Modality, the Progressive, and Discourse Structure

Eric McCready

Aoyama Gakuin University

July 21, 2008
This talk is concerned with the interaction of progressive aspect and modality in the context of verbs of creation.

- Imperfective paradox: existential entailments associated with indefinites fail in the context of verbs of creation (Dowty, 1979).

- (1) does not entail that any object exists that is actually a house, but merely that some object exists that, if the building process is not interrupted, will eventually become a house.

(1) John was building a house.
McCready (2006):

- What happens with anaphoric dependencies to indefinite objects of verbs of creation in the progressive like that in (1), in which an anaphoric element appears in a subsequent sentence?

- Anaphoric elements mean the same as the possible denotations of *a house* in (1) in most cases;
  - they can mean either the partially completed physical object that will eventually be a house,
  - or the abstract object the building event is intended to realize.

(2) John is building a house. It is next to a school. (concrete)

(3) John is building a house. He designed it himself. (abstract/plan)
The focus of today’s talk is a different case from that paper:

- The anaphor can ‘refer’ to ordinary complete houses.
  - These are cases in which the anaphor is in the scope of a futurate modal operator such as *might*, as in (2).

(4) John is building a house. He might let his sister live in it.

- *Might* has a futurate interpretation on which the event in its scope is understood as taking place at a time later than the interpretation time of the modal sentence itself (Condoravdi 2002).
- Evaluation time of the content in the scope of the modal can be understood as later than the time at which the house-building is completed.
- This means that *it* can be interpreted as meaning the complete house that is the result of the building.
But the presence of a modal in the sentence containing the anaphor does not always induce this sort of interpretation.

▶ In (5), *it* clearly indicates the partial house, not the completed one.

▶ Why does this happen, given the presence of *might*, which we have already seen moves reference time forward?

(5) John is building a house. He might use high quality shingles when he puts the roof on it.
The answer, according to McCready, lies in world knowledge about how building events work.

- People usually live in completed houses, and it is difficult or impossible to put a roof on a completed house.
- Futurate modals make available futurate readings, but have nothing to say about whether reference time is advanced to a point after the end of the building event.
  - Two options are thus available:
    - advancement to a point before the end of the creation, and
    - advancement to a point after its culmination.

- Which is selected is determined by world knowledge; the pragmatically best interpretation is selected, which seems intuitively right.

Other languages in which modals have this futurate character, such as Japanese, behave in much the same way.
This informal analysis above seems essentially right.

- But there are problems with the formal treatment presented by McCready within Segmented Discourse Representation Theory (SDRT; Asher and Lascarides 2003). Ultra crash course:

  - SDRT is a dynamic semantic theory of discourse structure.
  - Each utterance (or clause) introduces a *speech act discourse referent*.
  - These referents are connected by binary discourse relations, which are inferred in a system of nonmonotonic logic based on a ‘knowledge base’ of world-knowledge related axioms.
  - The relations also have ‘hard’ semantic consequences.
McCready’s analysis of the progressive facts essentially has two parts.

- The first is the content of the modal operator
  - This introduces a) the modalization and b) the temporal contribution of modals

- The second is the contribution of world knowledge/contextual factors
  - This allows distinguishing situations in which we want to shift before and after the culmination point
The details:

- First, (existential, epistemic) modality is defined as a discourse relation: $\mathcal{R}_{\Diamond}$.

- Effect: the second argument of $\mathcal{R}_{\Diamond}$ (the content understood as in the scope of the modal) is semantically modified by a possibility modal.

This is the modal part of the content.
Why do this?

- The main reason:
  - as will become clear, the logic in which discourse relations are computed needs access to the content in the scope of the modal
  - Factoring the modality out into a discourse relation makes this content available as a separate discourse constituent.
  - The logic can access this constituent and the needed discourse relations can be computed.
Second, the evaluation time of the content in the scope of the modal is moved forward in a manner analogous to that proposed by Condoravdi (2002).

- The content \( t \leq t' \land AT(t', \varphi) \) is added to the asserted content. Definition (Condoravdi, 2002):

\[
AT(t, \varphi) = \begin{cases} 
\exists e [\varphi(e) \land \tau(e) \subseteq t] & \text{if } \varphi \text{ is eventive} \\
\exists e [\varphi(e) \land \tau(e) \circ t] & \text{if } \varphi \text{ is stative}
\end{cases}
\]

According to \( AT \),

- the eventuality described by modalized stative sentences must overlap with the reference time,
- while the eventuality described by event-denoting sentences must be included in the reference time.
However, such inclusion is impossible if $t'$ denotes the instant at which the sentence is uttered;

- for both $t \leq t'$ and $\tau(e) \subseteq t'$ to hold,
  - $t'$ must be interpreted as an interval beginning at the reference time and extending into the future for a distance large enough to hold the entire runtime of the event $\varphi(e)$.

The two together have the desired result of modalizing the modalized content to shift in time while still allowing discourse relations access to it.
The second part of the analysis introduces world knowledge into the picture via SDRT discourse relations.

- McCready makes use of two discourse relations in this context: *Narration* and *Elaboration*.
- Semantic consequences for connected elements:
  - *Narration* requires that the two events connected occur in sequence, without temporal overlap
  - *Elaboration* requires that the two events overlap in time.
- Which relation is inferred depends on the content of the sentences (via axioms in the nonmonotonic logic).

As may be clear, discourses like (2) are analyzed as Narrations, while discourses like (3) are Elaborations. This gives the correct semantic results.
The problem lies in the combination of these two elements.

- As noted (in a brief footnote) by Asher and McCready (2007), the theory results in unwelcome consequences.
- Connecting two segments with both $R\Diamond$ and either Narration or Elaboration results in a veridicality clash.
  - Both Narration and Elaboration are veridical relations: when they connect two segments, the content of each segment is entailed to be true.
  - However, this results in obvious problems when combined with $R\Diamond$.
- The entire point of the modal relation is that it does not require truth of its second argument.
We thus have a situation where

- the content in the scope of the modal is entailed to be true by one relation
- and not entailed by the other.

This is bad/incoherent. What to do about it?

- The rest of the talk is devoted to exploring options.
- The conclusion will be that the discourse referents themselves should be part of the modal content in some cases.
There are multiple possible solutions to this problem.

► Here I will explore two general strategies.

1. The first involves retaining $R \Diamond$ and finding ways around the veridicality clash.
2. The second requires moving to a more traditional conception of modals as semantic operators rather than as discourse relations.

I believe the second option is to be preferred.
Option 1: Veridicality Rankings. The first possibility is to keep the idea of ‘modality as discourse relation,’ but to make changes in the metatheory of discourse relations.

- What is needed is a way to ensure that the second argument of $R$ is nonveridical.
- The most obvious is to impose a kind of preference ranking on discourse relations when multiple relations connect two discourse segments.
  - The idea: the less veridical relation determines veridicality in such situations, so if one relation is nonveridical and one is veridical, the argument is understood nonveridically.
  - Antiveridical relations like Correction, which entail falsity of their first arguments, would always take precedence over the other types.

This immediately solves the problem.
But is this really the best solution?

- It appears stipulative.
- It also introduces complications into the metatheory which, if not strictly necessary, would be better omitted.
- Further, this route works to save an analysis which is nonstandard in two distinct ways.
  1. First, unlike nearly every other theory on the market, it treats modality as discourse relational rather than as an operator.
     - The move was inspired partly by Frank (1997), but is motivated by the need to access submodal content.
     - If there is another way, we might take it.
  2. Second, and more importantly, it assumes that two relations can connect the same two segments.
     - We have seen that doing so causes problems and should be avoided if possible.
There is actually a worse problem here . . . the proposed analysis simply *does not work*.

We could realize this analysis in two ways.

1. In cases of veridicality clash, the more veridical relation is simply cancelled.

2. Or: the veridicality is cancelled, but the semantic effects of the relation hold.

The first option will not work.

- Without the discourse relations holding, the analysis falls apart: it is the temporal effects of the relation that distinguish between
  - cases where the modalized event happens after the conclusion of the creation,
  - and those where the event time precedes it.

If we remove the discourse relations completely, this generalization is lost. Option 1 fails.
On the other hand, what happens if we cancel only veridicality?

- We now have the following situation:
  - *Narration*(α, β) and *R♦*(α, β).
  - Because *R♦* is nonveridical, it ‘wins’ and β is no longer entailed.

- But what does this mean for the content of the *Narration*? How can a possibly nonexistent event stand in a narrative relation to anything?

- More formally, *Narration*(α, β) entails that τ(me(α)) < τ(me(β)).
  - But τ(me(β)) does not necessarily have any runtime.

This is nonsense. Let us look elsewhere.
Option 2: Modals as Dual Speech Acts. The following conditions must be satisfied for a successful analysis.

- We need a way to represent ‘hard’ temporal information
- We need a way to allow information from discourse relations to enter the picture

We saw above that both were needed in order to get the right interpretations for discourses like (2) and (3).

- This means that it is necessary for the nonmonotonic logic to have access to the content in the scope of the modal.
The McCready analysis:

- $\mathcal{R}_\Diamond$: content of the modal ‘factored out’ into a discourse relation.

But there is another possible route:

- We can give the content in the scope of the modal a distinct discourse referent of its own, which can be accessed directly by the nonmonotonic logic.

Then the modal itself can be given a more standard operator treatment.
This is the route we will now follow.
I spell out this treatment using the dynamic semantics for modals proposed by Asher and McCready (2007).

- On this semantics, information states are sets of 4-tuples: the standard world-assignment pairs, and also $F, G$, sets of epistemic possibilities (which take the same form as information states).

- As information is processed, it updates worlds and assignments as usual, and also alters the global set of epistemic possibilities $G$ to reflect the new information.

- Modal information serves to update $F$, the set of focused possibilities.
The definition of *Might* is as follows:

\[ \sigma[Might \phi] \mathcal{A} \sigma(\frac{G_\sigma \setminus M}{G_\sigma \setminus F'}), \]

where:

- \( F' \) is a \( \phi \) descendant of the maximal set \( M \) of elements in \( F_\sigma \) that have \( \phi \) descendants, and \( G' = (G_\sigma \setminus M) \cup F' \), provided there are \( \phi \) descendants in \( F_\sigma \).
- If not, then \( F' \) is a \( \phi \) descendant of the maximal set \( M \) of elements in \( G_\sigma \) that have \( \phi \) descendants, and \( G' = (G_\sigma \setminus M) \cup F' \), provided there are \( \phi \) descendants in \( G_\sigma \).

\[ \sigma[Might \phi] \mathcal{A} 0 \text{ otherwise.} \]

where a \( \phi \)-descendent is defined as

\[ \varepsilon' \text{ is a } \phi \text{ descendant of } \varepsilon \text{ iff } \forall \sigma \in \varepsilon \exists \sigma' \in \varepsilon' \sigma[\phi] \sigma' \text{ and } \forall \sigma' \in \varepsilon' \exists \sigma \in \varepsilon \sigma[\phi] \sigma'. \]
This choice of theory is made for concreteness.

- However, the Condoravdi observations about tense must be added. This is easily done by adding the $AT$ predicate to the semantics:

\[ \sigma[Might \phi] \sigma(\frac{G_\sigma}{G'} \frac{F_\sigma}{F'}) \text{, where:} \]

- $F'$ is a $\phi$ descendant of the maximal set $M$ of elements in $F_\sigma$ that have
  \[ \phi \land \exists t'[t \leq t' \land AT(t', \phi)] \]
- descendants ($t$ the reference time), and $G' = (G_\sigma - M) \cup F'$,
  provided there are $\phi$ descendants in $F_\sigma$.
- If not, then $F'$ is a $\phi$ descendant of the maximal set $M$ of elements in $G_\sigma$
  that have $\phi$ descendants, and $G' = (G_\sigma - M) \cup F'$,
  provided there are $\phi$ descendants in $G_\sigma$.

\[ \sigma[Might \phi] \sigma(\emptyset) \text{ otherwise.} \]
We now need access to a speech act referent for the content in the scope of the modal that can be used to derive discourse-level temporal interpretation.

- I therefore assume that the content $\varphi$ in the scope of the modal—$Might \varphi$—is associated with a referent of its own.

- At the DRS-level we therefore have

\[
\begin{array}{c}
\pi \\
\pi : Might \\
\pi' \\
\pi' : \varphi
\end{array}
\]
How will this work?

(6) John is building a house. He might use high quality shingles when he puts the roof on it.

(7) $\phi; \text{Might}(\psi)$

$\pi, \pi'$

$\pi : \phi$

$\pi' : \text{Might}$

$\pi'' : \psi$

Elaboration $(\pi, \pi'')$

$\rightarrow$ The possible roof-putting temporally overlaps the building, as desired.
There is an issue with the preceding representation that becomes more obvious with the other kind of example.

(8) John is building a house. He might let his sister live in it.

\[ \pi, \pi' \]
\[ \pi : \phi \]
\[ \pi' : \text{Might} \]
\[ \pi'' : \psi \]

Narration (\(\pi, \pi''\))

\[ \Rightarrow \text{The possible living-in follows the end of the actual building.} \]
This is plainly problematic.

- The progressive does not require completion of the event.
- Thus the previous representation could mean that the sister might be allowed to live in the partly completed house.
- This interpretation might be available—but it is not the only one.
- The same holds for the discourse before it: the actual event might never reach the stage of shingle-putting.

How to solve this problem?
For an answer, we must look deeper into the semantics of the first sentence.

\[ \phi = \text{Prog}(e, \psi) \]

- Suppose we use the following axiom (‘\( > \)’ a normality conditional):

\[ \text{Prog}(e, \psi) > \exists e' \sqsupset e[\lambda e. \psi(e)] \]

- Then the sentence can be read as follows:
  - There is a partial event of housebuilding and in normal worlds this event is completed.
We really, then, have this:

\[ \text{Prog}(e, \phi); \text{Might}\psi \]

where the event described by \( \phi \) is realized in incomplete form as \( e \) and in completed form in normal worlds.

Note the modal character to this last definition. This modal character means that we should expect what we find.
Modal subordination: situation where an object in the scope of a modal can access an object in the scope of a previous modal.

(9) A wolf might come in. # It is big.
(10) A wolf might come in. It might be big.
(11) A wolf came in. It might be big.

Given that the progressive has a modal semantics, it is not a surprise that the modalized content can be accessed by content in the scope of another modal.

This follows on the assumption that ‘normal’ possibilities are accessible to epistemic ones, which looks reasonable enough.
The picture then becomes the following:

- The progressive has a modal element; the progressivized event is completed in normal worlds
- The modal allows access to future times
  - when the event is in progress
  - and when it is complete
- Discourse relations distinguish between the two by introducing temporal relations
  - overlap for Elaboration, and precedence for Narration

possible because of the presence of a ‘speech act’ referent in the scope of the modal

These observations clean up the picture and eliminate the need for \( R_\Diamond \), solving the problem.


