Via

Kamal Lodaya
The Institute of Mathematical Sciences, Chennai 600113, India

This paper makes a proposal for a logic to describe paths in graphs and suggests another one to compare paths as well.

Keywords: branching temporal logic, computation tree logic, two-variable logic

1. Introduction

That modal logic on Kripke structures translates to two-variable logic has been known for a long time. Lutz, Sattler and Wolter extended modal logic with boolean and converse operations to obtain expressive completeness with respect to two-variable logic on graphs [19]. Marx and de Rijke, building on earlier work on words, showed there is a temporal logic on trees which is expressively complete for two-variable first-order logic [21].

In earlier work on words, we showed that one can extend two-variable first-order logic with "between" relations, and again get an expressively complete temporal logic [15]. We also showed that this fragment has EXPSPACE satisfiability.

In the present paper, we change the setting to directed acyclic graphs. We design a modal logic which is expressively complete for two-variable first-order logic with "via" relations, inspired by our earlier work. We show that this fragment has elementary satisfiability.

We also extend this modal logic to a richer framework with costs. Defining a suitable two-variable logic, so as to extend the above result to obtain expressive completeness, and proving satisfiability of decidability, is left as an open question.

1.1. *Setup*

Suppose I have to go from Bangalore, where I am, to China later this year. I look up various travel websites and discover various facts. For example, there may be a two-hop journey whose fare is cheaper than any one-hop journey. I can go via Delhi where I have several options but then I have to face the long queues at Delhi airport (which I don't care much for), or I can go via southeast Asia where there are longer stopovers.

Graded modal logics [10], logics with nominals (see [7] for an approach due to Arthur Prior, or the later work of Gargov and Goranko [11]), more generally description logics [5], provide facilities for knowledge representation and reasoning over graphs, and I might imagine specifying my requirements to a theorem prover in such a logic.

What if I was planning a longer journey, continuing from China via the Pacific to Canada? Now another such set of requirements comes up. Any logic which has a tree model property will fare badly with such repeated sets of requirements, because there are several branches reaching China and the entire tree of requirements from China to Canada has to be repeated at each China node. I could go to Dubai and then via the Atlantic to Canada, which is my eventual destination, satisfying my shopping urges along this alternative path.

1.2. *Rich modal logics*

Surely this could be done in graded CTL [2, 9, 23] or in dynamic logics, where there is a large literature [8, 13, 20] ? However, these logics all have the tree model property. We use nominals from hybrid logics to bring together paths. Indeed, our approach will be close to hybrid CTL [1, 14, 26]. We have to avoid enriching such logics to force the formation of unbounded-size grids, which leads to high undecidability (for an example from our early work on concurrency, see [17]). We would like our logic to be elementarily decidable.

In his article "Modality, si! modal logic, no!" [22], John McCarthy criticizes modal logic for not providing the richness required for representing human practice, where one sometimes introduces new modalities on an *ad hoc* basis. Some of his criticisms, about other kinds of knowledge than just *knowing that*, are dealt with in recent work of Yanjing Wang [28]. We view the requirement of expanding an edge to a *small* source-sink "via graph" as a similar source of richness which could be added to modal logics. Unlike coalgebraic modal logics [24], this kind of facility is not uniform, but as McCarthy says, on an *ad hoc* basis at a particular node of the model. We did not find such considerations in the literature.

What do we mean by *small* ? The main idea here is that one should not allow recursive specification of arbitrary properties of the logic inside the via graph, for then one allows the kind of generality which one is trying to avoid, such as the formation of complex structures. The way we do this below is to only allow constant structure. It is also possible to allow a hierarchy of structure, but we do not consider that here.

2. Logic

Our frames are directed graphs, although our examples and results will only talk about rooted acyclic ones.

$b ::= p \in Prop \mid o \in Nom \mid \neg b \mid b_1 \vee b_2$
$\pi ::= \textbf{via } b_1 \ldots \textbf{via } b_k \mid \pi_1, \pi_2$
$\alpha ::= b \mid \neg\alpha \mid \alpha \vee \beta \mid \langle\pi\rangle\alpha$

The boolean expressions b are evaluated at possible worlds as in any Kripke structure. They include *nominals* which denote single worlds. We will require that if a vertex is the target of edges from two distinct sources and is also the source of another edge, that vertex is identified by a nominal. Frames of models meeting this requirement are called *via* dags.

The expressions π evaluate to sets of (nontrivial) paths. The expression **via** $b_1 \ldots$ **via** b_k collects the set of all paths between a source and target vertex having intermediate nodes in sequence (not including the source and target) satisfying $b_1, \ldots, b_k$. Thus **via** $false$ specifies a single edge.

The formulae α put all these together to describe graph properties. In particular sets of paths from a source node converge to a common target: we have $w \models \langle\pi_1, \pi_2\rangle\alpha$ iff there is a node $x \models \alpha$ accessible from w through two paths evaluating to π_1 and π_2 respectively.

We can define unary CTL modalities $EX\alpha = \langle\textbf{via } false\rangle\alpha$, $EXEF\alpha = \langle\textbf{via } true\rangle\alpha$, $EF\alpha = \alpha \vee EXEF\alpha$. We will freely use these abbreviations, recall also that $AX\alpha = \neg EX\neg\alpha$ and $AG\alpha = \neg EF\neg\alpha$.

Note that the formula $\langle\pi,\pi\rangle\alpha$ simplifies to $\langle\pi\rangle\alpha$. We could read this as requiring two paths, but that takes us into graded CTL [2, 9]. We duck the issue by not having any multiset requirements. Since there is an easy extension of the logic we will continue to use graded modalities informally.

This buys us a theorem. Let $\mathrm{FO}^2[desc, child, Nom]$ stand for two-variable first-order logic with child (edge) and descendant (transitive closure of edges) relations together with specified nominal unary predicates. By the future fragment of this logic we mean that in a quantifier, say $\exists y\alpha(x, y)$ with free variable x, it is required that y is a descendant of x.

Theorem 2.1. *On* via *dags, our logic is exactly as expressive as the future fragment of two-variable first-order logic $FO^2[desc, child, Nom]$, with* *descendant and child relations and specified nominal unary predicates. Satisfiability is decidable in* EXPTIME.

Proof. We first translate our modalities into a hybrid unary CTL. The path modality $\langle \mathbf{via}\ b \rangle \alpha$ is expressible in CTL as:

$$(b \equiv false \supset EX\alpha) \wedge (b \not\equiv false \supset EXEF(b \wedge EXEF\alpha)).$$

For a longer path from x to y, the specified boolean conditions occur as a subword in sequence, again possibly but not necessarily consecutively. (We thank a referee for pointing out a mistake in our earlier formulation.)

The two-path modality $\langle \pi_1, \pi_2 \rangle \alpha$ can be described using conjunction of the translations for $\langle \pi_1 \rangle \alpha$ and $\langle \pi_2 \rangle \alpha$, but such a CTL formula can have a tree model which is not the intention, here we want the two paths to have a common target node. By using a fresh nominal o, we translate to $\langle \pi_1 \rangle o \wedge \langle \pi_2 \rangle o \wedge AG(o \supset \alpha)$, which is a formula in hybrid CTL [14], where we continue to use unary CTL modalities.

Sattler and Vardi give a translation from hybrid mu-calculus (a richer logic than CTL) to two-way alternating tree automata and an EXPTIME decision procedure for satisfiability [26].

There is an easy translation $Tran(\alpha)$ of a unary CTL formula α into the future fragment of two-variable first-order logic. When we translate a nominal, the locution $AG(o \supset \alpha)$ at a world x translates to the subformula $\forall y\ desc\ x(o(y) \supset Tran(\alpha(y)))$. Because a nominal is interpreted as a single world, we pull out such subformulae as outermost conjuncts which can be seen as enforcing global constraints that we have a *via* dag. The formula is in the future fragment of $\mathrm{FO}^2[desc, child, Nom]$.

The converse direction is an extension of the technique of Marx and de Rijke for the descendant and child axes of Core XPath [21], which itself comes from earlier work. The extension is to specified nominals, and we recall that our formula is interpreted on *via* dags. Every formula of the future fragment of two-variable logic $\mathrm{FO}^2[desc, child, Nom]$ can be put into normal form as a set of formulae, one for the original formula and one for each nominal. Each normal form is a disjunction of conjuncts ranging over order types consisting of children, the descendant relation beyond children and not going beyond specified nominals. Intuitively each normal form describes a tree-like portion of a dag, rooted either at a node identified by a nominal, or at the root of the dag where the original formula holds. For each order type, a corresponding unary CTL formula in the logic is constructed. Finally all these formulae are put together as a single formula of our logic using the nominals, which holds at the root of the *via* dag. $\square$

The main reason we restricted ourselves to the future fragment is that using negation we can talk about two nodes, none of which is a descendant of the other, and we do not know how to extend the Marx-de Rijke proof idea to this case. Some complexity lower bounds we examined require a significant amount of hybrid logic or CTL [14, 26] and do not go through for our logic. Hence it is possible that the complexity is lower, there is a PSPACE lower bound for unary CTL.

Let us now see how the examples which started this paper fare in this logic. A two-stop journey from Bangalore to China can be expressed as $Bangalore \wedge EXEF\ EXEF\ EXEF\ China$. Having several options to go to China at Delhi can be expressed by nominals for Delhi and China (to represent several cities in China we can use a proposition), and using a graded formula to express side properties, as in $\langle \mathbf{via}\ Delhi \rangle China \wedge AG(Delhi \supset EX^{\geq 2}EF\ China)$ or directly in graded CTL as $EXEF(Delhi \wedge EX^{\geq 2}EF\ China)$. To describe my plight as a nervous flier, that the journey from Bangalore to Delhi should not go over any sea, I can use $\neg\langle \mathbf{via}\ \neg land \rangle Delhi$ which turns out equivalent, given that seas exist, to the binary CTL formula $land\ AU\ Delhi$.

2.1. *Costs*

To model fare amounts, lengths of queues and duration of stopovers one could use propositions, or techniques developed for weighted CTL [6, 16]. For example, a simple change in the evaluation of paths is possible, by allowing the syntax to express the **cost** of a path compared to a constant c, which comes from an ordered abelian cancellative semigroup, for simplicity we assume the natural numbers with zero. Note that we make use of subtraction but not of negative values. Thus I can view the long queues at Delhi as a cost I have to bear when I pass through Delhi, but if I use some other intermediate airport, I do not consider the absence of long queues as a negative cost.

Our logic could be modified so that the expressions π now have sets of paths possibly with costs attached to them. Thus $\mathbf{cost} \sim c\ \mathbf{via}\ b_1 \ldots \mathbf{via}\ b_k$ gathers paths between a source and target vertex which additionally specify the cost constraint. The cost of a path is the sum of costs of intermediate edges and vertices. The Rescher **preference** modality [25, 27] compares paths, giving sets of paths in π_1 which are less costly than those in π_2. We require that the preference relation is a strict partial order. The cost is optional, if no cost is stated then we assume no cost is defined for the given path. Similarly, if no preference is provided, there is no preference relation between the two paths. We can ask whether there is a business class fare below a certain cost as well as an economy class fare below another cost.

The formulae α continue to model situations rather than necessarily performing some optimizing computations.

$$
\begin{aligned}
&b ::= p \in \mathit{Prop} \mid o \in \mathit{Nom} \mid \neg b \mid b_1 \vee b_2 \\
&\pi ::= \textbf{cost} \sim c \ \textbf{via}\ b_1 \ldots \textbf{via}\ b_k, \sim \in \{<, \leq, =, \geq, >\} \mid \pi_1, \pi_2 \mid \pi_1 \ \textbf{pref}\ \ \pi_2 \\
&\alpha ::= b \mid \neg\alpha \mid \alpha \vee \beta \mid \langle\pi\rangle\alpha
\end{aligned}
$$

Before we jump into the technical details we would like to ask this question: should one consider satisfiability as not just a matter of assigning truth values to propositions, but also as a matter of assigning costs to propositions in order to make a formula with preferences come out true? This is reminiscent of satisfiability questions in fuzzy [12], deontic [3] and weighted [4] logics. We confess to ignorance of the philosophical implications of adopting such a view. Hence we have not addressed the question of where the basic costs come from, leaving it to future work.

We would like that satisfiability of this extended logic is decidable with elementary complexity. Intuitively one expects that the procedure used for propagation of CTL eventualities through children can be lifted to propagate costs through descendants and nominals as well, after one nondeterministically guesses costs associated with each proposition and adds them up to compute the costs of an intermediate state. One of the referees suggested we could also consider model checking questions for motivation.

Acknowledgements

A first proposal in this direction was made at the end of an early paper on concurrency with Rohit Parikh, R. Ramanujam and P.S. Thiagarajan [17]. Another paper with Ramanujam considered navigation in graphs [18]. The present proposal generalizes work with Andreas Krebs, Paritosh Pandya and Howard Straubing [15], which I presented at the 15th ALC in Daejeon. I would like to thank all these collaborators for their inspiration.

Thanks also to Fenrong Liu and Byeong-uk Yi for the discussions at Daejeon. When Hans van Ditmarsch learned of my interest, he kindly made available his draft paper with Minghui Ma [20]. We thank the referees for encouraging this work.

References

[1] Mario R. Folhadela Benevides and Luis Menasché Schechter. Modal expressiveness of graph properties, *Proc. 2nd LSFA*, Ouro Preto (M. Ayala-Rincón and E.H. Haeusler, eds.), *ENTCS* **205**, 31–47 (2008).

[2] Alessandro Bianco, Fabio Mogavero and Aniello Murano. Graded computation tree logic, *ACM Trans. Comput. Log.* **13**, 3, 25:1–25:53 (2012).

[3] Johan van Benthem, Davide Grossi and Fenrong Liu. Deontics = betterness + priority, *Proc. 10th Deon*, Fiesole (Guido Governatori and Giovanni Sartor, eds.), *LNCS* **6181**, 50–65 (2010).

[4] Benedikt Bollig, Paul Gastin, Benjamin Monmege and Marc Zeitoun. Logical characterization of weighted pebble walking automata, *Proc. 23rd CSL & 29th LICS*, Vienna (Thomas A. Henzinger and Dale Miller, eds.), ACM, 19:1–19:10 (2014).

[5] Ronald J. Brachman and James G. Schmolze. An overview of the KL-ONE knowledge representation system, *Cog. Sci.* **9**, 171–216 (1985).

[6] Peter Buchholz and Peter Kemper. Model checking for a class of weighted automata, *Discr. Ev. Dyn. Syst.* **20**, 1, 103–137 (2010).

[7] Robert A. Bull. An approach to tense logic, *Theoria* **36**, 3, 282–300 (1970).

[8] Hans van Ditmarsch, Wiebe van der Hoek and Barteld Kooi. *Dynamic epistemic logic*, Springer (2008).

[9] Alessandro Ferrante, Margherita Napoli and Mimmo Parente. Model checking for graded CTL, *Fund. Inform.* **96**, 3, 323–339 (2009).

[10] Kit Fine. In so many possible worlds, *Notre Dame J. Formal Logic* **13**, 4, 516–520 (1972).

[11] Geroge Gargov and Valentin Goranko. Modal logic with names, *J. Philos. Logic* **22**, 6, 607–636 (1993).

[12] Petr Hájek and Dagmar Harmancová. A comparative fuzzy modal logic, *Proc. 8th FLAI*, Linz (Erich-Peter Klement and Wolfgang Slany, eds.), *LNCS* **695**, 27–34 (1993).

[13] David Harel, Dexter Kozen and Jerzy Tiuryn. *Dynamic logic*, MIT Press (2000).

[14] Ahmet Kara, Volker Weber, Martin Lange and Thomas Schwentick. On the hybrid extension of CTL and CTL+, *Proc. 34th*

MFCS, Nový Smokovec (RASTISLAV KRÁLOVIČ AND DAMIAN NIWIŃSKI, eds.), *LNCS* **5734**, 427–438 (2009).

[15] ANDREAS KREBS, KAMAL LODAYA, PARITOSH PANDYA AND HOWARD STRAUBING. Two-variable logic with a between relation, *Proc. 31st LICS*, New York (MARTIN GROHE, ERIK KOSKINEN AND NATARAJAN SHANKAR, eds.), ACM/IEEE , 106–115 (2016).

[16] KIM G. LARSEN, RADU MARDARE AND BINGTIAN XUE. Alternation-free weighted mu-calculus: decidability and completeness, *Proc. 31st MFPS*, Nijmegen (DAN R. GHICA, ed.), *ENTCS* **319**, 289–313 (2015).

[17] KAMAL LODAYA, ROHIT PARIKH, R. RAMANUJAM AND P.S. THIAGARAJAN. A logical study of distributed systems, *Inf. Comput.* **115**, 1, 91–115 (1995).

[18] KAMAL LODAYA AND R. RAMANUJAM. An automaton model of user-controlled navigation on the web, *Proc. 5th CIAA*, London (CA) (SHENG YU AND ANDREI PAUN, eds.), *LNCS* **2088**, 208–216 (2001).

[19] CARSTEN LUTZ, ULRIKE SATTLER AND FRANK WOLTER. Modal logic and the two-variable fragment, *Proc. 15th CSL*, Paris (LAURENT FRIBOURG, ed.), *LNCS* **2142**, 247–261 (2001).

[20] MINGHUI MA AND HANS VAN DITMARSCH. Dynamic graded epistemic logic, manuscript (2016).

[21] MAARTEN MARX AND MAARTEN DE RIJKE. Semantic characterizations of navigational XPath, *Sigmod Record* **34**, 2, 41–46 (2005).

[22] JOHN MCCARTHY. Modality, si! modal logic, no! *Studia Logica* **59**, 1, 29–32 (1997).

[23] FARON MOLLER AND ALEXANDER RABINOVICH. Counting on CTL*: on the expressive power of monadic path logic, *Inf. Comput.* **184**, 147–159 (2003).

[24] LAWRENCE S. MOSS. Coalgebraic logic, *Ann Pure Appl. Logic* **96**, 1-3, 277–317 (1999).

[25] NICHOLAS RESCHER. Plurality-quantification and quasi-categorical propositions, *J. Symb. Logic* **27**, 373–374 (1962).

[26] ULRIKE SATTLER AND MOSHE VARDI. The hybrid mu-calculus, *Proc. 1st IJCAR*, Siena (RAJEEV GORÉ, ALEXANDER LEITSCH AND TOBIAS NIPKOW, eds.), *LNAI* **2083**, 76–91 (2001).

[27] ROBERT C. STALNAKER. A theory of conditionals, *Studies in logical theory* (NICHOLAS RESCHER, ed.), Basil Blackwell , 98–112 (1968).

[28] YANJING WANG. A new modal framework for epistemic logic, *Proc. 16th TARK*, Liverpool (JÉRÔME LANG, ed.), *EPTCS* **251**, 515–534 (2017).