Reasoning with Embedded Formulas and Modalities in SUMO[1]

Benzmüller Christoph and **Pease Adam**[2]

Abstract. Reasoning with embedded formulas is relevant for the SUMO ontology but there is limited automation support so far. We investigate whether higher-order automated theorem provers are applicable for the task. Moreover, we point to a challenge that we have revealed as part of our experiments: modal operators in SUMO are in conflict with Boolean extensionality. A solution is proposed.

1 EMBEDDED FORMULAS IN SUMO

The open source Suggested Upper Merged Ontology[3] (SUMO) [9] (and similarly, proprietary Cyc [13]) contains a small but significant amount of higher-order representations. The approach taken in these systems to address higher-order challenges has been to employ specific translation 'tricks', possibly in combination or in addition to some pre-processing techniques. Examples of such means are the quoting techniques for embedded formulas as employed in SUMO [11] and the heuristic-level modules in CYC [13]. Unfortunately, however, these solutions are strongly limited. The effect is that many desirable inferences are currently not supported, so that many relevant queries cannot be answered.

This includes statements in which formulas are embedded as arguments of terms, for example, statements that employ epistemic operators such as `believes` or `knows`, temporal operators such as `holdsDuring`, and further operators such as `disapproves` or `hasPurpose`. While first-order automated theorem proving (FO-ATP) for SUMO has strongly improved recently [12], there is still only very limited support for reasoning with non-trivial embedded formulas; we give an example (free variables in premises are universal and those in the query are existential):

Ex. 1 (Reasoning in temporal contexts.) *What holds that holds at all times. Mary likes Bill.*[4] *During 2009 Sue liked whoever Mary liked. Is there a year in which Sue has liked somebody?*
A: `(=> ?P (holdsDuring ?Y ?P))`
B: `(lk Mary Bill)`
C: `(holdsDuring (YearFn 2009)`
`(forall (?X) (=> (lk Mary ?X) (lk Sue ?X))))`
Q: `(holdsDuring (YearFn ?Y) (lk Sue ?X))`

This example, which is a challenge for FO-ATP (note the embedded first-order formula), is actually trivial for higher-order automated theorem provers (HO-ATP): the prover LEO-II [5] can solve it in 0.16 sec. on a standard MacBook. A slight modification of Ex.1, which LEO-II proves in 0.08 sec., is:[5]

Ex. 2 (Ex.1 modified; A is replaced by 'True always holds'.)
A': `(holdsDuring ?Y True)`
B: `(lk Mary Bill)`
C: `(holdsDuring (YearFn 2009)`
`(forall (?X) (=> (lk Mary ?X) (lk Sue ?X))))`
Q: `(holdsDuring (YearFn ?Y) (lk Sue ?X))`

Further examples are studied in [6]; there we also outline the translation from SUMO's SUO-KIF representation language [10, 7] as used above to the new higher-order TPTP THF syntax [14] as supported by several HO-ATPs including LEO-II.

2 THE PROBLEM WITH MODAL OPERATORS

Validity of Ex.1 and Ex.2 is easily shown provided that Boolean extensionality[6] is assumed (this ensures that the denotation of each formula, also the embedded ones, is either true of false). This assumption has actually never been questioned for SUMO, neither in [7] nor in [10].

However, this assumption also leads to problematic effects as the following slight modification of Ex.2 illustrates:

Ex. 3 (Ex.2 modified; now formulated for an epistimec context)
A": `(knows ?Y True)`
B: `(lk Mary Bill)`
C': `(knows Chris`
`(forall (?X) (=> (lk Mary ?X) (lk Sue ?X)))`
Q': `(knows Chris (lk Sue Bill))`

Using Boolean extensionality the query is easily shown valid and LEO-II can prove it in 0.04 sec. However, now this inference is disturbing since we have not explicitly required that *(knows Chris (lk Mary Bill))* holds which intuitively seems mandatory. Hence, we here (re-)discover an issue that some logicians possibly claim as widely known: modalities have to be treated with great care in classical, extensional higher-order logic. Our ongoing work therefore studies how we can suitably adapt the modeling of affected modalities in SUMO in order to appropriately address this issue. A respective proposal is sketched next.

---

[1] This work is funded by the German Research Foundation under grant BE 2501/6-1.
[2] Articulate Software, email: cbenzmueller|apease@articulatesoftware.com
[3] SUMO is available at `http://www.ontologyportal.org`
[4] To save space 'likes' is written as 'lk'.
[5] It is important to note that $True$ in **A'** can actually be replaced by other tautologies, e.g. by $(equal\, Mary\, Mary)$; this may appear more natural and the example can still be proved by LEO-II in milliseconds.
[6] For a detailed discussion of functional and Boolean extensionality in classical higher-order logic we refer to [2].

3 REASONING WITH MODALITIES IN SUMO

The solution we currently explore is to treat SUMO reasoning problems hat involve modal operators as problems in quantified multimodal logics. Unfortunately there are only very few direct theorem provers for quantified multimodal logics available. We therefore exploit our recent embedding of quantified multimodal logics in classical higher-order logic [4, 3] and we investigate whether this embedding can fruitfully support the automation of modal operators in SUMO with off-the-shelf HO-ATPs (cf. [1] for first studies).

The idea of the embedding is simple: modal formulas are lifted to predicates over possible worlds, i.e. HO-terms of type $\iota \to o$, where ι is a reserved base type denoting the set of possible worlds. For individuals we reserve a second base type μ.[7]

Modal operators such as $\top$, $\bot$, $\neg$, $\vee$, $\supset$, $\Box$, and $\forall^{ind}$, $\forall^{prop}$ are then simply defined as abbreviations of proper HO-terms, e.g. $\neg = \lambda\phi_{\iota\to o}.\lambda W_\iota.\neg\phi W$ and $\Box = \lambda R_{\iota\to\iota\to o}.\lambda\phi_{\iota\to o}.\lambda W_\iota.\forall V_\iota.\neg R\,W\,V \vee \phi V$. In the following we write $\Box_r$ for the application of $\Box$ to an accessibility relation r of type $\iota \to \iota \to o$.[8]

Exploiting this embedding we can now suitably map SUMO problems containing modal operators: e.g. **C'** in Ex. 3 is lifted to $(\Box_{Chris}\,\forall^{ind}X_\mu.((lk\,Mary\,X) \supset (lk\,Sue\,X)))_{\iota\to o}$ and **B** simply becomes $(lk\,Mary\,Bill)_{\iota\to o}$. Since $\Box$ is here associated with knowledge we axiomatize it as an S4 or S5 modality below. Similarly, we could introduce further copies of $\Box$, e.g. for $believes$, and provide different axioms for it.

The final step is to ground the lifted terms. For this, T-Box like information in SUMO, such as the axiom $(instance\ \ instance\ \ BinaryPredicate)$, is interpreted as universally quantified over all possible worlds: $\forall W_\iota.((instance\ \ instance\ \ BinaryPredicate)\ \ W)$. A-Box like information and queries in contrast are modeled with respect to a current world cw (of type ι). Since our examples only contain local premises and queries, i.e. A-Box like information, Ex.3 is thus translated as:

Ex. 4 (Translated Ex.3)
A": $\forall Y_{\iota\to\iota\to o}.((\Box_Y\,\top)\ cw)$
B: $((lk\ Mary\ Bill)\ cw)$
C': $((\Box_{Chris}\,(\forall^{ind}X_\mu.((lk\ Mary\ X) \supset (lk\ Sue\ X))))\ cw)$
Q': $((\Box_{Chris}\,(lk\ Sue\ Bill))\ cw)$

The axioms for S4 (**T+4**) or S5 (**T+5**) can be added as follows:
T: $\forall W_\iota.((\forall^{prop}\phi_{\iota\to o}.\Box_{Chris}\,\phi \supset \phi)\ W)$
4: $\forall W_\iota.((\forall^{prop}\phi_{\iota\to o}.\Box_{Chris}\,\phi \supset \Box_{Chris}\,\Box_{Chris}\,\phi)\ W)$
5: $\forall W_\iota.((\forall^{prop}\phi_{\iota\to o}.\Diamond_{Chris}\,\phi \supset \Box_{Chris}\,\Diamond_{Chris}\,\phi)\ W)$

The above example is not valid, which we wanted to achieve, and LEO-II correctly fails to prove it (timeout). However, if we move premise **B** in the context of Chris' knowledge then we get:

Ex. 5 (Modified Ex.4)
A": $\forall Y_{\iota\to\iota\to o}.((\Box_Y\,\top)\ cw)$
B': $((\Box_{Chris}\,(lk\ Mary\ Bill))\ cw)$
C': $((\Box_{Chris}\,(\forall^{ind}X_\mu.((lk\ Mary\ X) \supset (lk\ Sue\ X))))\ cw)$
Q': $((\Box_{Chris}\,(lk\ Sue\ Bill))\ cw)$

Ex.5 is valid and it is proved by LEO-II in less than 0.15 sec.

[7] In [4] we use μ for possible worlds and ι for individuals; this syntactic switch is completely unproplematic.

[8] Note the elegant way in which indexation over different accessibility relations is facilitated via λ-abstraction in the definition of $\Box$; this is the basis for combining different $\Box_k$ operators in our framework; cf. [1].

4 CONCLUSION

Reasoning with embedded formulas is naturally supported in extensional HO-ATPs. However, this leads to a problem regarding the adequate treatment of modal operators. A potential solution has been outlined in this paper that we are currently investigating further. Our ongoing work in particular studies the scalability of HO-ATPs for the task. Due to the recent, strong improvements of HO-ATPs [14] – which will be further fostered by the new higher-order CASC competitions – we are quite optimistic though. The large theories challenge obviously requires the development or adaptation of strong relevance filters, such as SInE [8], to our higher-order logic setting.

We have to admit that we currently see few alternatives to HO-ATP for the automation of ontology reasoning problems with embedded formulas and modalities as presented in this paper and already our toy examples seem challenging for other approaches.

Acknowledgment: We would like to thank Geoff Sutcliffe and Larry Paulson for their support and their contributions to our work.

REFERENCES

[1] C. Benzmüller, 'Simple type theory as framework for combining logics', in *Contest paper at the World Congress and School on Universal Logic III (UNILOG'2010)*, Lisbon, Portugal, (2010). http://arxiv.org/abs/1004.5500.

[2] C. Benzmüller, C. Brown, and M. Kohlhase, 'Higher-order semantics and extensionality', *J. of Symbolic Logic*, **69**(4), 1027–1088, (2004).

[3] C. Benzmüller and L. C. Paulson, 'Multimodal and intuitionistic logics in simple type theory', *The Logic J. of the IGPL.* (2010).

[4] C. Benzmüller and L. C. Paulson, *Quantified Multimodal Logics in Simple Type Theory*, SEKI Report SR–2009–02 (ISSN 1437-4447), SEKI Publications, DFKI Bremen GmbH, Safe and Secure Cognitive Systems, Cartesium, Enrique Schmidt Str. 5, D–28359 Bremen, Germany, 2009. http://arxiv.org/abs/0905.2435.

[5] C. Benzmüller, L. C. Paulson, F. Theiss, and A. Fietzke, 'LEO-II — A Cooperative Automatic Theorem Prover for Higher-Order Logic', in *Proceedings of the 4th International Joint Conference on Automated Reasoning*, eds., P. Baumgartner, A. Armando, and D. Gilles, number 5195 in LNAI, pp. 162–170, Springer, (2008).

[6] C. Benzmüller and A. Pease, 'Progress in automating higher-order ontology reasoning', in *Proceedings of the Second International Workshop on Practical Aspects of Automated Reasoning, Edinburgh, UK, July 14, 2008*, eds., B. Konev, R. A. Schmidt, and S. Schulz, CEUR Workshop Proceedings. CEUR-WS.org, (2010).

[7] P. Hayes and C. Menzel, 'A semantics for the knowledge interchange format', in *IJCAI 2001 Workshop on the IEEE Standard Upper Ontology*, (2001).

[8] K. Hoder, *Automated Reasoning in Large Knowledge Bases*, Master's thesis, Department of Theoretical Computer Science and Mathematical Logic, Charles University, Prague, 2008.

[9] I. Niles and A. Pease, 'Towards a standard upper ontology', in 2nd International Conference on Formal Ontology in Information Systems (FOIS), eds., C. Welty and B. Smith, pp. 2–9, (2001). http://www.adampease.org/professional/FOIS.pdf.

[10] A. Pease. Standard upper ontology knowledge interchange format. http://sigmakee.cvs.sourceforge.net/*checkout*/sigmakee/sigma/suo-kif.pdf.

[11] A. Pease and G. Sutcliffe, 'First order reasoning on a large ontology', in *ESARLT*, eds., G. Sutcliffe, J. Urban, and S. Schulz, volume 257 of *CEUR Workshop Proceedings*. CEUR-WS.org, (2007).

[12] A. Pease, G. Sutcliffe, N. Siegel, and S. Trac, 'Large theory reasoning with SUMO at CASC', *AI Commun.*, **23**(2-3), 137–144, (2010).

[13] D. Ramachandran, P. Reagan, and K. Goolsbey, 'First-orderized ResearchCyc: Expressivity and efficiency in a common-sense ontology', in *Papers from the AAAI Workshop on Contexts and Ontologies: Theory, Practice and Applications*, Pittsburgh, Pennsylvania, (July 2005).

[14] G. Sutcliffe and C. Benzmüller, 'Automated reasoning in higher-order logic using the TPTP THF infrastructure.', *J. of Formalized Reasoning*, **3**(1), 1–27, (2010).