

From generic sentences to scripts

Lenhart Schubert
University of Rochester

Abstract

One way to tackle the problem of acquiring general world knowledge to support language understanding and commonsense reasoning is to derive this knowledge by direct interpretation of general statements in ordinary language. One of several problems encountered in such an effort is that general statements frequently involve “donkey anaphora”. Here a “dynamic Skolemization” approach is suggested that avoids dynamic semantics and leads naturally to script- or frame-like representations.

Introduction: Long-range goals, and the need for world knowledge

A group of us at the University of Rochester are pursuing the ambitious goal of creating a broadly knowledgeable dialog agent that is motivated by curiosity, by vicarious satisfaction in fulfilling user goals, and by the sheer “pleasure” (positive utility) of engaging in conversation. However, our achievements so far fall far short of this goal: we have (1) a general knowledge representation (episodic logic, or EL) intended to match the expressiveness of human languages (e.g., (Schubert and Hwang 2000; Schubert 2000)), (2) a functioning, but still incomplete inference engine for that representation, EPILOG (e.g., (Schaeffer et al. 1993; Morbini and Schubert 2008)), (3) compositional methods for generating logical forms from English parses, (4) a large base of general “factoids” (Schubert and Van Durme 2008), and (5) a self-motivated, reward-seeking agent in a small simulated world, capable of planning ahead and simple dialogs (Liu and Schubert 2009). One of the most glaring lacunae in all of this is a sizable world knowledge base.

Two related bottlenecks, and the KNEXT system

I believe that human-like language understanding by machines is encumbered by two varieties of the notorious “knowledge acquisition bottleneck”: a *pattern* acquisition bottleneck, and a *world knowledge* acquisition bottleneck.

Concerning the former, I suggest that we need to re-conceptualize parsing/interpretation as a massively pattern-based process, in which more specific idiomatic and semantic patterns select (and sometimes override) traditional, generic phrase structure patterns (Schubert 2009). Concerning world knowledge acquisition, we are confronted with a variety of practical and theoretical problems, some of which will be my focus here.

Our approach to knowledge acquisition is text-based, i.e., we are attempting to exploit the abundant textual materials available online nowadays. Most of our efforts so far, starting about 8 years ago, have been concerned with obtaining simple general factoids from arbitrary texts (see item (4) in the opening section). For example, the sentence “*The boy entered through the open door of the old man’s shack*”, when processed by our KNEXT system (e.g., (Schubert 2002; Schubert and Tong 2003; Van Durme and Schubert 2008)) yields a set of EL formulas that are automatically verbalized as

```
A BOY MAY ENTER THROUGH A DOOR OF A SHACK.  
A SHACK MAY HAVE A DOOR.  
A DOOR CAN BE OPEN.  
A MAN MAY HAVE A SHACK.  
A MAN CAN BE OLD.
```

(assuming that the initial statistical parse is correct). We have obtained many millions of such factoids, and we regard these as providing a direct means for alleviating the pattern acquisition bottleneck; i.e., they provide patterns of predication and modification that can be used to guide a parser and semantic interpreter to produce more reliable phrasal and logical form analyses. In this way they pave the way for acquiring more complex knowledge.

Generic knowledge and donkey anaphora

In particular, with improved parsing and interpretive accuracy, we should be able to reliably interpret sentences with multiple verbs, such as “*A car crashed into a tree, killing the driver*”, or “*If a car crashes into a tree, the driver is apt to be injured or killed*”. Note that the first sentence is specific, but should lend itself to tentative generalization, while the second is already explicitly generic. (Generic sentences similar to this can be found, for example, in the Open Mind Common Sense collection (Singh et al. 2002)). A key point

is that a generic conditional sentence such as the preceding one (unlike the earlier simple factoids) can enable *inferences*. For example, given a particular instance of a car crashing into a tree, we can infer that the driver of the car was probably injured or killed.

It is the acquisition of this kind of generic conditional knowledge about the world that is currently our primary concern. We are exploring a number of approaches, including abstraction from clusters of related factoids (Van Durme, Michalak, and Schubert 2009), abstraction from particular multi-verb sentences, and interpretation of available collections of generic sentences. But there are some profound difficulties afflicting all approaches that seek to derive inference-capable generic knowledge from generic sentences phrased in ordinary language:

- **lexical ambiguity**; e.g., “*Journalists write articles*” – *articles* in what sense?
- **temporal ambiguity**; e.g., “*People worry about the economy*” – now? always?
- **quantificational ambiguity**; e.g., “*Bears are rarely dangerous*” – few bears, or bears in general at few times?
- **radical underspecification**; e.g., “*Cats land on their feet*” – under what conditions? and
- **donkey anaphora**; e.g., “*If a car crashes into a tree, the driver (of the car) is apt to be hurt*” – which driver, which car?

There are plausible approaches to all five issues, but I will focus on the last, because it is equally important from the perspective of NL semantics and from the perspective of knowledge representation and reasoning (KR&R); my approach to donkey anaphora leads to a serendipitous convergence of these perspectives.

The problem with donkey anaphora is well-known (Geach 1962). For example, in the sentence

If a farmer buys a donkey, he rides it home,

we would like to bind the anaphoric pronouns *he* and *it* to the indefinites *a farmer* and *a donkey* in the antecedent. But if we interpret the sentence (using a generic event quantifier *Gen*) as

(Gen *e*: during(*e*, Extended-present)
 [[(∃*x*: farmer(*x*))(∃*y*: donkey(*y*)) buys(*x*,*y*,*e*)]
 → (∃*e'*: after(*e'*,*e*) rides-home(*x*,*y*,*e'*))],

we are left with free variables in the consequent, falling outside the scopes of the existential quantifiers in the antecedent. This difficulty sparked the development of various versions of *dynamic semantics*, including discourse representation theory (DRT) (Kamp 1981; Heim 1982) and dynamic predicate logic (DPL) (Groenendijk and Stokhof 1991). In these logics, formulas are thought of as transforming an “input” variable assignment. In the above formalized sentence, the antecedent formula would transform the values of *x* and *y*, for any given value of *e*, so as to satisfy *buys*(*x*, *y*, *e*) (if such values exist); these values would then be available in the consequent. We say in this case that *x* and *y* are *dynamically bound* by the existential quantifiers in the

antecedent, even though they lie outside the *static* scopes of those quantifiers.

A disadvantage of dynamic semantics from a KR&R perspective is loss of modularity. For example, even in a simple narrative such as

A farmer bought a donkey; he rode it home,

if we use free variables for *he* and *it* that are dynamically bound by the existential quantifiers in the first sentence, then the combined meaning of the two sentences would be lost if we interchanged their formal representations, let alone if we stored these representations separately in a KB, for independent use in inference processes. This difficulty is aggravated in extended generic contexts. For example, in the conjunctive sentence

Every man had a gun, and many used it,

the initial universal quantifier does not cover the second sentence, and consequently dynamic interpretation of the anaphoric *it* requires repetition of material from the first sentence, along the lines

*Every man had a gun, and many men had a gun
 and used it.*

Treatment of extended generic passages in DRT was pioneered in (Carlson and Spejewski 1997), but the need to repeat earlier material for each new quantifier that applied to some portion of the generic passage was very much in evidence in that study.

Dynamic Skolemization

These problems can be avoided by Skolemizing. For example, the Skolemized representations of the two-sentence examples above would be (neglecting tense)

farmer(*A*), donkey(*B*), bought(*A*,*B*,*E1*);
 after(*E2*,*E1*), rode-home(*A*,*B*,*E2*).
 (∀*x*: man(*x*)) gun(*G*(*x*)) ∧ had(*x*,*G*(*x*));
 (Many *x*: man(*x*)) used(*x*,*G*(*x*)),

where *A*, *B*, *E1*, and *E2* are Skolem constants and *G* is a Skolem function. Note that if we immediately Skolemize upon encountering an existentially quantified variable, the Skolem constant or function will be available for resolving subsequent anaphors.

But of course the recalcitrant cases are those involving donkey anaphora, where the existential quantifiers occur in a conditional antecedent, and thus in a negative-polarity environment. Standard Skolemization should not be expected to “work” in such a case, and at first glance, it does not. For example, consider a conditional version of the last example, tentatively Skolemized as before:

If every man has a gun, then many use it.
 [(∀*x*: man(*x*)) gun(*G*(*x*)) ∧ have(*x*,*G*(*x*))]
 → (Many *x*: man(*x*)) use(*x*,*G*(*x*)).

This formulation has many undesirable models, e.g., models in which there are men, each has a gun, and none uses it, simply because *G*(*x*) happens to denote something that is not a gun; in such a case, the conditional sentence is trivially true in virtue of the falsity of the antecedent.

But this observation, that Skolemized donkey sentences

are rendered trivially true by trivially false antecedents, also holds the key to a solution: We constrain the Skolem constants or functions in the antecedent so that they cannot be trivially false. In the present example, this can be done by stipulating (separately) the following *Skolem conditional*:

$$(\forall x: \text{man}(x)) [[(\exists y: \text{gun}(y)) \text{have}(x,y)] \rightarrow [\text{gun}(G(x)) \wedge \text{have}(x,G(x))]].$$

Thus $G(x)$ is constrained to refer to a gun that a man x has, if indeed x is a man who has a gun; and thus if in fact every man has a gun, then the Skolemized antecedent is true, and so the conditional sentence as a whole is true only if many men x use $G(x)$, which must then be a gun that x has.

In general, we define *dynamic Skolemization* as the replacement of indefinites (as soon as they are encountered in NL interpretation) by Skolem constants or functions, while simultaneously stipulating a corresponding Skolem conditional. More formally:

Given an occurrence of a formula $(\exists y : \Phi)\Psi$, containing free variables x_1, \dots, x_m (and no unresolved referring expressions) in the “provisional logical form” of an English sentence being interpreted,

- (a) Assert the Skolem conditional

$$(\forall x_1) \dots (\forall x_m) [[(\exists y : \Phi)\Psi] \rightarrow [\Phi_{F(x_1, \dots, x_m)/y} \wedge \Psi_{F(x_1, \dots, x_m)/y}]],$$
 where F is a new m -place function symbol;
- (b) Replace the original occurrence of $(\exists y : \Phi)\Psi$ by

$$[\Phi_{F(x_1, \dots, x_m)/y} \wedge \Psi_{F(x_1, \dots, x_m)/y}].$$

Some points worth noting are the following. First, as long as anaphors occur after (to the right of) their indefinite coreferents, dynamic Skolemization enables anaphora resolution “as we go”. Second, we can easily generalize the above definition so that multiple indefinites occurring in sequence, and lying within the scopes of the same non-existential quantifiers, are Skolemized simultaneously. Third, the Skolem conditional can be looked at as an equivalence, since its converse is assured by existential generalization. Fifth, dynamic Skolemization reduces to ordinary Skolemization if the existential formula $(\exists y : \Phi)\Psi$ being Skolemized lies in a positive-polarity (upward-entailing) environment; i.e., in such a case the Skolem conditional can be shown to be logically redundant.

Abbreviating dynamic Skolemization: concept definitions

There is a certain verbosity in the use of dynamic Skolemization: we are writing down variants of the formula to be Skolemized, $(\exists y : \Phi)\Psi$, three times – once on each side of the Skolem conditional, and once in the Skolemized version of the original existential sentence. We can condense dynamic Skolemization through *concept definitions* that essentially name the Skolemized material, and leave the Skolem conditional itself implicit.

For example, for the earlier sentence, “*If every man has a gun, then many use it*”, instead of writing down the stated

Skolem conditional, we would assert

$$(\text{Def HG } (x) (G) [\text{gun}(G) \wedge \text{has}(x,G)]),$$

and rewrite the logical form of the given sentence as

$$[(\forall x: \text{man}(x)) \text{HG}(x) \rightarrow (\text{Many } x: \text{man}(x)) \text{use}(x,G(x))].$$

The stated *Def*-schema is a shorthand for

$$(\forall x)[\text{HG}(x) \leftrightarrow [(\exists y: \text{gun}(y)) \text{have}(x,y)] \leftrightarrow [\text{gun}(G(x)) \wedge \text{have}(x,G(x))]],$$

i.e., it stipulates that $\text{HG}(x)$ holds just in case x has a gun, which in turn holds just in case $G(x)$ is a gun that x has. In general, a *Def*-schema for an existential formula $(\exists y_1) \dots (\exists y_n) \Phi$, containing free variables x_1, \dots, x_m ,

$$(\text{Def } \Pi (x_1, \dots, x_m) (F_1, \dots, F_n) \Phi_{\underline{F}/\underline{y}}),$$

is shorthand for

$$(\forall x_1) \dots (\forall x_m) [[\Pi(x_1, \dots, x_m) \leftrightarrow (\exists y_1) \dots (\exists y_n) \Phi] \leftrightarrow \Phi_{\underline{F}(x_1, \dots, x_m)/\underline{y}}],$$

where $\underline{F}(x_1, \dots, x_m) = (F_1(x_1, \dots, x_m), \dots, F_n(x_1, \dots, x_m))$ and $\underline{y} = (y_1, \dots, y_n)$.

Creating scripts

To see how these devices lead to scripts, suppose that we are given something like the following English generic passage as a description of the process of dining at a restaurant:

Generally, when a person dines at a restaurant,
they enter it,
they get seated at a table,
they make a meal selection from a menu,
they tell the order to a waiter,
etc.

A rough provisional logical form, with anaphors subscripted for clarity with the presumed referents, is as follows:

$$(\text{Gen } e: (\exists x: \text{person}(x)) (\exists y: \text{restaurant}(y)) \text{dine-at}(x,y,e)) \\ (\exists e_1: \text{initial-part}(e_1,e)) \text{enter}(x, \text{it}_y, e_1) \wedge \\ (\exists e_2: \text{part-of}(e_2,e) \wedge \text{before}(e_1, e_2)) \\ (\exists t: \text{table}(t)) \text{get-seated-at}(\text{they}_x, t, e_2) \wedge \\ \dots \text{etc} \dots$$

The restrictor of the generic sentence is not a positive polarity environment, and so we Skolemize it dynamically, using a concept definition for conciseness. Note that the only free variable in the restrictor is e ; for mnemonic reasons, we use X, Y for the Skolem functions (of e) corresponding to x and y :

$$(\text{Def PDR } (e) (X,Y) \\ [\text{person}(X) \wedge \text{restaurant}(y) \wedge \text{dine-at}(X,Y,e)]).$$

The body of the generic formula is a positive-polarity environment, and so we use ordinary Skolemization there. The rewritten logical form, with the anaphors resolved, thus becomes

$$\text{Gen } e: \text{PDR}(e) \\ \text{initial-part}(E_1(e),e) \wedge \text{enter}(X(e),Y(e),e) \wedge \\ \text{part-of}(E_2(e),e) \wedge \text{before}(E_1(e),E_2(e)) \wedge \\ \text{table}(T(e)) \wedge \text{get-seated-at}(X(e),T(e),e) \wedge \\ \dots \text{etc} \dots$$

But now it is clear that the *Def*-schema for predicate *PDR*, together with the condensed logical form for the generic

restaurant dining description, is remarkably script-like in structure and meaning. In fact, if we were to package the two items together, using a *Script* header, we would have

```
(Script PDR (e) (X,Y)
 [person(X) ∧ restaurant(Y) ∧ dine-at(X,Y,e)]
 initial-part(E1(e),e) ∧ enter(X(e),Y(e),e) ∧
 part-of(E2(e),e) ∧ before(E1(e),E2(e)) ∧
 table(T(e)) ∧ get-seated-at(X(e),T(e),e) ∧
 ...etc... ),
```

where by definition we take the first 4 items following “*Script*” to denote (respectively) the *script name* (i.e., the predicate defined in the implicit *Def*-schema), the *script parameters*, the *script roles*, and the *script summary* (much as proposed in (Lehnert 1977)). The advantages here, however, are that

- The representation is derived directly from English logical form.
- The constituents of the script have a precise semantics, via their defined relationship to the original logical representation.
- We can refer to functional entities in scripts *outside* the scripts; this provides a solution to many problems in the interpretation of *bridging anaphora*.
- Because of the preceding point, we have complete flexibility in elaborating the derived script or variants/specializations of it.

The last two points deserve brief illustration. Suppose that we have stored the above sort of script in a KB. and we come across the sentences

John dined at Mario’s. His table was wobbly.

In traditional script theory, as well as in a dynamic semantics approach, there is no simple way to equate the table referred to with the table John was seated at, if he followed the generic restaurant script. Essentially we would have to instantiate the general script, which would have John being seated at some table and eating at that table, and then identify that table with the one referred to here. In the present approach, we can simply refer to the table as $T(E)$, where E is John’s dining at Mario’s, and its proper role in the script instance – which need it not be made explicit – will be well-defined. The same method applies, for instance to parts of familiar kinds of objects. For example, in

John ran to a nearby house. The door was locked,

we can easily interpret *the door* as a Skolem function applied to the house in question, where that Skolem function was dynamically generated in the interpretation of a description along the lines, “*A house generally has a door at the front, through which people can enter the house*”. Descriptions such as this one, which quantify over objects rather than events, can be thought of as giving rise to Minsky-like *frames* rather than scripts, upon dynamic Skolemization, and packaging with the appropriate *Def*-schema.

Concerning variant scripts and specialized scripts, consider the following elaboration of the original generic passage:

When people dine at a fancy restaurant, their table

usually has a tablecloth on it.

Much as in the previous examples, we can use the available Skolem functions for this elaboration:

```
(Usually e: [PDR(e) ∧ fancy(Y(e))]
 (∃c: tablecloth(c)) on(c,T(e)).
```

We can in turn Skolemize the tablecloth:

```
(Usually e: [PDR(e) ∧ fancy(Y(e))]
 tablecloth(C(e)) ∧ on(C(e),T(e)).
```

Concluding remarks

I have indicated briefly that dynamic Skolemization can be used to deal systematically with donkey anaphora in generic sentences, without recourse to dynamic semantics or repetition of material. Moreover, the approach fortuitously leads to a convergence with traditional script- or frame-representations, while providing more flexibility and having a precise semantics. The connection between Skolem functions and roles in frames was noted early on in (Hayes 1981), but Hayes did not consider the possibility of Skolemizing indefinites in negative-polarity environments. Also, what has been added here is a systematic transduction from language to scripts or frames, and an elucidation of the role of Skolem constants/functions in anaphora, including donkey anaphora.

In previous work, I described both the dynamic Skolemization approach (Schubert 1999) and an approach called *implicit Skolemization* that avoids the stipulation of Skolem conditionals (Schubert 2007). But neither approach seemed fully satisfactory: Dynamic Skolemization seemed incapable of dealing with extended negative contexts where an indefinite phrase and anaphoric references to it both lie within the same downward-entailing environment; implicit Skolemization deals uniformly with all environments – but the semantics remains dynamic. However, I have recently found that negative-polarity environments can in fact be handled straightforwardly within the dynamic Skolemization approach. In essence, we distinguish *co-Skolem* constants/functions from the ordinary variety, and interpret these in terms of universal closure (whereas ordinary Skolem constants/functions are in effect interpreted in terms of existential closure). It turns out that co-Skolemization in donkey anaphora leads to “strong readings” of donkey sentences (where for instance “*If a man has a gun, he uses it*” is considered true only if multiple-gun owners use *all* their guns; in the weak readings I have restricted myself to herein, the multiple-gun owners need only use one of their guns). In current work, I am also looking at dynamic Skolemization of indefinites within attitudes such as belief, and this likewise appears feasible.

Acknowledgements

This work was supported by NSF grants IIS-0328849 and IIS-0535105.

References

- Carlson, G., and Spejewski, B. 1997. Generic passages. *Natural Language Semantics* 5(2):101–165.
- Geach, P. 1962. *Reference and Generality*. Ithaca, N.Y.: Cornell University Press.
- Groenendijk, J., and Stokhof, M. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:39–100.
- Hayes, P. 1981. The logic of frames. In Webber, B., and Nilsson, N., eds., *Readings in Artificial Intelligence*. Palo Alto, CA: Tioga Publishing Company.
- Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. Dissertation, U. Mass.
- Kamp, H. 1981. A theory of truth and semantic representation. In Groenendijk, J.; Janssen, T.; and Stokhof, M., eds., *Formal Methods in the Study of Language*. Amsterdam, Netherlands: Mathematical Centre-tracts, U. Amsterdam.
- Lehnert, W. 1977. Human and computational question answering. *Cognitive Science* 1(1):47–73.
- Liu, D., and Schubert, L. 2009. Incorporating planning and reasoning into a self-motivated, communicative agent. In *The 2nd Conf. on Artificial General Intelligence (AGI-09)*, 108–113.
- Morbini, F., and Schubert, L. 2008. Metareasoning as an integral part of commonsense and autocognitive reasoning. In *AAAI-08 Workshop on Metareasoning*, 155–162. Forthcoming as chapter in M. T. Cox & A. Raja (eds.), *Metareasoning: Thinking about Thinking*, MIT Press.
- Schaeffer, S.; Hwang, C.; de Haan, J.; and Schubert, L. 1993. EPILOG, the computational system for episodic logic: User’s guide. Technical report, Dept. of Computing Science, Univ. of Alberta, Edmonton, Canada. prepared for Boeing Computer Services, Seattle, WA.
- Schubert, L., and Hwang, C. 2000. Episodic Logic meets Little Red Riding Hood: A comprehensive, natural representation for language understanding. In Iwanska, L., and Shapiro, S., eds., *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*. Menlo Park, CA, and Cambridge, MA: MIT/AAAI Press. 111–174.
- Schubert, L., and Tong, M. 2003. Extracting and evaluating general world knowledge from the Brown Corpus. In *Proc. of the HLT-NAACL Workshop on Text Meaning*, 7–13.
- Schubert, L., and Van Durme, B. 2008. Open extraction of general knowledge through compositional semantics. In *Abstracts of the NSF Symposium on Semantic Knowledge Discovery, Organization and Use*.
- Schubert, L. 1999. Dynamic Skolemization. In Bunt, H., and Muskens, R., eds., *Computing Meaning, vol. 1*. Dordrecht: Kluwer Academic Press. 219–253.
- Schubert, L. 2000. The situations we talk about. In Minker, J., ed., *Logic-Based Artificial Intelligence*. Dordrecht: Kluwer. 407–439.
- Schubert, L. 2002. Can we derive general world knowledge from texts? In *Proc. of the 2nd Int. Conf. on Human Language Technology Research (HLT 2002)*, 94–97.
- Schubert, L. 2007. Implicit Skolemization: Efficient reference to dependent entities. *Research on Language and Computation* 5:69–86.
- Schubert, L. 2009. Language understanding as recognition and transduction of numerous overlaid patterns. In *AAAI Spring Symposium on Learning by Reading and Learning to Read*, 94–96.
- Singh, P.; Lin, T.; Mueller, E. T.; Lim, G.; Perkins, T.; and Zhu, W. L. 2002. Open Mind Common Sense: Knowledge acquisition from the general public. In *Proc. of the 1st Int. Conf. on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems*.
- Van Durme, B., and Schubert, L. 2008. Open knowledge extraction through compositional language processing. In *Symposium on Semantics in Systems for Text Processing (STEP 2008)*.
- Van Durme, B.; Michalak, P.; and Schubert, L. 2009. Deriving generalized knowledge from corpora using WordNet abstraction. In *12th Conf. of the European Chapter of the Assoc. for Computational Linguistics (EACL-09)*.