Unpublished Manuscript: Please contact author before citing.
This material was presented as a poster at SemDial 2015 (GoDial), Gothenburg, August 2015

How Questions and Answers Cohere

Mandy Simons
Carnegie Mellon University, Department of Philosophy
July 2015

Abstract

When a syntactically complete sentence is uttered in answer to a *wh*-question, the asserted content may go beyond what is compositionally derivable from the sentence uttered. This paper provides an account of this observation, arguing that the Direct Answer relation is semantically significant. When a discourse segment is construed as a direct answer, special update rules apply. The content derivable from the surface form of the answer is merged with the content of the question using a procedure to be specified. For the Direct Answer relation to be licensed, the merge procedure must result in assignment of a value to the *wh*-variable in the question, which is treated here as a forward looking anaphor. This constraint gives content to the notion of Direct Answerhood.

1 The basic observation

When a declarative sentence is uttered in response to a question, the asserted content may be richer than the compositionally derivable content for that sentence. Here are some illustrative examples:

(1) Q:What did Clara draw with her new pencil?
A: She drew a dragon.
Asserted content: Clara drew a dragon **with her new pencil**.

(2) Q: What's Jane wearing for the wedding?
A: She's wearing jeans and a t-shirt.
Asserted content: Jane is wearing jeans and a t-shirt **for the wedding**.

(3) Q: What's Harriet knitting for Henry?
A: She's knitting a scarf.
Asserted content: Harriet is knitting a scarf **for Henry**.

I make here the strong claim that in these cases, the richer content is asserted, and is not merely implied or implicated. The claim reflects the intuition that the A speaker is fully committed to the truth of the richer claim – as long as she is considered to be answering the question. The only way that the speaker of (1), for example, could evade commitment to the proposition that Clara drew a dragon *with her new pencil* (rather than the weaker claim that Clara drew a dragon, with no specification of the instrument) would be to make explicit that her assertion is not intended as an answer to the question (but perhaps instead as a potentially relevant non-answer). There is thus a clear connection between the observed enrichment and the existence of a question-answer relation between Q and A. More specifically, the enrichment relies on A being taken as a direct answer to the question, a notion which will gain substance in the course of the paper.

In this paper, I argue that the direct answer relation is semantically relevant. Specifically, I argue that there is a specialised update procedure for utterances construed as answers, and that it is this procedure which results in the enrichment seen in (1)-(3). The central idea is that an utterance which serves as a direct answer is not interpreted independently. Rather, contents of question and answer are merged in the interpretation procedure.

I spell out this account in DRT (Kamp 1981), enriched with notions from SDRT (Asher 1993, Lascarides & Asher 1993). While most of the discussion requires only the machinery of DRT, some of the additional resources of SDRT are required, specifically: (a) the mechanism for

distinguishing distinct discourse segments (via the assignment of labels to segments) and (b) procedures for inferring and representing discourse relations between segments. For the sake of perspicuity I will present the proposal (almost) entirely within the language of DRT. It will be important at various points, however, to recall that the DRT machinery is assumed to be embedded in the richer SDRT framework; and I will point out those places where these assumptions are necessary.

Articulation of the account requires that I provide an account of the representation of *wh*-questions in DRT. My goal is not, however, to offer any new account of the semantics of questions. The goal is to provide an account of the semantics of answers.

2 Preliminary assumptions: the discourse relation DirAns

Throughout this discussion, I will make the following assumption: When a question is asked and the next discourse move is an utterance of a declarative with assertoric intonation by the addressee of the question, then that utterance is assumed to be intended as a direct answer. The interpretation procedure therefore proceeds by specifying that the Direct Answer relation (*DirAns*) holds between the two segments.[1]

In SDRT, positing a particular discourse relation between two segments may introduce additional requirements on the interpretation of the segments. For example, if the relation *Narration* is established between two segments, S1 and S2, this imposes constraints on the temporal relations between events in the two segments. If these constraints cannot be met (perhaps because satisfying the constraints would introduce a contradiction), then the discourse relation posited between S1 and S2 must be revised. These additional requirements are expressed in *coherence constraints* associated with each discourse relation.

In this paper, I will propose coherence constraints on the DirAns relation which must be satisfied in order to maintain the assumption that this relation indeed holds between two segments.

3 Representing the content of questions

The first step is to specify a DRS representation for *wh*-questions (cf. Asher and Lascarides 1998). My treatment of questions is informed by the structured meaning approach (Krifka 2001), according to which questions are predicates. Part of the appeal of this approach is to highlight the tight semantic relation between questions and their answers, for on this view, answers provide the argument to which the question-predicate is applied to produce the asserted answer. The approach developed here can be seen as an attempt to articulate the questions-as-predicates view within a discourse-oriented framework, in a way which accommodates both short and long answers.

In developing a DRS construction rule for *wh*-questions, I assume that at some level of representation, a *wh*-question has the form shown in (4), where one or more *wh*-operators take scope over a sentence which contains *wh*-traces co-indexed with and grammatically linked to the wh-expressions.

(4) $[wh_1 ...[\ wh_n\ [_S ...x_1 ... x_n...]]$

I further assume that semantic type information (e.g. person for *who*, non-person for *what*, location for *where* etc.) is "left behind" as a feature on the variable.

The embedded S can then be treated using standard DRS construction rules. The only special rule that is required is for the *wh*-traces. These are treated by the rule in (5).

(5) **DRS construction rule for *wh*-traces** Given the syntactic configuration: $[_{XP}\ x_i$: $\varphi_1...\varphi_n\ ]$, ($\varphi_1...\varphi_n$ being the semantic type

[1] Asher and Lascarides 1998 propose a discourse relation *Question Answer Pair* (QAP) for similar purposes. They assign this relation a particular semantics, which I do not adopt for DirAns, hence I use a different term.

features derived from the *wh*-expression itself), where x_i is bound by a *wh*-expression:

i. introduce a new discourse referent x_i into the universe of the DRS under construction, and conditions $\varphi_1(x_i)...\varphi_n(x_i)$ into the set of conditions of that DRS.

ii. Then, add a condition of the form , $?x_i$ to the set of conditions of the DRS.

Having created a DRS for the content of the embedded S, we need to represent that this is the content of a question rather than asserted content. To do this, I use the notation shown in (6):

(6) λx…y […x, y…: …?x, ?y]

This structure is to be interpreted as a set of entities which have the properties characterized by the embedded DRS. This DRS is inert in the procedure for determining the truth of any larger DRS which contains it. It will makes its truth conditional contribution in the effect it has on the interpretation of utterances construed as direct answers.

More must be said regarding the new condition $?x_i$ introduced by part ii. of rule (5) above. This condition marks x_i as a *forward looking anaphor*. While a standard anaphor requires an antecedent in the existing DRS or at least in a superordinate position, a forward looking anaphor is an indication that some new predication containing information about the discourse referent in question is anticipated in the subsequent discourse. Generally, this will be accomplished by identifying the forward looking anaphor with some new discourse referent introduced in a later utterance. I propose that a central requirement for a segment to count as a direct answer to a question is that it provide a value for the forward looking anaphor in the question. I formulate this requirement as the *Answerhood Constraint*, proposed as a coherence constraint on *DirAns*.

Answerhood Constraint, version 1 Let π_i, π_j be discourse segments; let K_i be the DRS corresponding to π_i, where $?x_1 \in Con(K_i)$; and let K_j be the DRS corresponding to π_j. If DirAns(π_i, π_j) then the result of updating with K_j must provide a suitable value for x_i.

The remainder of the paper will involve specifying how this constraint is to be satisfied, and how a discourse referent in the answer is determined to be a suitable value for the forward looking anaphor.

4 First attempt: extracting short answers

The cases that are of interest to us are those in which *wh*-questions are given a full-sentence, or long, answer. However, such questions naturally invite a constituent, or short, answer, as illustrated in (1)' and (2)':

(1)' Q: What did Clara draw with her new pencil?
A: A dragon

(2)' Q: What is Jane wearing for the wedding?
A: A scarlet dress.

Cases like these provide a straightforward way to satisfy the Answerhood Constraint, as the response introduces a single discourse referent of the appropriate type to be identified with the forward looking anaphor of the question. Let's assume for current purposes that once the forward looking anaphor has been assigned a value, the λ-operator is deleted (as assigning a value to this anaphor is equivalent to applying the question-predicate to an argument), and the DRS originally in the scope of this operator comes to be asserted content. [2] The entire procedure is illustrated below for the case in (1)'. (DRSs are simplified for perspicuity.)

i. K_Q: λx_3 [e_1, x_1, x_2, x_3 : x_1=Clara, her-new-pencil(x_2), draw(e_1), Ag(e_1, x_1), Instr(e_1, x_2), Th(e_1, x_3), non-person(x_3), $?x_3$]

ii. K_A: [y : dragon(y)]

[2] We will be able to drop this assumption in the full version of the proposal laid out in section 5.

iii. DirAns(Q,A)

iv. K_{Q+A}: [e_1, x_1, x_2, x_3 : x_1=Clara, her-new-pencil(x_2), draw(e_1), Ag(e_1, x_1), Instr(e_1, x_2), Th(e_1, x_3), non-person(x_3), x_3=y]

The construction of the DRS for the question (step i) proceeds as specified in section 2. For the answer, I assume here that the initial DRS consists only of the representation of the NP content. At this point, we establish the DirAns relation between the two (step iii).[3] (This takes place in the construction of the SDRS, the details of which are omitted here.) At the next step, we must satisfy any constraints associated with the discourse relation introduced. In our case, this means attempting to satisfy the Answerhood Constraint. To do this, we must find a discourse referent in K_A which can be identified with x_3, the forward looking anaphor in the question. Only one discourse referent is available. We revise K_Q to K_{Q+A}. By identifying the theme argument of the event introduced in the question with the discourse referent introduced in the short answer, we guarantee that the structure including K_A and K_{Q+A} will entail that Clara drew a dragon with her new pencil, just as we want.

These observations about short answers lead naturally to the thought that a short answer is really all that a *wh*-question requires. When a *wh*-question is followed by a full sentence answer, perhaps we should think of everything which is not part of the answering constituent simply as a semantically redundant grammatical vehicle for that constituent. On this picture, the job of interpreting a sentence as a direct answer involves finding the short answer in the reply, and using it to satisfy the Answerhood Constraint.

But this idea cannot be quite right. First, long answers typically provide multiple discourse referents that could potentially provide a value for the forward looking anaphor, as illustrated by (7):

(7) Q: Who did Jane see?
λy [x, y, e_1 : x=Jane, see(e_1), Exp(e_1,x), Th(e_1,y), person(y), ?y]
A: Jane saw Bill.
[w, z, e_2 : w=Jane, z=Bill, see(e_2), Exp(e_2,w), Th(e_2,z)]

The DRS for the answer provides two referents which could be identified with y. Clearly, only one of those is actually "the answer." In the current framework, we can think of this as the problem of specifying precisely under what conditions a discourse referent is a "suitable value" for a forward looking anaphor in a question. One idea we might consider is that focus marking serves to identify the relevant discourse referent. (Question-answer congruence requires that the answer constituent in a full sentence answer carry semantic focus, which in English is typically indicated by prosody (Rooth 1991).)

However, prosodic marking of a constituent is not enough to guarantee answerhood, as shown by sequences like (8):

(8) Q: Who did Jane see?
A: Frankie loves [$_F$ Billie].

The presence of appropriate prosodic marking on an NP constituent does not render this sequence well formed as a question/answer pair. Obviously, the content of the answer must match the question. Identifying a constituent as "the answer" (or a discourse referent as a suitable value for the forward looking anaphor) requires there to be a parallel between the contents of question and answer.

A further problem is raised by examples like (9):

(9) Q:What did Clara draw with her new pencil?
A: In the morning, (she drew) a dragon, and in the afternoon, (she drew) a snake.

[3] This case is not covered by the assumption introduced earlier, that assertoric responses to questions are assumed to be direct answers. But it is natural to extend that assumption to cases where a *wh*-question is followed by utterance of a constituent of the same syntactic type as the *wh*-gap.

This shows that a direct answer, whether long or short, can contain additional information that is not contained in the question. We cannot arrive at the correct interpretation of the sequence in (9) simply by extracting "the answer" from the full sentence. On the other hand, our basic observation is that we cannot arrive at the correct interpretation simply by interpreting the answer in isolation. This suggests that what we need instead is a procedure that combines the contents of question and answer. We turn next to developing such a procedure.

5 Merge + Unification

There are several observations and intuitions we would like the account to capture. First: questions and their direct answers are a kind of compound discourse unit: an utterance construed as a direct answer should not be interpreted independently of its question. Second: for an utterance to count as a direct answer it must be construable as being, roughly, "about" the same thing as the question. In the sequence, *Who did Jane see yesterday? Jane saw Bill*, we understand that the seeing event mentioned in the answer is the same as the seeing event introduced in the question. It is this that allows us to infer that Bill is the theme of the event which the question asks about.

The first of these observations suggests that we should construct the content of answers by *merging* the representations of the linguistically expressed answer content with the content of the question. The second observation suggests that in this process of merge, we should seek to *unify* the content of the question with the content of the answer where possible [4] The final point suggests that the merge+unification procedure should guarantee an answer to the question. What follows is an attempt to formulate a procedure of this sort.

[4] This is in accord with the principle of Hobbs et al. 1993 to eliminate redundancies wherever possible.

5.1 Basic case

The procedure consists of two stages: merge followed by unification. The procedure is triggered by introduction into the SDRS of the condition DirAns(Q,A), where Q, A are discourse segments.

Merge
If DirAns(Q,A), then revise K(A) to K(Q+A) as follows:
i. $U(K(Q+A)) = U(K(Q)) \cup U(K(A))$
ii. $Con(K(Q+A)) = Con(K(Q)) \cup Con(K(A))$ [5]

Unify
$\forall x \in U(K(Q+A))$, if $\exists y \in U(K(Q+A))$ s.t. positing x=y does not lead to inconsistency, then add x=y to Con(K(Q+A)).[6]

Let's apply this procedure to a basic case, then consider some complications. We begin with example (10), a case where the assertion contains some content expressed explicitly in the question but not the answer, as well as content explicit in the answer but not the question.

(10) Q:What did Clara draw with her new pencil?
A: In the morning, she drew a dragon.

i. K_Q: λx_3 [e_1, x_1, x_2, x_3 : x_1=Clara, her-new-pencil(x_2), draw(e_1), Ag(e_1, x_1), Instr(e_1, x_2), Th(e_1, x_3), non-person(x_3), ?x_3]

ii. K_A: [e_2, y_1, y_2: female(y_1), y_1=?, draw(e_2), Ag(e_2, y_1), Th(e_2, y_2), dragon(y_2), $e_2 \subseteq$the-morning]

iii. Assume: DirAns(Q,A)

[5] Because this procedure involves adding the question content to the answer segment, the step of dropping the λ-binder from the question representation, introduced in section 4, is no longer required.

[6] In preparing this paper for submission, I came across a reference to work on unification in DRT by Kamp (2001). Unfortunately, I have not had time to study this paper.

iv. Merge: Revise K_A to K(Q+A):
[e_1, x_1, x_2, x_3, e_2, y_1, y_2: x_1=Clara, her-new-pencil(x_2), draw(e_1), Ag(e_1, x_1), Instr(e_1, x_2), Th(e_1, x_3), non-person(x_3), ?x_3, female(y_1), y_1=?, draw(e_2), Ag(e_2, y_1), Th(e_2, y_2), dragon(y_2), $e_2 \subseteq$the-morning]

v. Unify:
[e_1, x_1, x_2, x_3, e_2, y_1, y_2: x_1=Clara, her-new-pencil(x_2), draw(e_1), Ag(e_1, x_1), Instr(e_1, x_2), Th(e_1, x_3), non-person(x_3), female(y_1), draw(e_2), Ag(e_2, y_1), Th(e_2, y_2), $e_2 \subseteq$the-morning, dragon(y_2), $e_1=e_2$, $x_1=y_1$, $x_3=y_2$]

In this straightforward example, it is obvious that we can identify e_1 and e_2, as both are drawing events, and the information about the participants in the events is compatible. Having identified the two events, we must also identify the discourse referents corresponding to the participants. In particular, x_3, the forward looking anaphor introduced in the question, must be identified with y_2, the discourse referent introduced by *a dragon*. Hence, the Answerhood Constraint is satisfied, and the missing theme argument is provided. (As a by-product of unification, the anaphoric pronoun *she* in the answer is also resolved.)

Recall that I am treating questions as predicates, so that the question/answer relation should involve providing an argument for that predicate. The Merge+Unify procedure has this effect. In K(Q+A), the complex predicate represented by the DRS corresponding to the embedded S of the question is applied to the discourse referent y_2, which itself corresponds to the "answer constituent" in A.

5.2 Avoiding arbitrary solutions

In the above example, the identification of the forward looking anaphor with a particular discourse referent is necessitated by the overall pattern of unification of referents. This is a requirement for direct answerhood. To capture this requirement, I revise the Answerhood Constraint as follows:

Answerhood Constraint, revised Let π_i, π_j be discourse segments; let K_i be the DRS corresponding to π_i, where $?x_1 \in Con(K_i)$; and let K_j be the DRS corresponding to π_j. If DirAns(π_i, π_j), then application of Merge and Unify must necessitate x=y for some discourse referent $y \in U(K_{i+j})$

The following sequence shows why this revision is required.

(11) Q: Who ate a sandwich?
A: Jane answered a question.

Suppose the interpreter attempts to treat A as a direct answer to Q, and so merges the DRSs of the two segments. The event referents, however, cannot be identified, and neither can the referents corresponding to *a sandwich* and *a question*. But nothing prevents us from identifying the discourse referent corresponding to *Jane* with the forward looking anaphor corresponding to the agent of the eating event: it is perfectly consistent for Jane to be agent of both an eating event and an answering event.

However, while this identification of referents does not lead to any inconsistency, it is not necessitated by any other unification, and indeed failing to make that identification would be equally coherent. This reflects the fact that the explicit content of the A utterance provides no information which requires us to take Jane to be the agent of the eating event. Because this identification is not necessitated, it does not suffice to satisfy the Answerhood Constraint. By imposing this stronger constraint on the assignment of a value to the forward looking anaphor, we ensure that sequences like (11) cannot be treated as instances of DirAns.

It is worth pointing out, however, that the procedure allows for various kinds of restatements of question content in answers. Consider for example:

(12) Q: What did Jane eat?
A: She munched on some salad.

(13) Q: What did Cecily break?
A: She smashed my favorite vase.

An event of munching on something is a sub-type of eating event, and an event of smashing something is a sub-type of breaking event. This would allow the question-event and answer-event in these pairs to be identified, despite the change in verb.

A remaining problem is presented by sequences like the following:

(14) Q: Who saw Bill?
A: Jane saw someone.

In this case, unification of the events and their arguments is entirely consistent, and indeed once the events are unified, identification of the agents is necessitated. Nonetheless, A should not normally count as a direct answer to Q. A speaker who answers Q with A would not normally be claiming that Jane saw Bill (only suggesting that this is possible). It is also the case, though, that in this sequence, A would not be produced with neutral prosody, but perhaps with something like the following pattern:

(15) Jane saw someone.
H* L*+H L-H%

Plausibly, the additional pitch accent on *someone* marks it as new information, therefore not to be identified with any existing discourse referent. This would, in turn, prevent identification of the seeing events (as the theme of the answer event and the theme of the question event are distinct). There would then be no necessity to identify the agents of the events, and hence the Answerhood Constraint will not be satisfied.

5.3 Multiple and Plural Answers

Issues of plurality and quantification in answers raise some complications which cannot be dealt with in any depth at this point, but should nonetheless be noted. The central issue is that *wh*-questions are unspecified for number, as illustrated by the possibility of the answer in (16):

(16) Q: Who ate a sandwich for lunch?
A: Jane ate a sandwich and Lucy ate a sandwich.

While the question introduces only one event, the answer introduces two, and these must be distinct from one another.

A natural way to treat this case is to assume that each conjunct constitutes a distinct discourse segment, each of which is related to the question by the DirAns relation. Hence, Merge + Unify will be carried out separately for each conjunct (guaranteeing that both eating events were “for lunch”). Here, it is important to remember that the discourse referents introduced in the DRS of the question are simply part of a predicate, which can of course be predicated of multiple arguments. Identifying the question event with the event of the first conjunct and then with the event of the second conjunct does not entail identity of the events in the two conjuncts, but only provides a way to ensure that all the properties attributed to the event referent in the question are attributed to the event referent in each conjunct of the answer.

Similarly, questions may be given quantified answers, as in (17):

(17) Q: Who got a pencil from the teacher?
A: Every math student got a pencil.

This case requires a somewhat more complicated version of the Merge + Unify rules, which I will not pursue here.

Some cases of plural answers hide a great deal of complexity in the logical form of the proposition expressed, as in (18):

(18) Q: Who ate a sandwich for lunch?
A: Jane and Lucy (each) ate a sandwich.

The distributive semantics required here will guarantee that there are two distinct events of eating, each with a (different) sandwich as theme.

Unification should proceed just as in example (16). However, there are other cases which involve single events with multiple entities fulfilling a particular role in the event. Consider for example:

(19) Q: Who did Jane drive to school today?
A: She drove Cecily and Dave.

We might plausibly analyze (19)A as introducing a single driving event, with both Cecily and Dave as themes. After Merge, U(K(Q+A)) will contain at least three discourse referents: one for Cecily, one for Dave, and one for the theme argument of the driving event in the question. Exactly how unification should proceed here requires some further investigation. But it is also unclear whether the analysis suggested is correct. An alternative would be to treat the NP *Cecily and Dave* as introducing a complex individual which serves as the theme argument of the event. Clearly, this brings us into complex issues pertaining to the interpretation of plurals and the analysis of events, which are tangential to our concerns here.

6 Conclusion

The goal of this paper was to account for the observation that what is asserted by an answer to a question may include content not explicit in the sentence uttered. I have argued that this content results from a merging of the content of question and answer, which is triggered by the assumption that the discourse relation DirAns holds between them. I have argued further that in order to maintain this assumption, the process of merging the contents of question and answer must necessitate the identification of the *wh*-variable (represented in the DRS by a forward looking anaphor) with a discourse referent in the question: in effect, the merge must necessitate a particular answer to the question.

Although the focus of this paper has been on full sentential answers to *wh*-questions, I have also touched briefly on the treatment of short (constituent) answers. The proposal sketched in section 4 suggests how such answers can come to convey full propositional content. More work is required to fully flesh out that proposal

It should be noted that the treatment of short answers has an important application in a set of cases not discussed at all so far: full sentence answers to *how* and *why* questions. As examples (20) and (21) show, the asserted content of these answers is enriched:

(20) Q: How is John getting to Chicago?
A: He's taking the train.
Asserted content: John is getting to Chicago by taking the train.

(21) Q: Why is John going to Chicago by train?
A: He's afraid to fly.
Asserted content: John is going to Chicago by train because he's afraid to fly.

Although they are syntactically complete sentences, the answers in (20) and (21) are actually *short* answers: they do no more than provide the value for the forward looking anaphor in the questions. I leave further elaboration of these cases for future work.

Finally, this paper provides an account of the notion of a direct answer for a *wh*-question. A direct answer is an utterance whose interpretation results in satisfaction of the Answerhood Constraint (revised). No stipulations concerning either the form of the utterance or the logical form of its content are required.

References

Asher, Nicholas. 1993. *Reference To Abstract Objects in Discourse*. Kluwer Academic Publishers, Dordrecht.

Asher, Nicholas and Alex Lascarides. 1998. Questions in Dialogue. *Linguistics and Philosophy*, 21(3): 237-309

Hobbs, Jerry R., Stickel, M. E., Appelt, D. E, and Martin, P. 1993. Interpretation as abduction. *Artificial Intelligence* 63: 69-142.

Kamp, Hans 1981. A theory of truth and semantic representation. In J. Groenendijk, TH. Janssen and M. Stokhof (eds.) *Formal Methods in the Study of Language*, Part 1. Mathematisch Centrum. Amsterdam: 277-322.

Kamp, Hans. 2001. The Importance of Presupposition. In C. Rohrer, A. Rossdeutscher and H. Kamp (eds), *Linguistic Form and its Computation*. CSLI, Stanford: 207–254.

Krifka, Manfred 2001. For a structured meaning account of questions and answers. C. Féry and W. Sternefeld (eds), *Audiatur vox sapientiae: A festschrift for Arnim von Stechow.*Akademie Verlag, Berlin: 287-319.

Lascarides, Alex and Nicholas Asher. 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy* 16, pp. 437-93

Rooth, Mats. 1992. A Theory of Focus Interpretation. Natural Language Semantics 1: 75–116.