A Declarative Formalization of Knowledge Translation [1]

Saša Buvač
Computer Science Department
Stanford University
Stanford
California 94305-2140
buvac@cs.stanford.edu

Richard Fikes
Knowledge Systems Laboratory
Stanford University
701 Welch Road, Bldg. C
Palo Alto, California 94304
fikes@ksl.stanford.edu

Abstract

We describe an interlingua-based methodology for translating encoded knowledge and present a formalism for declaratively specifying vocabulary translations within a predicate logic interlingua. In this paper we (1) use the formalism to provide a semantics for translation, (2) show that the formalism enables translation to be done as deduction by a standard theorem prover, (3) describe a proof technique for determining whether a given set of rules for translating from one vocabulary to another is sufficient for performing that translation for any theory, and (4) describe techniques for precompiling translation rules that translate directly between two given vocabularies.

Motivation

Acquiring and representing knowledge is the key to building powerful intelligent systems. Unfortunately, knowledge base construction is difficult and time consuming. The development of most systems requires a new knowledge base to be constructed from scratch. As a result, most systems remain small to medium in size. A promising approach to removing this barrier to the building of intelligent systems is to develop techniques for encoding knowledge in a reusable form so that large portions of a knowledge base for a given application can be assembled from knowledge repositories and other systems.

For encoded knowledge to be incorporated into a system's knowledge base, the knowledge must either be represented in the receiving system's representation language or be translatable in some practical way into that language. Since an important means of achieving efficiency in application systems is to use specialized representation languages that directly support the knowledge processing requirements of the application, we cannot expect a standard knowledge representation language to emerge that would be used generally in application systems. Thus, we are confronted with a *heterogeneous language problem* whose solution requires a capability for translating encoded knowledge among specialized representation languages.

We are addressing this problem by participating in the development of a translation technology for knowledge representation languages based on the use of an *interlingua* for communicating knowledge among systems. Given such an interlingua, a sending system would translate knowledge from its application-specific representation into the interlingua for communication purposes and a receiving system would translate knowledge from the interlingua into its application-specific representation before use. In addition, the interlingua is the language in which libraries would provide reusable knowledge bases.

Translating knowledge from one representation to another involves translating both syntax and vocabulary. Thus, an interlingua must provide a syntax and a vocabulary into and out of which knowledge can be effectively translated. A linear ASCII predicate logic syntax appears to be a suitable interlingua syntax for a broad spectrum of applications. However, no such single interlingua vocabulary is likely to be suitable, since most encoded knowledge uses a rich vocabulary of complex domain-specific terms. An interlingua that requires such terms to be restated in a low-level vocabulary of sets, tuples, integers, etc. will not be practically useful for doing translation in most cases. Thus, we expect that in practice, most interlinguas will be *domain-specific interlinguas* comprised of a general-purpose core logic-based interlingua extended to include a foundation vocabulary and theory for an application domain of interest.

We consider in this paper the problem of translating logical theories from one *representation* to another using a logic-based domain-specific interlingua. By a representation, we mean (1) a declarative representation language; (2) a vocabulary of object and relation constants; and (3) a *base theory* that is considered to be included in any theory expressed in the representation. A representation's base theory is the union of the theory specified by the semantics of the representation language and whatever domain theory is being assumed in the representation.

[1] This paper is a short version of (Buvač & Fikes 1994), which is available at http://sail.stanford.edu/buvac.

A straightforward notion of translation might require that translation be *equivalence preserving*, i.e., that the translation of a source theory be logically equivalent to the source theory. However, translation may involve abstracting, approximating, or otherwise transforming the source theory. We describe here a multi-step translation methodology in which the translation steps into and out of the interlingua are assumed to be equivalence preserving, and translation steps within the interlingua may involve arbitrary transformations of the theory being translated, as specified by sets of translation rules.

We assume a first-order logic interlingua with a model-theoretic semantics such as the **K**nowledge **I**nterchange **F**ormat (KIF) (Genesereth & Fikes 1992) being developed in the ARPA Knowledge Sharing Initiative (Patil *et al.* 1992). We assume that the interlingua includes a domain-specific vocabulary and a domain theory. Finally, we assume that any theory representable in the source language is also representable in the interlingua and that any theory representable in the target language is also representable in the interlingua.

Defining Translation in Terms of Truth

In this section, we describe a formalism for specifying translations within a first-order logic language and use the formalism to define translation in terms of truth. Intuitively, we will say that a theory T' is a translation of theory T iff all the sentences in T' are true whenever all the sentences in T are true.

We assume a context formalism as being developed by McCarthy, et al, (McCarthy & Buvač 1994; Guha 1991; Buvač & Mason 1993; Attardi & Simi 1995; Buvač, Buvač, & Mason 1995). A context can be thought of as an object in a first-order language which has an associated *vocabulary* (i.e., a set of non-logical symbols) and denotes a set of models. Formulas in the vocabulary of a context are referred to as *meaningful* in that context. The formalism includes a truth modality $\mathtt{ist}(\kappa, \phi)$ (pronounced "is true"), which holds iff formula ϕ is true in all the models denoted by context κ. Furthermore, we extend the $\mathtt{ist}$ modality to finitely axiomatizable theories as follows. Assume T is a finitely axiomatizable theory. Let A be a finite set of axioms such that $T = Th(A)$. Then $\mathtt{ist}(\kappa, T)$ iff $\mathtt{ist}(\kappa, \bigwedge\!\!\!\bigwedge A)$, i.e., iff the conjunction of the axioms is true in context κ. Note that the choice of A among the various axiomatizations of T is irrelevant.

A particular translation is specified by giving a set of translation rules from a "source" context to a "target" context. The rules can be thought of as describing syntactic transformations, like replacement rules or grammar rules. Formally, the rules enable one to derive sentences in the target context that are equivalent to sentences in the source context.

Definition (translation rule): A translation rule from context κ to context κ' is any axiom which has the form: $\mathtt{ist}(\kappa, \phi) \leftrightarrow \mathtt{ist}(\kappa', \phi')$.

In context formalisms, a *lifting axiom* is defined to be any axiom which relates the truth in one context to the truth in another context. Our translation rules are lifting axioms which have a particular syntactic form.

Now we can define what we mean by translating a theory.

Definition (theory translation): A set of translation rules Tr from context κ to context κ' translates theory T into theory T' iff $Tr \vdash \mathtt{ist}(\kappa, T) \leftrightarrow \mathtt{ist}(\kappa', T')$.

We are assuming a sound and complete logic of contexts (like the one given in (Buvač, Buvač, & Mason 1995)). Thus, we could have used $\models$ instead of $\vdash$ in the above formula.

Translation is typically thought of in terms of individual sentences rather than theories. However, since the top level task we are addressing is the translation of theories, we need only concern ourselves with translating the theory defined by a set of sentences, rather than with translating the sentences themselves, even when the set contains only one sentence.

We now state some properties of our translation formalism which we have found to be useful for characterizing particular translations.

Definition (translatable): Theory T is *translatable* from context κ into context κ' by translation rules Tr iff there exists some theory T' such that translation rules Tr from context κ to context κ' translate theory T into theory T'. Context κ is translatable into context κ' by translation rules Tr iff every theory meaningful in context κ is translatable into context κ' by translation rules Tr from context κ to context κ'.

Definition (equivalence preserving): A translation specified by a set of translation rules Tr from context κ into context κ' is *equivalence preserving* iff for any ϕ and ψ for which there exist ϕ' and ψ' such that $Tr \vdash \mathtt{ist}(\kappa, \phi) \leftrightarrow \mathtt{ist}(\kappa', \phi')$ and $Tr \vdash \mathtt{ist}(\kappa, \psi) \leftrightarrow \mathtt{ist}(\kappa', \psi')$, it is the case that $\phi \vdash \psi$ iff $\phi' \vdash \psi'$.

Since we have defined translation in terms of truth, translations in this formalism are "almost always" equivalence preserving. Only translations which can not be represented by a function (because they map a single sentence to two non-equivalent sentences) are not equivalence preserving.

Translation as Deduction

One of the advantages of declaratively representing translation is that a standard theorem prover can be used to derive translations. Moreover, translation axioms specified for individual predicates will allow us to infer the translation of an entire theory using only standard logical deduction. In this section we illustrate this with an example of translating data bases.

We assume the following properties hold for data base contexts:

$$\mathtt{ist}(c, \phi \rightarrow \psi) \leftrightarrow (\mathtt{ist}(c, \phi) \rightarrow \mathtt{ist}(c, \psi)) \qquad (1)$$

$$(\forall x)\mathtt{ist}(c, \phi(x)) \leftrightarrow \mathtt{ist}(c, (\forall x)\phi(x)). \quad (2)$$

The $\Leftarrow$ direction of axiom 1 implies the closed world assumption (which would more commonly be expressed as $\neg\mathtt{ist}(c, \phi) \rightarrow \mathtt{ist}(c, \neg\phi)$). The closed world assumption is valid for data base contexts, but poses fairly severe constraints on the semantics of contexts, and will not hold in general for other types of contexts (see (Buvač, Buvač, & Mason 1995) for a discussion).

Here is a hypothetical example. McCarthy's diary data base and the United Airlines reservations data base use different vocabularies to represent facts about flights. In order to book a flight, McCarthy's artificially intelligent travel agent will need to translate between the two representations.

Assume McCarthy's diary contains the formula

$$(\forall d)\, \mathit{Thursday}(d) \rightarrow \quad (3)$$
$$\mathit{fly}(\mathit{UA}, d, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)$$

expressing the fact that every Thursday McCarthy flies to Los Angeles on flight 921, scheduled to leave San Francisco at 7:00 and to arrive in LA at 8:21. Note that although the term "McCarthy" is not mentioned in the above formula, the entry implicitly pertains to McCarthy since the formula is given in the context of McCarthy's diary.

United Airlines uses a relational data base to store reservation and flight information. The vocabulary of this relational data base contains relations *passenger-record* and *flight-record*. Using the relations in this data base, the fact given by formula 3 is represented by

$$(\forall d)\, \mathit{Thursday}(d) \rightarrow \quad (4)$$
$$\mathit{passenger\text{-}record}(921, d, \mathit{McCarthy})$$

$$\mathit{flight\text{-}record}(921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21). \quad (5)$$

Note that since all the flights listed in the United Airlines data base will be United Airlines flights, this information is implicitly assumed by the context of that data base.

If we take *diary*(p) to denote the context of person p's diary, and we take *UA-db* to denote the context of the United Airlines relational data base, then we can use the translation axioms

$$\mathtt{ist}(\mathit{diary}(p), \mathit{fly}(\mathit{UA}, d, f, l1, t1, l2, t2)) \leftrightarrow \quad (6)$$
$$\mathtt{ist}(\mathit{UA\text{-}db}, \mathit{passenger\text{-}record}(f, d, p) \wedge$$
$$\mathit{flight\text{-}record}(f, l1, t1, l2, t2))$$

$$\mathtt{ist}(\mathit{diary}(p), \mathit{Thursday}(d)) \leftrightarrow \quad (7)$$
$$\mathtt{ist}(\mathit{UA\text{-}db}, \mathit{Thursday}(d))$$

to show that the theory defined by formula 3 in the *diary*(*McCarthy*) context translates into the theory defined by the conjunction of formulas 4 and 5 in the *UA-db* context.

We proceed to prove this. Applying the quantifier rule given in formula 2 to formula 3 gives

$$(\forall d)\mathtt{ist}(\mathit{diary}(\mathit{McCarthy}), \mathit{Thursday}(d) \rightarrow \quad (8)$$
$$\mathit{fly}(\mathit{UA}, d, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)).$$

From formulas 8 and 1 we can infer

$$(\forall d)\mathtt{ist}(\mathit{diary}(\mathit{McCarthy}), \mathit{Thursday}(d)) \rightarrow \quad (9)$$
$$\mathtt{ist}(\mathit{diary}(\mathit{McCarthy}),$$
$$\mathit{fly}(\mathit{UA}, d, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)),$$

which in conjunction with translation axioms 6 and 7 gives

$$(\forall d)\mathtt{ist}(\mathit{UA\text{-}db}, \mathit{Thursday}(d)) \rightarrow \quad (10)$$
$$\mathtt{ist}(\mathit{UA\text{-}db}, \mathit{passenger\text{-}record}(921, d, \mathit{McCarthy}) \wedge$$
$$\mathit{flight\text{-}record}(921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)).$$

Again, applying the property of `ist` given in formula 1 to formula 10 gives

$$(\forall d)\mathtt{ist}(\mathit{UA\text{-}db}, \mathit{Thursday}(d) \rightarrow \quad (11)$$
$$(\mathit{passenger\text{-}record}(921, d, \mathit{McCarthy}) \wedge$$
$$\mathit{flight\text{-}record}(921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)),$$

and, making use of the quantificational property of contexts described in formula 2, we get

$$\mathtt{ist}(\mathit{UA\text{-}db}, (\forall d)\, \mathit{Thursday}(d) \rightarrow \quad (12)$$
$$(\mathit{passenger\text{-}record}(921, d, \mathit{McCarthy}) \wedge$$
$$\mathit{flight\text{-}record}(921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)),$$

which is equivalent to the conjunction of formulas 4 and 5. Since all the steps in this proof are reversible, the other direction follows in the same fashion. Therefore, we have proved that

$$\mathtt{ist}(\mathit{diary}(\mathit{McCarthy}), (\forall d)\, \mathit{Thursday}(d) \rightarrow \quad (13)$$
$$\mathit{fly}(\mathit{UA}, d, 921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)) \leftrightarrow$$
$$\mathtt{ist}(\mathit{UA\text{-}db}, (\forall d)\, \mathit{Thursday}(d) \rightarrow$$
$$\mathit{passenger\text{-}record}(921, d, \mathit{McCarthy})$$
$$\wedge \mathit{flight\text{-}record}(921, \mathit{San\text{-}Francisco}, 7{:}00, \mathit{LA}, 8{:}21)),$$

which is the desired result.

Proving Translatability

Another advantage of declaratively representing translation in first-order logic is that a standard theorem prover can be used to show that any theory in one context is (or is not, as the case may be) translatable to another context via a given set of translation rules. In the above example, if we assume that context *diary*(p) contains only the non-logical predicates listed in formula 3, then any theory in the context *diary*(p) (for any person p) is translatable into the context *UA-db* by translation rules 6 and 7. Thus, any context *diary*(p) is translatable into the context *UA-db* by translation rules 6 and 7.

Since data bases commonly make a closed world assumption, strong results about their translatability will

be provable. To show that the context of one data base is translatable into another by a set of translation rules, we use the following proposition:

Proposition (compositionality of data base translation): If c and c' are data base contexts (i.e., they satisfy formula 1), then context c is translatable to c' via translation rules Tr iff all theories axiomatized by atomic formuli are translatable from c to c' via translation rules Tr.

Returning to our example, note that translatability does not hold in the other direction: it is not the case that any theory in context *UA-db* is translatable into the *diary*(p) context via the translation axioms 6 and 7. A simple counter example will best illustrate this point. The theory axiomatized by formula 5 is not translatable into *diary*(p) context because the translation axioms tell us only how to translate conjunctions from the *UA-db* context into the *diary*(p) context. Note that we can also use deduction and a theorem prover to show that one context is *not translatable* to another context. This would be done by giving a counter example, i.e., a theory which is not translatable.

Deriving Translation Rules

Representing translation rules declaratively enables new translation rules to be derived from other translation rules. Rules can be automatically derived for translating directly between two vocabularies from the rules for translating each of the vocabularies into and out of an interlingua vocabulary. For example, the *Engineering Math Ontology* (EMO), (Gruber 1994), is an example of a domain specific interlingua for the domain of engineering mathematics. Assume that force is represented as a vector in SI units in the EMO. Now, assume we are given two other ontologies which we want to translate between. One uses English units, and the other deals only with a single dimensional space and therefore represents force as a scalar. If translation rules are given for these ontologies into and out of the EMO, a theorem prover could derive translation rules that would translate directly between the ontologies. The derived rules could then be used to perform translation more efficiently by avoiding the steps of translating into and out of the EMO.

We now give a way of characterizing derived translations.

Proposition (translatability): Assume context c_1 is translatable into context c by translation rules Tr_1, and context c is translatable into context c_2 by translation rules Tr_2. Then we can derive translation rules Tr such that context c_1 is translatable into context c_2 by translation rules Tr.

Summary

We have described an interlingua-based methodology for translating encoded knowledge and presented a formalism for declaratively specifying vocabulary translations within a predicate logic interlingua. We used the formalism to provide a formal semantics for translation; noted that the formalism enables translation to be done as deduction by a standard theorem prover; described a proof technique for determining whether a given set of rules for translating from one vocabulary to another is sufficient for performing that translation for any theory; and described techniques for precompiling translation rules that translate directly between two given vocabularies. In the full version of this paper, (Buvač & Fikes 1994), we also provided examples of how the formalism can be used to specify various forms of translation, including abstractions and approximations, as well as showed how the formalism might also be used to specify the syntactic translations, like those typically described by grammar-based translators (Van Baalen & Fikes 1994).

References

Attardi, G., and Simi, M. 1995. A formalization of viewpoints. *Fundamenta Informaticae* 23(3).

Buvač, S., and Fikes, R. 1994. A declarative formalization of knowledge translation. Technical Report KSL-94-59, Knowledge Systems Laboratory, Stanford University.

Buvač, S., and Mason, I. A. 1993. Propositional logic of context. In *Proceedings of the Eleventh National Conference on Artificial Intelligence.*

Buvač, S.; Buvač, V.; and Mason, I. A. 1995. Metamathematics of contexts. *Fundamenta Informaticae* 23(3).

Genesereth, M. R., and Fikes, R. E. 1992. Knowledge interchange format, version 3.0, reference manual. Technical Report Logic Group Report Logic-92-1, Stanford University.

Gruber, T. R. 1994. An ontology for engineering mathematics. In *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning.*

Guha, R. V. 1991. *Contexts: A Formalization and Some Applications.* Ph.D. Dissertation, Stanford University.

McCarthy, J., and Buvač, S. 1994. Formalizing Context (Expanded Notes). Technical Note STAN-CS-TN-94-13, Stanford University.

Patil, R. S.; Fikes, R. E.; Patel-Schneider, P. F.; Mckay, D.; Finin, T.; Gruber, T.; and Neches, R. 1992. The DARPA knowledge sharing effort: Progress report. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning.*

Van Baalen, J., and Fikes, R. E. 1994. The role of reversible grammars in translating between representation languages. In *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning.*