INFORMATION STATES AS FIRST CLASS CITIZENS

Jørgen Villadsen

Centre for Language Technology, University of Copenhagen
Njalsgade 80, DK-2300 Copenhagen S, Denmark
Internet: jv@cst.ku.dk

ABSTRACT

The information state of an agent is changed when a text (in natural language) is processed. The meaning of a text can be taken to be this information state change potential. The inference of a consequence make explicit something already implicit in the premises — i.e. that no information state change occurs if the (assumed) consequence text is processed after the (given) premise texts have been processed. Elementary logic (i.e. first-order logic) can be used as a logical representation language for texts, but the notion of a information state (a set of possibilities — namely first-order models) is not available from the object language (belongs to the meta language). This means that texts with other texts as parts (e.g. propositional attitudes with embedded sentences) cannot be treated directly. Traditional intensional logics (i.e. modal logic) allow (via modal operators) access to the information states from the object language, but the access is limited and interference with (extensional) notions like (standard) identity, variables etc. is introduced. This does not mean that the ideas present in intensional logics will not work (possibly improved by adding a notion of partiality), but rather that often a formalisation in the simple type theory (with sorts for entities and indices making information states first class citizens — like individuals) is more comprehensible, flexible and logically well-behaved.

INTRODUCTION

Classical first-order logic (hereafter called elementary logic) is often used as logical representation language. For instance, elementary logic has proven very useful when formalising mathematical structures like in axiomatic set theory, number theory etc. Also, in natural language processing (NLP) systems, "toy" examples are easily formalised in elementary logic:

Every man lies. John is a man.
So, John lies. (1)

$\forall x (man(x) \rightarrow lie(x)), \quad man(John)$
$\vdash \quad lie(John)$ (2)

The formalisation is judged adequate since the model theory of elementary logic is in correspondence with intuitions (when some logical maturity is gained and some logical innocence is lost) — moreover the proof theory gives a reasonable notion of entailment for the "toy" examples.

Extending this success story to linguistically more complicated cases is difficult. Two problematic topics are:

Anaphora
It must be explained how, in a text, a dependent manages to pick up a referent that was introduced by its antecedent.

Every man lies. John is a man.
So, he lies. (3)

Attitude reports
Propositional attitudes involves reports about cognition (belief/knowledge), perception etc.

Mary believes that every man lies.
John is a man.
So, Mary believes that John lies. (4)

It is a characteristic that if one starts with the "toy" examples in elementary logic it is very difficult to make progress for the above-mentioned problematic topics. Much of the work on the first three topics comes from the last decade — in case of the last topic pioneering work by Hintikka, Kripke and Montague started in the sixties.

The aim of this paper is to show that by taking an abstract notion of information states as starting point the "toy" examples and the limitations of elementary logic are better understood. We argue that information states are to be taken serious in logic-based approaches to NLP. Furthermore, we think that information states can be regarded as sets of possibilities (structural aspects can be added, but should not be taken as stand-alone).

Information states are at the meta-level only when elementary logic is used. Information states are still mainly at the meta-level when intensional logics (e.g. modal logic) are used, but *some* manipulations are available at the object level.

This limited access is problematic in connection with (extensional) notions like (standard) identity, variables etc. Information states can be put at object level by using a so-called simple type theory (a classical higher-order logic based on the simply typed λ-calculus) — this gives a very elegant framework for NLP applications.

The point is not that elementary or the various intensional logics are wrong — on the contrary they include many important ideas — but for the purpose of understanding, integrating and implementing a formalisation one is better off with a simple type theory (stronger type theories are possible, of course).

AGENTS AND TEXTS

Consider an agent processing the texts $t_1, \ldots, t_n$.

By processing we mean that the agent accepts the information conveyed by the texts. The texts are assumed to be declarative (purely informative) and unambiguous (uniquely informative). The texts are processed one by one (dynamically) — not considered as a whole (statically). The dynamic interpretation of texts seems more realistic than the static interpretation.

By a text we consider (complete) discourses — although as examples we use only single (complete) sentences. We take the completeness to mean that the order of the texts is irrelevant. In general texts have expressions as parts whose order is important — the completeness requirement only means that the (top level) texts are complete units.

INFORMATION STATES

We first consider an abstract notion of an information state (often called a knowledge state or a belief state). The initial information state I_0 is assumed known (or assumed irrelevant). Changes are of the information states of the agent as follows:

$$I_0 \xrightarrow{\tau_1} I_1 \xrightarrow{\tau_2} I_2 \xrightarrow{\tau_3} \ldots \xrightarrow{\tau_n} I_n$$

where τ_i is the change in the information state when the text t_i is processed.

An obvious approach is to identify information states with the set of texts already processed — hence nothing lost. Some improvements are possible (normalisation and the like). Since the texts are concrete objects they are easy to treat computationally. We call this approach the syntactical approach.

An orthogonal approach (the semantical approach) identifies information states with sets of possibilities. This is the approach followed here.

Note that a possibility need not be a so-called "possible world" — partiality and similar notions can be introduced, see Muskens (1989).

A combination of the two approaches might be the optimal solution. Many of these aspects are discussed in Konolige (1986).

Observe that the universal and empty sets are understood as opposites: the empty set of possibility and the universal set of texts represent the (absolute) inconsistent information state; and the universal set of possibility and the empty set of texts represent the (absolute) initial information state. Other notions of consistency and initiality can be defined.

A partial order on information states ("getting better informed") is easy obtained. For the syntactical approach this is trivial — more texts make one better informed. For the semantical approach one could introduce previously eliminated possibilities in the information state, but we assume eliminative information state changes: $\tau(I) \subseteq I$ for all I (this does not necessarily hold for non-monotonic logics / belief revision / anaphora(?) — see Groenendijk and Stokhof (1991) for further details).

Given the texts $t_1, \ldots, t_n$ the agent is asked whether a text t can be inferred; i.e. whether processing t after processing $t_1, \ldots, t_n$ would change the information state or not:

$$I_n \xrightarrow{\tau} I_n$$

Here τ is the identity function.

ELEMENTARY LOGIC

When elementary logic is used as logical representation language for texts, information states are identified with sets of models.

Let the formulas $\phi_1, \ldots, \phi_n, \phi$ be the translations of the texts $t_1, \ldots, t_n, t$. The information state when $t_1, \ldots, t_k$ has been processed is the set of all models in which $\phi_1, \ldots, \phi_n$ are all true. $t_1, \ldots, t_n$ entails t if the model set corresponding to the processing of $t_1, \ldots, t_n$ does not change when t is processed. I.e. alternatively, consider a particular model M — if $\phi_1, \ldots, \phi_n$ are all true in M then ϕ must be true in M as well (this is the usual formulation of entailment).

Hence, although any proof theory for elementary logic matches the notion of entailment for "toy" example texts, the notion of information states is purely a notion of the model theory (hence in the meta-language; not available from the object language). This is problematic when texts have other texts as parts, like the embedded sentence in propositional attitudes, since a direct formalisation in elementary logic is ruled out.

304

TRADITIONAL APPROACH

When traditional intensional logics (e.g. modal logics) are used as logical representation languages for texts, information states are identified with sets of possible worlds relative to a model $M = \langle W, \ldots \rangle$, where W is the considered set of possible worlds.

The information state when $t_1, \ldots, t_k$ has been processed is, relative to a model, the set of possible worlds in which $\phi_1, \ldots, \phi_k$ are all true.

The truth definition for a formula ϕ allows for modal operators, say $\heartsuit$, such that if ϕ is $\heartsuit\psi$ then ϕ is true in the possible worlds $W_\phi \subseteq W$ if ψ is true in the possible worlds $W_\psi \subseteq W$, where $W_\psi = f_\heartsuit(W_\phi)$ for some function $f_\heartsuit : \mathcal{P}(W) \to \mathcal{P}(W)$ (hence $M = \langle W, f_\heartsuit, \ldots \rangle$).

For the usual modal operator $\square$ the function $f_\square$ reduces to a relation $R_\square : W \times W$ such that:

$$W_\psi = f_\square(W_\phi) = \bigcup_{w_\phi \in W_\phi} \{w_\psi \mid R_\square(w_\phi, w_\psi)\}$$

By introducing more modal operators the information states can be manipulated further (a small set of "permutational" and "quantificational" modal operators would suffice — compare combinatory logic and variable-free formulations of predicate logic). However, the information states as well as the possible worlds are never directly accessible from the object language.

Another complication is that the $f_\heartsuit$ function cannot be specified in the object language directly (although equivalent object language formulas can often be found — cf. the correspondence theory for modal logic).

Perhaps the most annoying complication is the possible interference with (extensional) notions like (standard) identity, where Leibniz's Law fails (for non-modally closed formulas) — see Muskens (1989) for examples. If variables are present the inference rule of $\forall$-Introduction fails in a similar way.

SIMPLE TYPE THEORY

The above-mentioned complications becomes even more evident if elementary logic is replaced by a simple type theory while *keeping the modal operators* (cf. Montague's Intensional Logic). The λ-calculus in the simple type theory allows for an elegant compositionality methodology (category to type correspondence over the two algebras). Often the higher-order logic (quantificational power) facilities of the simple type theory are not necessary — or so-called general models are sufficient.

The complication regarding variables mentioned above manifests itself in the way that β-reduction does not hold for the λ-calculus (again,

see Muskens (1989) and references herein). Even more damaging: The (simply typed!) λ-calculus is not Church-Rosser (due to the limited α-renaming capabilities of the modal operators).

What seems needed is a logical representation language in which the information states are explicit manipulable, like the individuals in elementary logic. This point of view is forcefully defended by Cresswell (1990), where the possibilities of the information states are optimised using the well-known technique of indexing. Hence we obtain an ontology of entities and indices.

In recent papers we have presented and discussed a categorial grammar formalism capable of (in a strict compositional way) parsing and translating natural language texts, see Villadsen (1991a,b,c). The resulting formulas are terms in a many-sorted simple type theory. An example of a translation (simplified):

Mary believes that John lies. (5)

$\lambda i.believe(i, Mary, (\lambda j.lie(j, John)))$ (6)

Adding partiality along the lines in Muskens (1989) is currently under investigation.

ACKNOWLEDGMENTS

REFERENCES

M. J. Cresswell (1990). *Entities and Indices.* Kluwer Academic Publishers.

J. Groenendijk and M. Stokhof (1991). *Two Theories of Dynamic Semantics.* In J. van Eijck, editor, Logics in AI - 91, Amsterdam. Springer-Verlag (Lecture Notes in Computer Science 478).

K. Konolige (1986) *A Deduction Model of Belief.* Pitman.

R. Muskens (1989). *Meaning and Partiality.* PhD thesis, University of Amsterdam.

J. Villadsen (1991a). *Combinatory Categorial Grammar for Intensional Fragment of Natural Language.* In B. Mayoh, editor, Scandinavian Conference on Artificial Intelligence - 91, Roskilde. IOS Press.

J. Villadsen (1991b). *Categorial Grammar and Intensionality.* In Annual Meeting of the Danish Association for Computational Linguistics - 91, Aalborg. Department of Computational Linguistics, Århus Business School.

J. Villadsen (1991c). *Anaphora and Intensionality in Classical Logic.* In Nordic Computational Linguistics Conference - 91, Bergen. To appear.