Dynamic Semantics Meets Game Theory, by Paul Dekker

Abstract In his 1997 Linguistics and Philosophy paper, Gabriel Sandu compared the dynamic semantic and the game-theoretical treatment of anaphora, and argued in favor of the latter. In this paper we take up Sandu's findings and show that principles from Game Theoretical Semantics instead can be used to motivate and improve procedures from Dynamic Semantics.

Introduction Dynamic Semantics is a branch of linguistics originating from the model-theoretic and referentially-based work of Montague and his followers. Key feature of a dynamic semantics is that it systematically takes into account certain pragmatic aspects of interpretation. Next to that of the (truth-conditional) *content* of an utterance, some versions focus on the *update* which an utterance of a sentence may bring about in the information state of a hearer; others instead focus on the *support* or evidence which a speaker may have to motivate an utterance.

In Dekker 1999 I have shown that, when we restrict our attention to a fragment of natural language which can be modeled by a language of predicate logic, the three notions interact elegantly. Content of, update with, and support for the utterance of predicate logical formulas can be defined separately, and anyone of these notions can be defined in terms of another. Besides, exchange of information can be shown to be safe. If an utterance is supported, then the update which the utterance brings about does not introduce any misinformation which the interlocutors did not have before the exchange.

The sketched outlook upon first order interpretation hinges upon two assumptions. First, indefinite noun phrase are interpreted as a certain kind of free variables, and, second, the terms are required to be used with referential intentions. Thus, support for the use of indefinite (and other) terms can be seen to relate to the game-theoretical notion of a witness. The dynamic semantic formulation of these assumptions is attractive because it does not require the existence (knowledge, disposal) of such witnesses, since it suffices to model the reliance upon or commitment to their existence.

Principles of Dynamic Interpretation If, as we hold, indefinite noun phrases associate with holes in (open) propositions, sentential operators have to be adjusted to cope with them, preferably in a motivated fashion. For conjunction this is easy. The dynamic semantic notion of conjunction can be understood to acknowledge the order of indefinite terms (referential intentions) in a discourse. The possible values of these expressions are ordered, and the information about them is conjoined. Thus, dynamic conjunction can still be understood to involve intersection, basically. (As proved in Dekker 1999)

The interaction of indefinites with other sentential connectives and operators (negation, implication, modalities) is less well understood. According to the standard dynamic semantic definition of these operators, they all induce an existential closure of the open positions deriving from indefinite terms. Thus, e.g., the dynamic semantic negation of ϕ corresponds to the classical negation of the existential closure of ϕ.

Sofar it has remained unclear what motivates precisely these closures. It has been claimed that the involved operators introduce anaphoric islands, inaccessible for subsequent anaphoric pronouns. But, as Sandu observes, this is dubious in many cases. The operators often appear to be anaphoric peninsulae. Besides, in the cases where we do find existential closure, the rationale of this is unclear.

Game Theoretical Motivation In my talk I argue that further pragmatic principles, and principles from game theoretical semantics can be used to motivate these dynamic semantic principles of interpretation, as well as further qualifications of them. In the first place, intuitive rules of conversation can be seen to conspire against referential ('wide scope') readings of indefinites in a variety of (downward entailing) contexts. In the second place, when we think, game-theoretically, of negation as switching the burden of proof, the *support* for a negated formula can be spelled out in terms of evidence against the *update* with that formula. Thus, the following definition suggest itself:

(1) $\sigma \models \neg\phi$ iff $[\![\phi]\!](\sigma) = \bot$ (state σ supports $\neg\phi$ iff the update with ϕ is absurd)

Under this interpretation, indefinites under a negation do not (need to) come with referential intentions, because they are not the speaker's responsibility. We see, furthermore, that the information which is required to support a negated formula, equals precisely the information which the formula is said to convey upon its standard dynamic semantic definition.

The support rule for $\neg$ can also be invoked to define support for an implication: $(\phi \to \psi) = \neg(\phi \wedge \neg\psi)$. Thus, we find that:

(2) $\sigma \models (\phi \to \psi)$ iff $[\![\phi]\!](\sigma) \models \psi$, that is, iff $\forall e$: if $e \in [\![\phi]\!](\sigma)$ then $\exists c$: $ce \in [\![\psi]\!]([\![\phi]\!](\sigma))$

This definition can be seen to capture the game-theoretical insight that support for a conditional consists in a procedure to transform a proof of the antecedent into a proof of the consequent. In (2) the update function $[\![\bullet]\!](\sigma)$ on σ itself turns evidence for ϕ into possible support for ψ. We furthermore find that the information required to support an implication, equals the content of the implication upon its standard dynamic semantic definition.

Functional Dependencies Proof or support for a conditional can be substantiated further, if the supporting state disposes of a function which actually turns evidence for the antecedent into proof (witness) of the consequent. This function can be indicated by a linking function $l(\to)$:

(3) $\sigma \models_l (\phi \to \psi)$ iff $[\![\phi]\!](\sigma) \models_{l(\to)(n(\phi))} \psi$, that is, iff $\forall ce$: if $e \in \sigma$ and $ce \in [\![\phi]\!](\sigma)$ then $e_{l(\to)}(c)ce \in [\![\psi]\!]([\![\phi]\!](\sigma))$

Upon this definition indefinites in the consequent of an implication are still associated with referential intentions, be it that they may be functionally dependent upon possible values of indefinites in the antecedent.

(4) If a book is printed with Elsevier it has a table of contents.

A state σ supports this implication under a link l if it disposes over a witness-function (indicated by $e_{l(\to)}$), which associates possible books printed with Elsevier with their tables of contents. Interestingly, this function can be referred back to, if we quantify again over the same set of books:

(5) In some books it is at the end.

This example is supported iff, in any case considered possible in σ, in some of the Elsevier books the associated table of contents is at the end of the book.

Further Dependencies More in general, a referential interpretation of indefinites can be maintained, in an upward entailing context, when we make them functionally dependent upon governing quantifiers or operators. Thus, we can dismiss with existential closures under, e.g., $\forall x$ and B_x, in the following way:

(6) $\sigma \models_{l,g} \forall x \phi$ iff $\forall d$: $\sigma \models_{l',g[x/d]} \phi$ with $[l'(y)]_e = e_{l(y)}(d)$ (for $y \in D(\phi)$)
(7) $\sigma \models_l B_x \phi$ iff $\forall e \in \sigma$: $\mathcal{R}_{e_v,e_x} \models_{l'} \phi$ with $[l'(y)]_e = e_{l(y)}(e_y)$ (for $y \in D(\phi)$)

Upon this definition, we get a dynamic Skolem equivalence between

(8) $\forall x \exists y \phi(x,y)$ and $\exists f \forall y \phi(x, f(x))$

such that the witnessing Skolem function f is accessible for subsequent pronouns. As predicted, dependent indefinites also provide suitable antecedents for subsequent pronouns which figure in similar contexts:

(9) Harvey courts a girl at every convention. She always comes to the banquet with him. (Karttunen 1968)
(10) If every man is given a gun, then some man will fire it. (Sandu 1997)
(11) Mary thinks there is a burglar in the house. She thinks he came in through the chimney.
(12) A woulf might come in. He would eat you first.

Support for (9) may draw from a function associating conventions with girls Harvey dates, and support for (11) relates to witnesses of Mary's, which represents burglars in her belief worlds.

Evidently, the qualified definitions of $\to$, $\forall$ and B_a require more structure in the supporting information states then we had before, but the definitions are more transparent themselves since no underlying existential closure is required.

Conclusion In this paper we have given a further elaboration of the idea that indefinite noun phrases, like other terms, are used with referential intentions. The basic idea is that the use of these expressions is supported by witnesses in the information states of the speaker, and this idea can be used to motivate the dynamic semantic notion of conjunction as a 'pragmatically' infected form of intersection. Next we have called upon principles of game-theory to motivate and improve dynamic semantic notions of implication, quantification and modalization. The notion of role-switches or turn taking, and the call for witness-functions, provides the key to understanding (and improving the formulation of) the dynamics of these operators.