclosures), as indicated in the diagram legend.

Given that Figure 1 conveys precisely the desired information, Figures 2 and 3 are variants of Figure 1 that might be said to carry unwanted conversational implicatures. For example, in Figure 2 the pen-width used in drawing the channel-facility queue symbol differs from the pen width used for all other queue (and server) symbols in the diagram. A viewer of this diagram might conclude that the channel-facility queue is somehow different from the other queues (and the servers), because otherwise the diagram designer could simply have used the same pen width to draw the symbol. Other unwanted conversational implicatures in Figure 2 result from perceptual grouping. For example, the way the disk symbols are perceptually grouped by proximity to form two ‘gestalts’ might lead the viewer to conclude that there is some semantic basis to the perceived grouping, because otherwise the designer would not have positioned the disk symbols in this way. In addition, the layout of the device-queue symbols violates the Gestalt Principle of Good Continuation (Kaufman 1974), resulting in the implicature that one of the device queues is uniquely different, because otherwise the designer could have positioned the symbols as in Figure 1, for example. Finally, in Figure 3 there is an ordering (by size) of the node symbols: this ordering implicates that there must be a similar ordering relation among the vertices in the network model, which is not true (given Figure 1).

A thorough analysis that relates these graphical phenomena to known linguistic phenomena is a topic for future research. For now, we note that in each instance described above information is conveyed that differs from what is communicated directly by the symbols in the diagram. Furthermore, the ultimate source of implicature is the assumption that the designer and viewer are following Grice’s Cooperative Principle (Grice 1975) in their discourse, i.e., that the designer is not trying to mislead the viewer, and that his design decisions can therefore be considered as appropriate contributions to the discourse.

Analysis and Discussion

Grice (1975) proposed a number of maxims of conversation that cooperative communicating agents usually obey (cooperative agents might disobey the maxims to achieve a particular communicative goal). These maxims fall into four categories: Quality, Quantity, Relation, and Manner. We consider how these maxims can be applied to avoid unwanted conversational implicatures in text and graphics. Although Grice’s maxims apply to both information-content determination and surface-content determination, we will consider here only their application to the surface-content-determination tasks performed by the FN and ANDD systems.

Quality: Grice’s maxims of Quality require utterances to be truthful. For our purposes, these maxims translate into a constraint that everything that is explicitly communicated to the user must be true.

Quantity: The maxims of Quantity require utterances to be neither more or less informative than is necessary for the purpose of the exchange. For natural language, these maxims forbid utterances from containing elements that are irrelevant, can be inferred from other parts of the utterance, or are otherwise redundant. For instance, in the referring expression example of Section 2.1, the adjectives “brown” and “wooden” are unnecessary for fulfilling the referring goal, and hence their inclusion in utterance (2b) leads to false implicatures.

The application of Grice’s Quantity maxims to surface-content determination in graphics is less clear. In some cases extra information in graphical diagrams leads to false implicatures: for example, communicating the same network-model attribute via two distinct graphical properties (e.g., using both symbol color and shape to communicate the same information) could lead to unwanted implicatures, because the viewer might conclude that there must have been some reason the designer used two graphical properties instead of one. In other cases, however, extra information is acceptable and even desirable; for example, if the communicative goal is to identify the overloaded servers in the disk subsystem, the network diagram in Figure 1 (which depicts not only the overloaded servers, but also servers that are not overloaded, various other objects and how they interconnect, with...
and subsystem affiliation) is preferable to a diagram that depicts only the two overloaded servers and nothing else. In yet other cases, the inclusion of 'redundant' information may actually be necessary to avoid leading the user to make unwanted implicatures: for example, the use of perceptual grouping to reinforce information communicated in other ways is sometimes essential (in Figure 1 the perceptual groupings of the node symbols reinforce the subsystem-affiliation information, whereas in Figure 2 they do not), even though the perceptual grouping is in some sense redundant.

Relation: The maxim of Relation requires utterances to be relevant to the discourse. This maxim primarily affects the information-content determination task, but it also has some impact on surface-content determination. For example, natural-language utterances should not contain irrelevant elements, and graphical displays should avoid the use of spurious graphical-property values (the use of a distinct pen-type for the channel-facility queue symbol in Figure 2 is an example of a spurious or irrelevant graphical-property value that can lead to an unwanted conversational implicature.)

Manner: Grice's maxims of Manner concern obscurity, ambiguity, brevity, and orderliness. The concept that probably has the most impact on natural-language generation is brevity: shorter descriptions are preferred over longer ones, because the use of an unnecessarily long utterance may imply that a shorter utterance could not be used (e.g., the use of utterance (1b) implicates that utterance (1a) could not have been used). The category of Manner is also important for graphics, but some additional concepts are needed to cover issues that are unique to graphic design. Two important concepts are appropriate perceptual organization and perceptual limits. To achieve appropriate perceptual organization spurious perceptual organizations that are orthogonal to the information being conveyed in a diagram should be avoided. This is necessary to avoid the kinds of unwanted implicatures caused by perceptual grouping and ordering that are illustrated in Figures 2 and 3. The concept of perceptual limits also concerns human visual perception: a designer should limit the number of graphical-property values used in a diagram so that the values may be easily distinguished, and to ensure that the values are perceptually dissimilar (well-known design heuristics that concern perceptual limits can be found in (Bertin 1983)).

Basic Level: An additional source of conversational implicature, which was proposed by Cruse (1977) and Hirschberg (1985), is the failure to use basic-level classes (Roch 1978) in an utterance. For example, consider the utterances

3a) "I have a red shirt"
3b) "I have a red T-shirt"
3c) "I have a red piece of clothing"
3d) "I have a carmine shirt"

Red and shirt are probably basic-level for most urban Americans. Accordingly, utterance (3b) carries the conversational implicature that it is important that the object is a T-shirt and not some other kind of shirt; utterance (3c) carries the conversational implicature that the speaker was not able to categorize the object as a shirt or any other type of commonly used clothing; and utterance (3d) carries the conversational implicature that it is relevant that the object is carmine and not some other shade of red. If none of these implicatures are desired, then utterance (3a) should be generated.

A similar phenomenon can occur in the design of graphical displays. For example, a network diagram in which all the symbols are blinking will likely cause the viewer to conclude that there is some reason why the diagram could not be drawn with non-blinking symbols. Similarly, a network diagram in which all the node symbols are tiny (or huge) will lead to the conclusion that symbols of "normal" size could not be used. In other words, some graphical-property values seem to be preferred, and the use of a non-preferred value in a graphical display may implicate that the preferred value could not have been used.

The FN System

The FN system (Reiter 1990a) generates natural-language descriptions that are attributive descriptions of individuals, i.e., that communicate to the hearer that a particular object has certain attributes. Utterances (1a) and (1b) are examples of such descriptions. [Note that FN does not generate referring expressions such as (2a) or (2b).] FN assumes that the user has some domain knowledge, and takes advantage of this domain knowledge to form better descriptions. For instance (cf. utterances (1a) and (1b)), if FN wished to inform a user that a certain object is a computer network that had the attributes (data-rate: 100 Mbps, circuit-type: packet-switched), and the object being described is in fact an Ethernet, then FN would generate (1b) if it believed that the user knew that Ethernets were networks that had these attributes, and (1a) otherwise.

FN formalizes the problem of avoiding unwanted conversational implicatures by requiring generated utterances to be maximal elements under a preference function. More precisely, let \( \gg \) be a preference function that prefers utterances that do not contain unnecessary elements, that use basic-level classes, and so forth. Let \( D \) be the set of utterances that are truthful and that successfully fulfill the system's communicative goal (e.g., inform the hearer that the object is a computer network with the attributes (data-rate: 100 Mbps, circuit-type: packet-switched)). Then, an utterance in \( D \) is said to be free of false implicatures if it is a maximal element of \( D \) with respect to \( \gg \). Hence, an utterance \( B \) in \( D \) is free of false
implicatures if there is no utterance \( A \) in \( D \), such that \( A \gg B \). The preference-function formalization of conversational implicature is similar to the partially-ordered sets that Hirschberg (1984, 1985) used to formalize scalar implicatures (Gazdar 1979).

FN's overall preference function consists of three separate preference rules: No Unnecessary Components, Lexical Preference, and Local Brevity. We make the assumption that there are no conflicts between these preference rules, i.e., that it is never the case that utterance \( A \) is preferred over utterance \( B \) by one preference rule, but \( B \) is preferred over \( A \) by another preference rule.

No Unnecessary Components: The No Unnecessary Components rule is motivated by Grice's Quantity maxim. Formally, it states that \( A \gg B \) if \( A \) uses a subset of the modifiers (e.g., adjectives, prepositional phrases, relative clauses) that \( B \) uses. Hence, a 'free-of-false-implicatures' description cannot contain any unnecessary modifiers.

Utterance (2b) is an example of a referring expression that contains unnecessary components (the adjectives “brown” and “wooden”) and hence carries false implicatures. An example of an attributive description that contains an unnecessary component is the following:

1c) "packet-switched Ethernet"

If the user knows that Ethernets have the attribute circuit-type:packet-switched, the modifier “packet-switched” is redundant, and hence utterance (1b) is preferred over utterance (1c) by the No Unnecessary Components preference rule. Therefore, FN will not generate utterance (1c) for such a knowledgeable user, since it might lead her to draw incorrect conversational implicatures (e.g., that there were some Ethernets that did not have the attribute circuit-type:packet-switched).

Lexical Preference: The Lexical Preference rule is an extension of the basic-level implicature principle. This rule assumes there is a lexical preference hierarchy among the different lexical units that can be used in an utterance; the lexical-preference hierarchy always marks basic-level classes as preferred over other lexical units, and it may include other preference relations as well. The Lexical Preference rule then states that \( A \gg B \) if \( A \) and \( B \) have the same syntactic and semantic structure, and every lexical unit used in \( A \) is equal to or lexically preferred over the corresponding lexical unit in \( B \). Thus, free-of-false-implicatures descriptions need to use preferred lexical units whenever possible. An example of Lexical Preference is that utterance (3a) is preferred over utterances (3b), (3c), and (3d), because (3a) uses basic-level classes.

Local Brevity: This rule states that it should not be possible to generate a shorter description by introducing a single new classification or modifier and then eliminating old modifiers that are no longer necessary. Formally, say \( A \gg B \) if length(\( A \)) < length(\( B \)), and size(components(\( A \)) - components(\( B \))) = 1 (i.e. \( A \) has exactly one component not present in \( B \); \( B \) can have an arbitrary number of components not present in \( A \)). Then, \( \gg \) is defined as the transitive closure of \( \gg \). FN uses the number of open class words as its measure of description length.

This rule is an approximation to the more powerful rule of requiring descriptions to be as short as possible. The latter rule is rejected for complexity-theoretic reasons: it is impossible to find the shortest description with a polynomial-time algorithm (Reiter 1990b). An example of Local Brevity is that utterance (1b) is preferred over utterance (1a), if the user has appropriate domain knowledge, because the single class Ethernet in utterance (1b) replaces both the class computer network, and the attributes [data-rate:10Mbit/sec, circuit-type:packet-switched], in utterance (1a).

The generation algorithm used by FN is described in Reiter (1990a). The algorithm works by incremental improvement: it starts with an initial description, and then repeatedly replaces the current description by a preferred description, as long as such a description can be found. The algorithm terminates when it reaches a point where there are no descriptions that are preferred over the current one: this means the current description must be maximal and hence free of false implicatures.

The ANDD System

The ANDD (Automated Network-Diagram Designer) system automatically designs network diagrams. Its input is a network model and a list of relations that describe the information to be communicated to the user. These relations indicate which attribute values need to be communicated, whether the actual values of the quantitative attributes or just their relative ordering are important, and which aspects of the network interconnection should be emphasized (e.g., source-sink paths, feedback loops, hub structures).

The ANDD system uses an original formulation of syntax and semantics for network diagrams (Marks 1990) to relate graphical-display symbols to elements and attributes of the network model. This formulation is crucial to ANDD’s ability to automatically design network diagrams that are free of unwanted conversational implicatures. It also leads to a natural breakdown of the overall design problem into distinct design tasks. We first provide an overview of syntax and semantics, and then describe the various design tasks, how they are automated, and the ways in which they try to eliminate unwanted conversational implicatures.

The Syntax of Network Diagrams: The syntax of network diagrams used by ANDD describes the symbols, graphical properties, and perceptual-organization phenomena that a designer can use to communicate information that displays of data can be found in (Friedell, Barmett, & Kramlich 1982), (Friedell 1984), (Feiner 1985), and (Mackinlay 1986).

3 Previous research on automating the generation of graphical displays of data can be found in (Friedell, Barmett, & Kramlich 1982), (Friedell 1984), (Feiner 1985), and (Mackinlay 1986).
The morphological elements of network diagrams are node symbols, link symbols, text labels, and diacritical symbols. Symbols from each class have their own set of graphical properties (e.g., node symbols have the graphical properties of shape, size, pen color, and fill color). The most novel aspect of the syntax used by ANDD is the inclusion of relations that describe the perceptual organization of symbols, such as sequential layout (top-to-bottom or left-to-right), proximity grouping, alignment, symmetry, similarity, and ordering.

The reason for including perceptual grouping and ordering relations in the syntax is that perceptual organization is a property of the human visual system that we cannot disable: it is virtually impossible to design meaningful network diagrams for which no perceptual organization will occur. Instead, the ANDD system tries to actively exploit this property of the visual system (and thereby satisfy the maxim of Manner) by designing diagrams that contain appropriate perceptual groupings and orderings that communicate or emphasize certain information. The two vertical, evenly spaced node alignments in Figure 4 (designed by ANDD) are an example of this technique: the node interconnections are clear from the links in the diagram, but two special paths in the network are visually emphasized by the active use of perceptual grouping.

The Semantics of Network Diagrams: A major difference between language and graphics is that language has a fixed vocabulary (the English lexicon), while graphics, in general, does not. The meaning of a graphical property such as shape or color usually changes from one graphical display to the next; a graphic designer decides what information is to be communicated by each graphical property, and uses the diagram legend to inform the viewer of his design decisions.

In the ANDD system, the semantics of a network diagram is described in terms of an expressive mapping, which relates the network-model information to be communicated to a syntactic description of the network diagram. An expressive mapping includes the obvious direct mapping from vertices and edges in a network model to node and link symbols in a network diagram. In addition, it maps network-model attributes and relations onto network-diagram graphical properties and perceptual-organization relations.

The expressive mapping for the network diagram in Figure 4 maps vertex names onto text labels, vertex types onto node-symbol shapes, and edge types onto link pen types. In addition, node fill intensity is used to visually emphasize some nodes relative to others (emphasis is one of the vertex quantitative attributes in the network model). Finally, relations of sequential layout and alignment are used to help distinguish the two significant source-to-sink paths in the network model.

Automating the Design Tasks: ANDD's syntactic and semantic formulation leads naturally to two central design tasks: (i) creating an expressive mapping, and (ii) instantiating a diagram that conforms to the expressive mapping.

ANDD uses a rule-based system to construct expressive mappings: the rules encode the conditions under which the various network-model attributes and relations should be mapped onto certain graphical properties and syntactic relations. The rules incorporate a limited amount of graphic-design expertise to enable the system to avoid the kind of unwanted conversational implicature shown in Figure 3: for example, no rule maps a non-quantitative attribute onto a graphical property that will be perceived as ordered. In addition, the rules avoid violations of the maxims of Quantity (e.g., the use of multiple graphical properties to communicate a single network-model attribute), and other well-known design flaws that qualify as violations of the maxims of Manner (e.g., the use of too many colors, intensities, or shapes).

Once an expressive mapping is created, ANDD extracts the set of syntactic relations that must appear in the network diagram to communicate the desired information. It then attempts to instantiate a network diagram in which exactly these required syntactic relations appear. The diagram-instantiation task can be thought of as a constraint-satisfaction problem, in which the constraints are the required syntactic relations.

The diagram-instantiation task has two subtasks: picking locations for the symbols, and choosing suitable values for the other graphical properties. The latter task is performed first, by another rule-based system, which chooses graphical-property values to satisfy syntactic relations governing the perceptual-organization phenomena of similarity and ordering. The various symbol shapes, node fill intensities, and link pen types shown in Figure 4 were chosen from a palette of graphical-property values by these rules. Again, the rules must encode a certain degree of graphic-design expertise to avoid unwanted conversational implicatures due to violations of various maxims of Lexical Preference and Manner, e.g., the inappropriate use of preferred graphical-property values, or graphical-property values that are hard to distinguish visually.

Computing diagram layout is computationally much more difficult. The ideal algorithm for this task would compute a diagram layout with exactly the required syntactic relations governing such things as proximity grouping, alignment, and symmetry. However, such an algorithm appears to be computationally prohibitive. As a reasonable tradeoff between speed and performance (the eventual context is that of an interactive system), the current ANDD system uses a third rule-based system that heuristically picks locations for symbols based on the desired syntactic constraints and the existing layout of local regions of the nascent diagram. In principle, this heuristic approach will not always produce a network diagram with exactly the desired syntactic characteristics (and consequent absence of unwanted conversational implicatures); in practice, the system has worked well in many cases. We continue to refine our current approach.
to the layout task, and to investigate possible alternative approaches, including the use of mathematical-programming and simulated-annealing techniques.

It is interesting to note that FN and ANDD seem to face similar computational-tractability problems. For both systems, some of the most obvious formalizations of the unwanted-conversational-implicature problem (requiring NL descriptions to be as short as possible; requiring network diagrams to contain certain perceptual organizations) result in computationally intractable solutions. This has been formally proven for FN (Reiter 1990b); a formal complexity-theoretic analysis has not yet been carried out for ANDD, but it seems likely that such an analysis would show that the idealized version of the diagram-instantiation task is NP-Hard.

Conclusion
Grice observed that conversational implicatures are not just a linguistic phenomenon, but that they can occur in all modes of human communication. People expect communications that are directed towards them to follow certain stylistic/pragmatic rules (e.g., NL descriptions should not contain irrelevant components; network diagrams should not contain spurious perceptual groupings), and a computer system that violates these rules may lead its users, in an attempt to explain the violations, to draw unwanted and incorrect implicatures. The task of avoiding unwanted implicatures is central for both the FN natural-language generation and the ANDD automated graphic-design systems, and it seems likely that it will be a fundamental problem for any system that engages in complex computer-human communication, regardless of the communication medium.

Acknowledgments
Many thanks to Cecile Balkanski, Mark Friedell, Barbara Grosz, Julia Hirschberg, Sandeep Kochhar, Candy Sidner, and Bill Woods for their help. The first author was partially supported by a contract from US West Advanced Technologies; the second author was partially supported by a National Science Foundation Graduate Fellowship, an IBM Graduate Fellowship, and a contract from US West Advanced Technologies. Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation, IBM, or US West Advanced Technologies.

References
Figure 1 The disk subsystem of a computer [after Herring & Prather (1986)].

Figure 2 A variation of Figure 1.

Figure 3 Another variation of Figure 1.

Figure 4 A network diagram designed by ANDD [after Sanden (1989), p.337].