Modal Characterization of a First Order Logic for Topology

Alberto Gatto

Imperial College, London, UK
alberto.gatto@imperial.ac.uk

If one wants to describe topological spaces in first order terms the following language $\mathfrak{L}_2$ is probably one of the most 'natural'. $\mathfrak{L}_2$ is a two-sort first order language: we have first sort variables $x, y, ...$, that are assigned to points, and second sort variables $X, Y, ...$, that are assigned to open sets. $\mathfrak{L}_2$ may be defined over the desired signature of relational and functional symbols. But we always have a symbol $=$, that is interpreted as the equality relation, and a symbol ε, that is interpreted as the set membership relation. $\mathfrak{L}_2$ has the 'usual' boolean connectives, and quantifiers $\forall x$, $\exists x$ for first sort variables and quantifiers $\forall X$, $\exists X$ for second sort variables with the 'usual' meaning. As we would like to characterize (a fragment of) $\mathfrak{L}_2$ in modal terms, we restrict the signature of $\mathfrak{L}_2$ to a countable set $\mathfrak{Prop}$ of unary relation symbols.

If we add some restrictions to $\mathfrak{L}_2$, we obtain the first order language $\mathfrak{L}_t$ of [3, Part 1 §2]. $\mathfrak{L}_t$ is just as $\mathfrak{L}_2$ apart from the definition of second sort quantification. For $\mathfrak{L}_t$, second sort quantification is defined by:

- If φ is positive[1] in X, $\forall X(x \varepsilon X \to \varphi)$ is a formula of $\mathfrak{L}_t$;
- If φ is negative in X, $\exists X(x \varepsilon X \wedge \varphi)$ is a formula of $\mathfrak{L}_t$.

The language $\mathfrak{L}_t$, unlike $\mathfrak{L}_2$, interpreted over topological spaces enjoys 'important' properties 'characterizing' first order logic[2]: compactness and Löwenheim-Skolem Theorem [3, Part 1 §2, 3]. In fact, there is no language for describing topological spaces that is more expressive than $\mathfrak{L}_t$ and enjoys compactness and Löwenheim Skolem Theorem [3, Part 1 §8].

Moreover $\mathfrak{L}_t$ can express 'non-trivial' topological properties: e.g. (among others) T_0, T_1, T_2 and T_3 axioms, triviality, discreteness, etc. (However $\mathfrak{L}_t$ cannot express normality, connectedness and compactness.) (See [3, Part 1 §3].)

Furthermore the $\mathfrak{L}_t$ theory of all T_3 topological spaces is decidable. (However, for $i = 0, 1, 2$, the $\mathfrak{L}_t$ theory of all T_i topological spaces is undecidable, even without unary relations.) (See [3, Part 2 §1].)

Finally, $\mathfrak{L}_t$ is equivalent over topological spaces to the base-invariant fragment of $\mathfrak{L}_2$ [3, Part 1, Theorem 4.19], where 'base-invariance' is defined as follows. Call a *basoid model* every structure $(A, \mathfrak{B})$ where A is a set and $\mathfrak{B}$ is a base for a

[1] An $\mathfrak{L}_2$ formula is *positive* (*negative*) *in a second sort variable* X provided all free occurrences of X are under an even (odd) number of negation signs.

[2] Recall that, according to the Lindström Theorem, first order logic is (roughly) 'the strongest logic (satisfying certain conditions) that enjoys compactness and satisfies the Löwenheim-Skolem Theorem'.

topology over A. Let $\widehat{\mathfrak{B}}$ denote the topology generated by $\mathfrak{B}$. Let us interpret $\mathfrak{L}_2$ over basoid models by interpreting second sort variables as elements of $\mathfrak{B}$. A formula $\varphi(x_1, ..., x_n, X_1, ..., X_m)$ of $\mathfrak{L}_2$ is said *base-invariant* provided for every basoid model $(A, \mathfrak{B})$, $a_1, ..., a_n \in A$ and $O_1, ..., O_m \in \mathfrak{B}$,

$$(A, \mathfrak{B}) \models \varphi[a_1, ..., a_n, O_1, ..., O_m] \text{ iff } (A, \widehat{\mathfrak{B}}) \models \varphi[a_1, ..., a_n, O_1, ..., O_m]. \quad (1)$$

Languages to talk about topological spaces have been defined in modal terms as well. They have a long history (see e.g. the seminal [5]) and there is ongoing interest in the field - e.g. (among others) [2]. The main idea is to associate propositional variables to points of a topological space and give a topologically flavored semantics to modal operators.

We consider the derivative operator $\langle d \rangle$: $\langle d \rangle \varphi$ holds at a point a provided for all opens O containing a there is a point $a' \in O \setminus \{a\}$ where φ holds. Together with $\langle d \rangle$, we consider the graded modalities $\Diamond^n$ (for all $n \in \omega$): $\Diamond^n$ holds at a point a provided there are (at least) n different points at which φ holds. Let $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ denote the modal language in the signature $\{\langle d \rangle, \Diamond^n \mid n \in \omega\}$. We prove that over the class of all T_3 topological spaces $\mathfrak{L}_t$ and $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ are equivalent:

Theorem 1. *The following facts hold:*

1. *For all sentences φ of $\mathfrak{L}_t$ there is a sentence*[3] *$\alpha \in \mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ such that φ and α are equivalent over T_3 models*[4] *- i.e. for all T_3 models $\mathfrak{A}$ we have that $\mathfrak{A} \models \varphi$ if and only if $\mathfrak{A} \models \alpha$.*
2. *For all sentences $\alpha \in \mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ there is a sentence $\varphi \in \mathfrak{L}_t$ such that α and φ are equivalent over T_3 models.*
3. *For all formulas $\varphi(x) \in \mathfrak{L}_t$ there is a formula $\alpha \in \mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ such that $\varphi(x)$ and α are equivalent over T_3 models - i.e. for all T_3 models $\mathfrak{A}$ and points $a \in \mathfrak{A}$ we have that $\mathfrak{A} \models \varphi[a]$ if and only if $\mathfrak{A}, a \models \alpha$.*
4. *For all formulas $\alpha \in \mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ there is a formula $\varphi(x) \in \mathfrak{L}_t$ such that α and φ are equivalent over T_3 models.*

Moreover, there is a computable procedure that translates formulas into equivalent formulas between $\mathfrak{L}_t$ and $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$.

We prove this result by using a game *à la* Ehrenfeucht-Fraïssé.

There are at least two interpretations of this result that are worth mentioning. We can read this result as a van Benthem characterization theorem[5]: over T_3

[3] Call a *sentence of* $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ every formula of $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ of the form $\Diamond^n \psi$ of $\neg \Diamond^n \psi$ ($n \in \omega$). Note that the truth of sentences of $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ does not depend on the point at which they are evaluated.

[4] Call a T_3 *model* every tuple $\mathfrak{A} = (A, \sigma, \{p^{\mathfrak{A}}\}_{p \in \mathfrak{Prop}})$ where (A, σ) is a T_3 topological space and $\{p^{\mathfrak{A}}\}_{p \in \mathfrak{Prop}}$ is the interpretation in $\mathfrak{A}$ of the unary relation symbols in $\mathfrak{Prop}$.

[5] Recall that the van Benthem characterization theorem states (roughly) that basic modal logic is equivalent to the bisimulation invariant fragment of first order logic [1].

topological spaces $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ is the base invariant fragment of $\mathfrak{L}_2$. We can read this result also as a Kamp theorem[6]: over T_3 topological spaces, $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ 'captures' $\mathfrak{L}_t$.

This result opens a number of problems: e.g. (among others) since the $\mathfrak{L}_t$ theory of all T_3 topological spaces is decidable, we have that the $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ theory of all T_3 topological spaces is decidable as well, but what is its complexity? Is there some 'nice' axiomatization of the $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ theory of all T_3 topological spaces (note that this would axiomatize the $\mathfrak{L}_t$ theory of all T_3 topological spaces as well). What is the complexity of translating between $\mathfrak{L}_t$ and $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$? What happens if we replace T_3-ness with other conditions? As a first partial answer we prove that:

Theorem 2. *Over every class of topological spaces including all T_2 topological spaces, we have that $\mathfrak{L}_t$ and $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ are not equivalent.*

The proof uses classical (topo-)bisimulation arguments to show that, unlike in $\mathfrak{L}_t$, T_3-ness is no expressible in $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$. This leads to the following question: what about increasing the expressive power of $\mathfrak{L}_{\langle d \rangle \Diamond^\omega}$ to 'capture' $\mathfrak{L}_t$ over classes including all T_2 topological spaces?

Acknowledgments The author thanks Ian Hodkinson for several useful discussions.

References

1. van Benthem, J. *Modal Correspondence Theory.* PhD Thesis, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.
2. van Benthem, J., Bezhanishvili, G.: Modal Logic of Space. In: Handbook of Spatial Logic. Aiello, M., Pratt-Hartmann, I. E., van Benthem, J. (eds.), 217-298, Springer, 2007.
3. Flum, J., Ziegler, M.: Topological Model Theory (Lecture Notes in Mathematics). Dold, A., Eckmann, B. (eds.), Springer-Verlag, 1980.
4. Kamp, H.: Tense Logic and the Theory of Linear Order. PhD thesis, University of California, Los Angeles, 1968.
5. McKinsey, J., Tarski, A.: The algebra of topology. Ann. Math. 45(1), 141-191, 1944.

[6] Recall that the Kamp theorem states (roughly) that, over the naturals or the reals, the linear temporal logic with 'until' and 'since' is equivalent to first order logic [4].