AN INTENTION-BASED SEMANTICS FOR IMPERATIVES

Daniel W. Harris
Hunter College, CUNY
(1) Do the right thing
(2) Snow is white
(3) Fly me to the moon and let me play among the stars
(4) Make us omelettes or I'll get us some bagels
(5) Help me if you can
\[\text{Snow is white}] = \text{???}
\text{DECLARATIVE}

\[\text{Do the right thing}] = \text{???}
\text{IMPERATIVE}
TWO ASSUMPTIONS

1. Clauses factor, at LF, into a mood-marker and a moodless sentence radical.
TWO ASSUMPTIONS

1. Clauses factor, at LF, into a mood-marker and a moodless sentence radical.
TWO ASSUMPTIONS

2. The semantic value of a sentence radical is a set of possible worlds.

\[(\forall \varphi) [\varphi] \in W\]
POSITIVE VIEW: A SKETCH

cf. Charlow (2014)
For any sentence radical $\varphi$:

$[[\top \varphi]]^c$ is a belief
(namely: the belief that $[[\varphi]]^c$ is true)

$[[\Diamond \varphi]]^c$ is an intention
(namely: the intention to make $[[\varphi]]^c$ true)
POSITIVE VIEW: A SKETCH

\[\lbrack \text{snow is white}\rbrack^C = \]

The belief that snow is white.
POSITIVE VIEW: A SKETCH

$[[\text{buy me a drink}]]^c =$

the intention to buy speaker$_c$ a drink
How to formalize this?
POSITIVE VIEW: A SKETCH

M_A determines:  
(in a way to be explained)

B_A  
A’s BELIEF STATE  
= the set of worlds compatible with what A believes

M_A  
A’s COGNITIVE MODEL  
= a set-theoretic representation of A’s beliefs and plans

I_A  
A’s INTENTION STATE  
= the set of worlds compatible with what A intends
Beliefs and intentions are formalized as properties of cognitive models:

Belief that dogs are better than cats:
$$\lambda M_A . B_A \subseteq \{ w : \text{dogs are better than cats at } w \}$$

Intention to high-five Beyoncé:
$$\lambda M_A . I_A \subseteq \{ w : \text{A high-fives Beyonce at } w \}$$
So, the semantic values of clauses are properties of cognitive models too.

For any sentence radical \( \varphi \),

\[
\models_c \varphi = \lambda M_A . B_A \subseteq [\varphi]_c
\]

\[
\models_c \neg \varphi = \lambda M_A . I_A \subseteq [\varphi]_c
\]
ALTERNATIVE TREATMENTS
FIRST ALTERNATIVE:

**STATIC SEMANTICS, DYNAMIC PRAGMATICS**

(Portner; von Fintel & Iatradou; Roberts)

\[ \text{[Snow is white.]} = \lambda w_{st} . \text{snow is white in } w \]

(The proposition that snow is white.)

\[ \text{[Do the right thing!]} = \lambda w_{st} . \lambda x_{e} : x = \alpha_{c} . \text{x does the right thing in } w \]

(A property (restricted to the addressee) of doing the right thing.)
(We’re about to go into the bar together. I say:) 

Buy us drinks and I’ll find a table.

NOTE: 

• Needn’t have a conditional meaning.
• Can mean roughly: ‘I’ll find a table. Buy me a drink.’
• Can be the consequent of a conditional: ‘If your friend is tending bar, buy us drinks and I’ll find a table.’
PROBLEM FOR STATIC VIEWS:

MIXED COORDINATION

(We’re at a book store. Each of us has three books, but we only have enough money for five, total:)

Put back Naked Lunch or I’ll put back Waverley.

(Starr ms)

NOTE:

• Needn’t have a conditional meaning.

• Can be the consequent of a conditional: ‘If we only have $5, put back Naked Lunch or I’ll put back Waverley.’
Buy me a drink.
You won’t buy me a drink unless you go to the bar.
→So, go to the bar!

Attack if the weather is good.
The weather is good.
→So, attack!
SECOND ALTERNATIVE: DYNAMIC SEMANTICS
(e.g., Starr)

Starr’s clauses:

\[ R[\top \varphi] = \{ r_\varphi \mid r \in R \& r_\varphi \neq \emptyset \} \]

(where \( r_\varphi = \{ \langle a[\varphi], a'[\varphi] \rangle \mid \langle a, a' \rangle \in r \& a[\varphi] \neq \emptyset \} \))

(A CCP that adds the content of \( \varphi \) to the context’s information.)

\[ R[\not\top \varphi] = \{ r \cup \{ \langle c_r[\varphi], c_r-c_r[\varphi] \rangle \} \mid r \in R \} \]

(A CCP that adds a preference for \( [\varphi]^c \) over \( [\text{not } \varphi]^c \) to the context)
“Stop what you’re doing.”

“We’ve been all wrong about this.”
“...it would be a mistake to augment the theory of assertion with the fruits of the epistemological literature on belief-revision.”

—N. Charlow (2014: §5.6.2)
Solution:
Semantics specifies properties of cognitive models, but leaves it up to a theory of non-monotonic reasoning to sort out how addressees should satisfy those properties on particular occasions.
SECOND PROBLEM FOR DYNAMIC VIEWS:
BAD THEORY OF SPEECH ACTS
context
STALNAKER (1978, 2014):
To **assert** \( q \) is to propose adding \( p \) to the Common Ground (CG).

ROBERTS (1996/2012):
To **ask** \( q \) is to proffer \( q \), intending that it be adopted as the new Question Under Discussion (QUD).

PORTNER (2004):
To **direct A to \( \varphi \)** is to propose that \( \varphi \) be added to their section of the conversation’s To-Do List (TDL).
A proposition $p$ is common ground of a conversation iff the participants commonly accept $p$:

- each accepts $p$;
- each accepts that each accepts $p$;
- etc.

(Stalnaker 2014)
A Pragmatic Argument

Roughly: We regularly perform speech acts and successfully communicate, in situations where we can’t, and can’t expect to, change the common ground.
The Coordinated Attack Problem
(The Byzantine Generals Problem)
Messenger
Dear General B,

The attack will be at dawn tomorrow.

Please confirm.

with love, General A.
Dear General A,
I got your message. The attack will be at dawn.
Please confirm.

your best bro, General B.
Dear General B,
Got it. I love the smell of battle in the morning.
Please confirm.
bros 4 life, General A.
Dear General A,
Roger. Lock and load.
Please confirm.
bro grabs, General B.
THEOREM

Given reasonable assumptions about the generals’ utility functions and epistemic standards, they will never achieve common knowledge or common belief. (Akkoyunlu et al., 1975; Gray, 1978; Halpern and Moses, 1990; Yemini and Cohen, 1979)

A (PRETTY CLEAR) COROLLARY

They won’t achieve common acceptance, either.
Dear General B,

I’ve been reading some theoretical computer science papers, and it turns out that this is never going to work.

Anyway, my men have come down with cholera. Do you know the cure?

kisses, General A
Dear General A,
Shame about the attack.
Wash your hands and don’t eat so close to the latrines.
💙💙💙, General B.
Last Will and Testament
(Rubenstein 1989; Binmore 1998)
CONCLUSIONS

Successful communication doesn’t require changing the context, if the context is built out of common (or even shared) attitudes.

Performing a speech act doesn’t require intending or proposing to update the context, either.
CONCLUSIONS

Context change can result from communication only in certain special circumstances.

Which circumstances?

When the speaker and addressee are in a shared situation (Schiffer 1972; Clark & Marshall 1981).
It’s 3:00.
SHARED SITUATIONS

S said $\varphi$. 
Solution:

Adopt a (slightly) different theory of speech acts.

My preferred option is Grice’s original view.
INTENTION-BASED SEMANTICS

cf. Grice, Strawson, Schiffer, Bach & Harnish, Neale
INTEENTION-BASED SEMANTICS

A PRAGMATIC VIEW
To perform a speech act is to produce an utterance with an addressee-directed communicative intention.

A METASEMANTIC VIEW
The semantic properties of expression types are to be explained in terms of the psychological states involved in using them to perform speech acts.
MEANING AND INTENDING

By doing something, x, S, \textbf{MEANT} something iff, for some audience, A, and response \textbf{R}, S did x intending

(i) A to to have a certain response \textbf{R}

(ii) A to recognise that S did x intending (1)
MEANING AND INTENDING

By doing something, x, S, \textbf{MEANT} something iff, for some audience, A, and response \textbf{R}, S did x intending

(i) A to have a certain response \textbf{R}
(ii) A to recognise that S did x intending (1)

Different kinds of speech act are aimed at different kinds of responses

i.e., different values for \textbf{R}
MEANING AND INTENDING

By doing something, x, S, **MEANT** something iff, for some audience, A, and response R, S did x intending

(i) A to have a certain response R
(ii) A to recognise that S did x intending (1)

To assert p is to communicatively intend for one’s addressee to form a belief that p.

\[ R = \lambda M_A . B_A \subseteq p \]
MEANING AND INTENDING

By doing something, x, S, MEANT something iff, for some audience, A, and response R, S did x intending

(i) A to to have a certain response R
(ii) A to recognise that S did x intending (1)

To direct someone to $\psi$ is to communicatively intend for them to form an intention to $\psi$.

$$R = \lambda M_A . I_A \subseteq \{ w : A \psi s at w \}$$
...is the study of a component of the mind that computes partial and defeasible evidence about what speakers intend by their utterances.

Specifically, it computes $R$ values (or at least properties of $R$ values).
A clause $\Phi$ has $[\Phi]$ as its semantic value for a speaker $S$ in virtue of the fact that:

a. If $S$ were to communicatively intend to produce $[\Phi]$ in an addressee, they might utter an unembedded token of $\Phi$.

b. If $S$ were to perceive an unembedded utterance of $\Phi$, they would conclude that, if the speaker is being literal and direct (etc.), the speaker intends to produce $[\Phi]$ in their addressee.

c. (a) and (b) are true in virtue of facts about the semantic component of $S$’s faculty of language.
For any intuitive instance of logical consequence, $\Phi \models \Psi$, the fact that it strikes us as valid is explained by our sensitivity to the following fact:

A structurally rational* agent who is in mental state $\llbracket \Phi \rrbracket$ is also in mental state $\llbracket \Psi \rrbracket$.

*A structurally rational agent is one who exemplifies Bratman-style coherence relations.
A's cognitive model = a set-theoretic representation of A's beliefs and plans

\( M_A \)

A's belief state = the set of worlds compatible with what A believes

\( B_A \)

A's intention state = the set of worlds compatible with what A intends

\( I_A \)

\( M_A \) determines: (in a way to be explained)
under the idealized assumption that A is structurally rational.

$M_A$

A’s COGNITIVE MODEL
=a set-theoretic representation of A’s beliefs and plans

...under the idealized assumption that A is structurally rational.
TWO KINDS OF IDEALIZATION
FIRST: GALILEAN

Conflate mental states with necessarily equivalent contents.

( Distinction between Galilean and Minimal idealization due to Weisberg 2007 )
TWO KINDS OF IDEALIZATION
SECOND: MINIMALIST

Mental states are consistent and closed under entailment.
TWO KINDS OF IDEALIZATION
SECOND: MINIMALIST

$I_A \subseteq B_A$
TWO KINDS OF IDEALIZATION
SECOND: MINIMALIST

\[ I_A \subseteq B_A \]

Doxastic Constraint on Intending
A can’t intend to \( \psi \) if it is ruled out by A’s beliefs that A will \( \psi \).
TWO KINDS OF IDEALIZATION
SECOND: MINIMALIST

\[
I_A \subseteq B_A
\]

Strict Means-End Coherence

If:
(i) A intends to \( \varphi \).
(ii) A believes that \( \psi \) is necessary for \( \varphi \)ing.

then:
(iii) A intends to \( \psi \).
In other words:

We model an agent’s plans as a selection function.

This function maps each belief state to a subset of itself, which is the set of worlds compatible with what the agent intends.
How to define plans as selection functions?

1. Stipulate it as a primitive

2. $M_A = \langle B_A, X \rangle$, where $X$ is a set of intentions or preferences that imposes an ordering $<X$ on $B_A$. $I_A$ is $\max_{<X}(B_A)$. (Charlow 2014)

3. Construct $M_A$ from $A$'s belief worlds and the HYPERPLANS compatible with $A$'s plans.
A’s planning state, $P_A$, is a set of HYPERPLANS.

A HYPERPLAN for A is a selection function that maps each of A’s possible belief states to the intersection of that belief state with one of A’s CHOICE WORLDS. (cf. Yalcin 2012)

A’s CHOICE WORLDS are sets of worlds that are equivalent with respect to all of the choices that A could ever have to make.

Intuitively: each of A’s HYPERPLANS makes every choice that A could ever have to make.
A's **COGNITIVE MODEL** is a set of world/hyperplan pairs.

A's **BELIEF STATE** is the set of all of the world coordinates.

A's **PLANNING STATE** is the set of all of the hyperplan coordinates.
PLANS AND HYPERPLANS

Intuitively: $P_A$ is the set of $A$’s current fully specified practical options—ways of turning possible belief states into full life plans.

$$I_A = \{w : (\exists h \in P_A)(w \in h(B_A))\}$$

$A$’s **INTENTION STATE** is the set of worlds in $A$’s **BELIEF STATE** not ruled out by all of $A$’s current practical options.
The semantic values of clauses are properties of cognitive models:

For any sentence radical \( \varphi \),

\[
[[\top \varphi]]^c = \lambda M_A . B_A \subseteq [[\varphi]]^c
\]

\[
[[!\varphi]]^c = \lambda M_A . I_A \subseteq [[\varphi]]^c
\]
(We’re about to go into the bar together. I say:) 

**Buy us drinks and I’ll find a table.**

**NOTE:**

- Needn’t have a conditional meaning.
- Can mean roughly: ‘I’ll find a table. Buy me a drink.’
- Can be the consequent of a conditional: ‘If your friend is tending bar, buy us drinks and I’ll find a table.’
CONJUNCTION

Where $\Phi, \Psi$ are sentences that may be imperatives, declaratives, or combinations of the two:

What is $\llbracket \Phi \text{ and } \Psi \rrbracket$?
CONJUNCTION

\[ [\Phi \text{ and } \Psi] = \lambda M. [\Phi](M) = 1 \text{ and } [\Psi](M) = 1 \]
$[\text{Buy us drinks and I'll find a table}]^c = \lambda M . M \text{ intends to buy drinks and } M \text{ believes that speaker}_c \text{ will find a table}$

(The property of being a mind that intends to buy drinks and believes that speaker$_c$ find a table.)
PROBLEM FOR STATIC VIEWS:
MIXED COORDINATION

(We’re at a book store. Each of us has three books, but we only have enough money for five, total:)

Put back Naked Lunch or I’ll put back Waverley.

(Starr ms)

NOTE:

• Needn’t have a conditional meaning.

• Can be the consequent of a conditional: ‘If we only have $5, put back Naked Lunch or I’ll put back Waverly.’
Where $\Phi, \Psi$ are sentences that may be imperatives, declaratives, or combinations of the two:

What is $[\Phi \text{ or } \Psi]$?
WEAK DISJUNCTION

\[ [\Phi \text{ or } \Psi] = \lambda M \cdot (\exists M^1, M^2) \begin{cases} [\Phi](M^1) = 1 \text{ and } \\ [\Psi](M^2) = 1 \text{ and} \\ M^1 \cup M^2 : = M \end{cases} \]
WEAK DISJUNCTION

\[
\left[ \text{Put back Naked Lunch or I’ll put back Waverly} \right]^c = \\
\lambda M . M \text{ is in a state of either:}
\]

(a) intending to put back Naked Lunch; or
(b) believing that speakers will put back Waverly; or
(c) indecision between options (a) and (b), but commitment to at least one.
A problem with **weak disjunction**:

(1) A: I’ll put back Naked Lunch.
(2) B: Put back Naked Lunch *or* I’ll put back Waverly.

If A is being sincere with (1), then B’s response is redundant. So why doesn’t it *sound* redundant?
A problem with weak disjunction:

(1) A: Class is in room 505.
(2) B: It’s in 505 or it’s in 506.

If A is being sincere with (1), then B’s response is redundant. So why doesn’t it sound redundant?
DISJUNCTION

WEAK DISJUNCTION

$\left[ \text{Put back Naked Lunch or I’ll put back Waverly} \right]_c = \lambda M . M \text{ is in a state of either:}$

(a) intending to put back Naked Lunch; or

(b) believing that speakers will put back Waverly; or

(c) indecision between options (a) and (b), but commitment to at least one.
DISJUNCTION

WEAK DISJUNCTION

\[ \text{[Put back Naked Lunch or I’ll put back Waverly]} \]^c = \lambda M . \text{M is in a state of either:} \\
\text{(a) intending to put back Naked Lunch; or} \\
\text{(b) believing that speakers will put back Waverly; or} \\
\text{(c) indecision between options (a) and (b), but} \\
\text{commitment to at least one.} \]
DISJUNCTION

STRONG DISJUNCTION

\[ [\phi \text{ or } \psi] = \lambda M. (\exists M^1: I_{M^1} \neq \emptyset) (\exists M^2: I_{M^2} \neq \emptyset): \]

\[
\begin{cases}
[\phi](M^1) = 1 \text{ and } \\
[\psi](M^2) = 1 \text{ and } \\
M^1 \cup M^2: = M
\end{cases}
\]

Intuitively: In uttering a disjunction, I intend for you to take both alternatives seriously, at least initially and for the purposes of practical reasoning.
**STRONG DISJUNCTION**

\[ \{ \text{Put back Naked Lunch or I’ll put back Waverly} \}^c = \lambda M . \text{M is in a state of indecision between intending to put back Naked Lunch and believing that speaker}_c \text{ will put back Waverly, but commitment to at least one of these options.} \]
Where $\Phi$ is a declarative and $\Psi$ is a declarative, imperative, or combination of the two:

What is $⟦\text{if } \Phi \text{ then } \Psi⟧$?
MAXIMAL SUBMODEL

A maximal $\Phi$-supporting submodel $M^\Phi$ of $M$ meets the following conditions:

(i) $M^\Phi \subseteq M$;

(ii) $\langle \Phi \rangle(M^\Phi) = 1$;

(iii) There is no $M^*$ such that:

- $M^\Phi \subseteq M^* \subseteq M$; and

- $\langle \Phi \rangle(M^*) = 1$
$[[\text{if } \Phi \text{ then } \Psi]] = \lambda M . (\forall M^\Phi)[[\Psi]](M^\Phi)$

M satisfies $[[\text{if } \Phi \text{ then } \Psi]]$ iff every maximal $\Phi$-satisfying submodel of M satisfies $\Psi$.

Intuitively:
In uttering ‘if $\Phi$ then $\Psi$’, I intend you to enter a state of mind such that, if you were to also form the belief $[[\Phi]]$, it would be irrational for you not to also enter the state $[[\Psi]]$. 
If Quinn is bartending, buy the first round] = \lambda M_A. If A were to be in mental state $M_A$ and believe that Quinn is bartending, A would intend to buy the first round.
CONSEQUENCE (QUICKLY)
\{\Phi_1...\Phi_n\} \models \Psi \text{ iff: }

(\forall M) \text{ if } \llbracket \Phi_1 \rrbracket(M) = 1, \ldots, \llbracket \Phi_n \rrbracket(M) = 1, \text{ then } \llbracket \Psi \rrbracket(M) = 1

\Psi \text{ follows from } \{\Phi_1...\Phi_n\} \text{ iff every cognitive model that satisfies all of the premises also satisfied the conclusion.}
Buy me a drink.

You won’t buy me a drink unless you go to the bar.

\[ \Rightarrow \text{So, go to the bar!} \]

Attack if the weather is good.

The weather is good.

\[ \Rightarrow \text{So, attack!} \]
ROSS'S PARADOX

Post the letter
≠ Post the letter or burn the letter.

*Note: this inference is blocked only if we adopt STRONG DISJUNCTION.
CONSEQUENCE (QUICKLY)

FREE CHOICE PERMISSION

Have tea or coffee.
≡ Have tea.

Sounds valid only if the conclusion is read as a weak use of the imperative: a permission, acquiescence, invitation, or instruction.

Here’s clause for weak imperatives that would validate this inference:

\[ [[i\varphi]]^c = \lambda M_{IA} . I_A \cap [[\varphi]]^c \neq \emptyset \]

Intuitively: in saying ‘have tea’, my aim was for you to make having tea compatible with your plans (at least tentatively and for the purposes of practical reasoning).
THANKS