Felicity Conditions

Performative speech is neither true nor false, as we've argued, but it can certainly "fail" in some sense of the word -- there's some sense in which a performative must be uttered under "appropriate circumstances." By way of example, to bet is not merely to utter the words "I bet …, etc.": someone might do that all right, and yet we might still not agree that he had in fact succeeded in "making a bet," or at least not entirely, succeeded in betting. To satisfy ourselves of this, we have only to to go to a horse race and, for example, announce our bet after the race is over. None of your interlocutors, you may be sure, will believe you to have succeeded in betting. Nonetheless, under the right circumstances, saying "I bet …, etc." is precisely what betting is.

Sometimes performative speech goes wrong and the intended act -- marrying, betting, bequeathing, christening, baptizing, etc. -- is therefore to some extent a failure. The utterance is then, we may say, not indeed false but in general *unhappy*. In trying to classify the ways in which things can go wrong we arrive at the *Doctrine of the Infelicities*. Or, as they are now know, felicity conditions.

(A. I) There must exist an accepted conventional procedure having a certain conventional effect, that procedure to

include the uttering of certain words by certain persons in certain circumstances, and further,
(A. 2) the particular persons and circumstances in a given case must be appropriate for the invocation of the particular procedure invoked.
(B. I) The procedure must be executed by all participants both correctly and
(B. 2) completely.
(Γ. I) Where, as often, the procedure is designed for use by persons having certain thoughts or feelings, or for the inauguration of certain consequential conduct on the part of any participant, then a person participating in and so invoking the procedure must in fact have those thoughts or feelings, and the participants must intend so to conduct themselves, and further
(Γ. 2) must actually so conduct themselves subsequently.

Indirect Speech Acts

[...]

Sometimes explicit evidence of our construal of indirect speech acts rises to the linguistic surface, and we get adjacency pairs like:

A: “I’d love to help.”

B: “Thanks for the **offer**.”

Or:

A: “I could eat the whole cake.”

B: “Thank you (for the **compliment**)!”
Such pairs might seem like the worst kind of non-sequitur to the unprepared artificial intelligence. Humans rarely, if ever, notice them.

This is because we are so good at inferring, divining, simulating the intentions of other human beings -- we are so deeply and fundamentally cooperative, in our linguistic behavior no less.

Gricean meaning-nn
[...]
Now once you’ve truly fathomed this, it should be an *earth-shattering* revelation to you -- that language doesn’t effect communication by my *encoding* my thoughts into linguistic symbols, and then *sending* those linguistic symbols across the language pipeline for you to *decode* back into my original message. This is the common-sense, folk notion of language. But this notion of Gricean *meaning-nn*, and all the subsequent advancements in cognitive linguistics and cognitive semantics, should finally and ultimately disabuse us of those folk notions. Language, the words I choose, rather give you the hearer some minimal, but sufficient clues for finding the domains and principles appropriate for reconstructing my own conceptualization. Once these clues are combined with

already existing configurations, available cognitive principles, context, background framing, the appropriate mental construction can then take place -- and the results far exceed any overt explicit information.

Gricean *meaning-nn* allows me to divorce entirely the literally semantic content of my utterance from my intention in uttering it -- and most of the time, speakers and hearers don't even notice this rift. We're quite blind to it, in fact. It took centuries for anyone to even notice (or at least consider it theoretically significant) that we routinely make requests by asking questions or making vague statements -- with sometimes several layers worth of indirection away from the implicit request -- and that this is consistent across *many human languages*.

Dialogue Act Annotation
ISO 24617-2 Standard is XML-based, has rich theoretical grounding in cognitive-linguistic and conversation analytical work in hidden structure of dialogues.
General-purpose communicative function: *inform, agree, disagree, correct, answer, confirm, disconfirm, question, offer, accept-offer, decline-offer, promise, request, suggest, instruct,* and many more.
Dimension-specific communicative functions: *AutoPositive, AutoNegative, AlloPositive, AlloNegative,*

FeedbackElicitation, Stalling, Pausing, TurnTake, TurnGrab, TurnAccept, TurnKeep, TurnGive, TurnRelease, SelfCorrection, SelfError, Retraction, Completion (of partner), CorrectMispeaking (of partner), InitGreeting, ReturnGreeting, Apology, Thanking, etc.
Across 8 dimensions: *AutoFeedback, AlloFeedback, Time Management, Turn Management, Own Communication Management, Partner Communciation Management, Social Obligation Management, Discourse Structuring.*

If enough and large enough corpora are so annotated, this scheme (and others like it) obviously have the potential to give rise to rich theoretical and practical (engineering) work in the field of dialogue, broadly construed.

Coming up: reasoning from context
The modern work on deixis really beginning with (John) Lyons (1977 monumental work on Semantics) and (Chuck) Fillmore.

The work on pragmatic implicature, conversational and conventional, really beginning with (Herbert Paul) Grice

The work on conceptual metaphor primarily from (George) Lakoff, (Mark) Johnson, (Gilles) Fauconnier, (Zoltán) Kövecses, etc.

The work on politeness theory primarily from (Penelope) Brown, (Stephen) Levinson, (Dan) Sperber, (Deirdre) Wilson (the last two a relevance-theoretic interpretation)

Deixis (**background**)
The concentric circles shift — take, e.g., "The book is over there [on that desk, in the classroom, in this building, at UMD]" *vs* "All the students are here by now [at the University of Maryland]." Clearly the speaker-proximal area is far smaller in the first sentence. In other words, the meaning of these deictic words and expressions is non-conventional, and approximately inferred.

Or I might shrink my zone of intimacy/familiarity in a language that grammatical uses these notions (German Sie/du, Spanish tú/usted, etc.) — for instance my child in one moment might be addressed by *du*, but in the next (supposing he or she'd done something to upset me, or I needed to discipline him or her) I might address him/her as *Sie*, shrinking my circle of intimacy for the purpose of driving home the disciplinary nature of my action/utterance.

Deixis of place

Here/there — “the pragmatically given space, proximal/distal to speaker’s location at CT(, *that includes the point or location gesturally indicated*).”
E.g.: “Place it here,” “Place it there.”

This/that — [glossed by Lyons 1977a: 647 as] “the object in a pragmatically given area close to/beyond the speaker’s location at CT”
E.g.: “Bring me that book,” “That’s very delicious.”

Empathetic deixis: the speaker can shift the deictic centre to that of the addressee/hearer to show emotional closeness or empathy, or perhaps to garner empathy or help one’s argument to go through by conflating one’s own thoughts with those of the hearer. Such a shift has to be recognized contextually; and appropriate inferences may be drawn therefrom.

Various languages discretize space along this deictic dimension differently; English has only two pragmatically given discrete locations, the NW Amerindian language Tlingit has four (*right here, nearby, over there, way over there*), and Malagasy has a six-way contrast along the same dimension.

Finally, much as **time** deixis is inherent in verb tense as much as in temporal adverbs and phrases, **place** deixis also plays a role in the meaning of certain verbs whose precise interpretation is pragmatically inferred from context.

"He's coming" — *'He is moving towards the speaker's location at CT'* (as opposed to "He's going")
"I'm coming" is different, though — what does it mean? Not *'the speaker is moving towards the location of the speaker,'* but that *'the speaker is moving towards the location of the* addressee *at CT.'*
This may have arisen diachronically from a polite empathetic deictic shift to the addressee's point of view.
So we adjust our gloss to *'movement towards either the location of the speaker, or towards the location of the addressee, at CT.'*
But what about:
"When I'm in the office, you can still *come* to see me."
Perhaps we should adjust our gloss to *'movement towards the location of the speaker, or politely the addressee, at either CT or perhaps the time of some other specified event — call it* **reference time**.*'*
But what about:
"I came over several times to visit you, but you were never there." Our analysis is clearly still incomplete, and we should perhaps adjust our gloss further by adding the

clause *'or motion towards the home-base maintained at CT, or at some reference equally salient in context, by either speaker or addressee.'*

And you should feel, rightly, that that is an awful lot of inference to be done by a human just in order to understand the intended interpretation of a single word in a rather simple sentence.

Is all language context-dependent?
Perhaps there are indeed no reliable conventional (or *timeless*, or *literal*) meanings to be found in human linguistic communication. In other words, perhaps there are no straightforward ways to convert a sentence to a truth-conventional proposition that an AI could rely on. Perhaps meaning is all contextually inferred, with that inference relying heavily on inference of speaker intention.

Examples (consider the problem of definite description).

This problem of definite description reduces really to the problem of "the," which in turn thrusts us back upon the problem of grounding, salience, referential inference, etc.

'The' ambiguous between deictic axes

But it still serves to indicate to the hearer — to assure them, in a sense — that the intended referent is *available*, even in some sense *salient* in the contextual space. Consider the utterance: "The cathedral was built by the Medicis." What cathedral? It depends on context, obviously; but note that now *how we search for the intended referent also depends on context*. If we're in Italy, we'll likely assume that the proper cathedral is deictically available along the spatial axis — perhaps it's '*the cathedral in this city we're currently in*.' If we're in a classroom in America, during a history class, and we fell asleep, we might assume that the intended referent is available along the discourse axis — perhaps it's '*the cathedral that was grounded a few utterances ago*.' Or if we just watched a movie that involved a cathedral, perhaps the cathedral is temporally proximal — '*the cathedral that we just saw a few minutes ago in the film*.'

In any case, we can agree that *the* indicates that the correct referent is *salient*, in the sense (now that we can formally articulate it) of *deictically relatively proximal*.

Salience examples: salience decays over time (if two women were grounded, probably more recent one in absence of more explicit description); more likely to refer to say, someone in the room or someone mentioned

recently than say, to suddenly refer to *Hitler* or *Mahatma Ghandi*, although he's certainly somewhere in the common ground (as a part of our collective cultural knowledge).

Frame Semantics
Cf. also the notion of *profiling* found in Ron Langacker's theory of cognitive grammar (CG), for which consult the eponymous *Cognitive Grammar*.

Now at first blush you're likely to think, "Well this is some kind of philosopher's utopia — no one talks like this." And certainly people don't always tell the truth, and sometimes we ramble on about irrelevancies, and sometimes we get confused and use more words than were strictly necessary to convey a thought.

But this isn't what Grice is talking about. That's all at a very superficial level. The fact of it is that even when we're at each other's throats verbally, we're cooperating in the most wonderful way.

Suppose I say something to you in the course of a conversation that's blatantly false — I violate the maxim of quality — "Queen Victoria was made of **solid steel**." All of a sudden a metaphor springs into existence — that is, you assume I'm not saying what you and I know to be false and

instantly find an interpretation that renders what I say adherent to the maxim of quality.

Or suppose I appear to say something completely uncooperative and irrelevant — you ask "What time is it?" And I say "Well the milkman just came." Strictly speaking, this is a non sequitur if we're to take the second sentence at literal face value. But you never even *consider* that I might not be cooperating. You find a set of inferences that connects the two sentences.

Or suppose we were British, and you say in the course of a political conversation: "What if Putin were to blockade the gulf and keep all the oil from us?" — "Oh come now, Britain rules the seas!" Of course that hasn't been true since perhaps before WWII, but you don't suppose I'm uninformed or lying — you try to find a nearby proposition that I might really be trying to get across, and you hit upon the exact negation of what I said — "Britain has no naval clout anymore; we're screwed if that happens."

Other quality flouting:
"Tehran's in Turkey, isn't it, teacher?"
"And London's in Armenia I suppose."

Some quantity flouting:

“War is war.” -> “terrible things always happen in war, that’s its nature, and it’s no good lamenting that particular disaster.”
“Either John will come or he won’t.”
“If he does it, he does it” -> “Calm down, there’s no point in worrying about whether’s going to come because there’s nothing we can do about it”

On their face value these are tautologies that would seem to be supremely uncooperative in that they convey no information at all. But in fact they can convey a great deal, because we automatically assume relevance and an appropriate quantity of information.

Some relevance flouting:
“Hey where’s Susan?”
“Well I saw a yellow VW in front of John’s house again.”
This is literally a made up sentence — but look at all the presuppositions that sprang automatically into place in your mind to make the two sentences not nonsequiturs — that (1) Susan drives a yellow VW, that (2) she often goes to John’s house, that (3) I’m speculating — I don’t really know for sure where she is, but I have a hunch.
I said none of that.

Most interestingly, look what happens when I flout manner, particularly the sub-maxim to 'be brief' and say no more than is necessary for the other person to understand:
"Miss Singer sang an area from Rigoletto."
"Miss Singer produced a series of sounds corresponding closely to the score of an aria from Rigoletto."
What if we find the second in a musical review rather than the first? Do we assume the writer just had a stroke and forgot about the stylistic niceties of his craft? No. We almost automatically infer that there was in fact some considerable difference between Miss Singer's performance and those to which the term singing is usually applied.

Examples of Observing Gricean Maxims
[[[PLEASE FILL ME IN]]]

Examples of Exploiting/Flouting Gricean Maxims

Conceptual Metaphor Theory: Introduction
We got a metaphor to go through last slide by exploiting the maxim of quality -- and subsequently had merely to trust that the addressee would try and preserve their assumption of cooperativity by trying to find a "nearby"

proposition (in a sense that we haven't time to define or formalize) that he/she could safely infer we meant.

In general, this led to theoreticians beginning to increasingly notice the (1) pervasiveness and (2) strangeness of NL metaphor. Combined with notions from cognitive psychology, in particular a series of studies and papers in favor of the **perceptual symbol system** view of cognition, and the concomittent assumption of the fundamental importance of **embodiment** to human cognitive processes and reasoning, these observations began to coalesce into an important field of linguistics known as conceptual metaphor theory.

We mention here briefly simply because it rounds out our picture of the **hardness** of a general regime of natural language processing techniques, and (for our purposes) completes our short-list of things that a dialogue system will have to be able to cope with.

Can you think of any other metaphor schemata?
*An argument is a **building**.*

We've got the *framework* for a *solid* argument.

If you don't *support* your argument with *solid* facts, the whole thing will *collapse*.

He's trying to *buttress* his argument with a lot of irrelevant facts, but it's still so *shaky* that it will easily *fall apart* under criticism.

With the *groundwork* you've got [*foundation* you've laid], you can *construct* a pretty *strong* argument.

Arguments can be *undermined* ...

An argument is a ***journey***.

So far, we haven't *covered much ground* [*terrain*].

This is a *roundabout* [*circuitous*] argument.

We need to *go into this further* in order to *see clearly* what's involved.

As we *go further into* the topic, we find ...

We have *come to a point* where we must *explore* the issues much more deeply.

Understanding is ***seeing*** [combines with the above to produce ...]

[**combine with journey metaphor**]

Having come this far, we can now *see* how Hegel went wrong.

We will now *show* [guides do this] that Green misinterpreted Kant's account of will.

We ought to *point out* [a guide does this] that no such proof has been *found*.

Dig further into his argument and you'll *discover* a great deal.

We can *see* this only if we *delve into* the issues.

Shallow arguments are practically worthless, since they don't *show* us very much.

[**combine with building metaphor**]

We can now *see* the *outline* of the argument.

If we *look* carefully at the *structure* of the argument …

[**combine with container metaphor**]

That is a remarkably *transparent* argument.

I didn't *see* that point *in* your argument.

Since your argument isn't very *clear*, I can't *see* what you're getting at.

Your argument has no *content* at all -- I can *see right through* it.

Embodied Construction Grammar (**ECG**)

This is of course work undertaken by (Nancy) Chang and (Benjamin) Bergen -- originally under the direction of Jerry Feldman at UC Berkeley. It builds on *construction* based approaches to syntax/semantics, which are due to Fillmore, Kay, Goldberg, Croft and a number of others -- this is a huge and (still) productive area of research that's crossed over into NLP via the algorithmic notion of a unification grammar.

You don't need to understand any of that unless you're genuinely interested and want to know more. The point for our purposes is that Chang and Bergen, along with Feldman's team, have written an ECG analyzer that (1) analyzes a sentence *in context* to produce a semantic specification of the sentence in terms of *grammatical constructions and semantic frames,* and then (2) uses the semantic specification it just produced to run a dynamic simulation using active embodied structures; the meaning of the utterance consists of the simulation and the inferences it produces.

There's a series of nice papers by Srini Narayanan (and various collaborators, including Chang and Bergen at various points) about her work at processing conceptual metaphor and understanding semantic frames using ECG (Embodied Construction Grammar) and the ECG tools. The work begins really with her dissertation in 1997.

Conversation Analysis

Turn-relevance places (TRPs): actually, predicting TRPs was a machine learning task at a recent SIGDIAL conference. In any case, how do humans select the next speaker at TRPs? It typically happens very smoothly, but how do humans pull this off?

Sequence expansion:
Request/acknowledgment is a common adjacency-pair type. (Others are question/answer, offer/acceptance, offer/refusal, statement/disagreement — these are all surprisingly stereotypical). Obviously, though, all talk doesn't proceed in two-utterance pairs. That being said, much of the structure of talk seems to come from **sequence expansions** that are almost equally stereotypical. The three main way to expand a sequence are with a **pre-expansion**, an **insert-expansion** and a **post-expansion**.

Let's take as our example the request/granting sequence:

C: "[At Starbucks] Can I get a scone, please?"
E: "Sure thing."

How might we expand this sequence? With a pre-sequence insertion as follows:

Greeting/greeting:
E: "Good morning, sir!"
C: "Good morning — can I get …"
Or vice-versa

Or with a summons/answer pre-sequence:

E: "Sir? How may I help you?"
C: "Oh hello — yes, can I get …"
Or vice-versa

An insertion sequence might look like:

> …
> C: “Can I get a scone?”
> E: “Vanilla or strawberry?”
> C: “Uhhhh — strawberry”
> E: “Sure thing, sir.”

If E were at this point to begin asking about methods of payment, or talking about the weather, we’d consider this the start of a new conversation unit, a new sequence. But there are ways to extend it without opening a new sequence — using a **post-sequence** expansion, like a farewell/farewell:

> …
> E: “Sure thing, sir.”
> C: “Have a nice day!”
> E: “You do the same.”

At this point, it’s probably impossible to start a new sequence (on E’s part) without some kind of pre-sequence like another summons/answer (“Hey — umm”, “Oh yes?”).

Or another very common pre-sequence opening that you’re probably very familiar with is: “**Are you doing anything tonight?**” The dreaded question. The whole

point of the exchange is what's coming next turn — some kind of dreaded invitation to go do something. In fact, we often answer "Yes" even when it's not true — and the only reason we do this is because we recognize this as a common, polite pre-expansion to the invitation/accept or invitation/reject sequence. We know what's coming.

Or consider the following sequence collapse, which only occurs because the speakers are subconsciously familiar with the structure of common sequences:

"Do you smoke?"

"Nah I left them [my cigarettes] at home."

He's not answering "nah" to whether or not he smokes — he's answering "nah" to the part of the sequence that was collapsed, the fpp (*first pair-part*) of the request sequence ("can I bum a cigarette?"). What's been collapsed is the answer to the pre-expansion context question, which serves no other purpose than to determine whether the request is likely to be felicitous, or is applicable, as well as the fpp (*first pair-part*) of the request sequence itself.