Issues in

Algorithmic Model Theory

for

Specific Domains

Martin Otto
RWTH Aachen
otto@informatik.rwth-aachen.de

Trento, 7-99

Outline

Specific Domains

motivation from semantics, examples

Some Issues and Comparative Studies

- syntax vs. semantics:
  characterization theorems
  semantic completeness & complexity
- satisfiability
- recursion mechanisms

focus of results:

modal / non-modal 2-variable / guarded

mathematical modelling

in logic and model theory

---

relevant entities encoded as structures

databases	relational structures
processes (temporal/causal)	graphs transition systems
semantic dependencies (epistemic)	graphs Kripke structures
words	ordered monadic structures
...	

---

if modelling requires
exclusive attention to finite structures
$\longrightarrow$ finite model theory FMT

when two structures encode the same thing

never distinguish between isomorphic structures

isomorphism types $[\mathfrak{A}]^{\simeq}$

$\simeq$ may be too fine, e.g.

process behaviour – bisimulation type $[\mathfrak{A}]^{\sim}$

the modelling task

the semantic domain

$$\text{determine} \left\{ \begin{array}{l} \text{basic class of structures} \\ \text{and notion of equivalence} \end{array} \right.$$

$$\text{to match} \left\{ \begin{array}{l} \text{the right objects} \\ \text{at the right level of abstraction} \end{array} \right.$$

the logic

chosen to respect prescribed equivalence (domain)
to meet expressiveness requirements within

choice of logic subordinate to choice of semantic domain

examples of semantic domains

modal	—	bisimulation eq. Milner,van Benthem	$\sim$
2-variable	—	2-pebble game eq. Barwise,Poizat,Immerman	$\simeq^2$
guarded	—	guarded bisim. eq. Andreka,van Benthem,Nemeti	$\sim^{\mathrm{G}}$

and many others: k-variable,
guarded k-variable,
k-var. with counting,
topological, affine, . . .

domain inclusions

in particular

modal domain — guarded 2-var. domain

guarded 2-var. domain — 2-var. domain

guarded 2-var. domain — guarded domain

some (families of) logics

modal logic, ML

$\Diamond$, $\Box$, $\wedge$, $\vee$, $\neg$

$$\Diamond\varphi(x) \equiv \exists y(Exy \wedge \varphi(y)) \equiv \exists y Exy.\varphi(y)$$
$$\Box\varphi(x) \equiv \forall y(Exy \rightarrow \varphi(y)) \equiv \forall y Exy.\varphi(y)$$

infinitary variant **ML$_\infty$** with infinite $\bigvee$, $\bigwedge$

2-variable first-order logic, FO^2

with variables x,y relational, with $=$

infinitary variant **FO$^2_\infty$**

guarded first-order logic, GF

with only guarded quantification

$$\exists \boldsymbol{y}\big(\alpha(\boldsymbol{x},\boldsymbol{y}) \wedge \varphi(\boldsymbol{x},\boldsymbol{y})\big) \equiv \exists \boldsymbol{y}\alpha(\boldsymbol{x},\boldsymbol{y}).\varphi(\boldsymbol{x},\boldsymbol{y})$$
$$\forall \boldsymbol{y}\big(\alpha(\boldsymbol{x},\boldsymbol{y}) \rightarrow \varphi(\boldsymbol{x},\boldsymbol{y})\big) \equiv \forall \boldsymbol{y}\alpha(\boldsymbol{x},\boldsymbol{y}).\varphi(\boldsymbol{x},\boldsymbol{y})$$

for atoms α with all free variables of φ

infinitary variant **GF$_\infty$**

semantics vs. syntax

algebraic vs. logical equivalences

characterization thms (Ehrenfeucht-Fraïssé style)

Barwise, van Benthem, Poizat, Immerman, ...

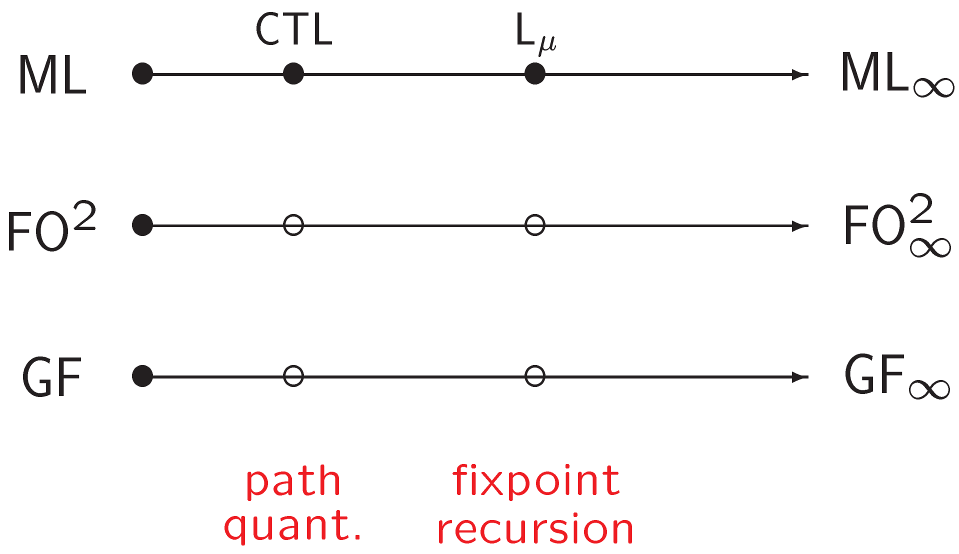
$\equiv^{\mathsf{ML}_\infty} \quad \sim \quad \equiv^{\mathsf{ML}}$ (FIN or ω-sat.)

$\equiv^{\mathsf{FO}^k_\infty} \quad \simeq^k \quad \equiv^{\mathsf{FO}^k}$ (FIN or ω-sat.)

$\equiv^{\mathsf{GF}_\infty} \quad \sim^{\mathsf{G}} \quad \equiv^{\mathsf{GF}}$ (FIN or ω-sat.)

domains come with spectra of logics

$\rightarrow$ one scenario for algorithmic concerns



semantics vs. syntax (ctd.)

FO-fragments

characterization thms (preservation style)

$\mathrm{FO}/\sim$	$\equiv$	ML		van Benthem
$\mathrm{FO}/\sim$	$\equiv$	ML	(FIN)	Rosen
$\mathrm{FO}/\simeq^k$	$\equiv$	FO^k		Barwise,van Benthem
$\mathrm{FO}/\sim^{\mathsf{G}}$	$\equiv$	GF		Andreka,van Benthem,Nemeti

classical proof idea

$\varphi \in$ FO preserved under $\approx^*$ $\quad\Rightarrow\quad$ φ $*$-expressible

show $\{\psi^* : \varphi \models \psi^*\} \models \varphi$ and use compactness

else, upgrade $\mathfrak{A} \equiv^* \mathfrak{B}$ $\quad \mathfrak{A} \models \varphi,\ \mathfrak{B} \not\models \varphi,$ to

$\mathfrak{A}' \approx^* \mathfrak{B}'$ in ω-sat. extns $\quad \#$

semantics vs. syntax (ctd.)

SO-fragments

$$\mathrm{MSO}/\!\sim \;\equiv\; \mathrm{MSO} \restriction \sim\text{-closed} \;\equiv\; \mathrm{L}_\mu$$

Janin, Walukiewicz

$$\mathrm{SO}/\!\sim^{\mathrm{G}} \;\equiv\; \mathrm{SO} \restriction \text{guarded} \;\stackrel{?}{\equiv}\; \mu\mathrm{GF}$$

conjecture/work in progress

$\mathrm{SO} \restriction \simeq^k\text{-closed} \quad \rightarrow$ Dawar

semantic completeness & complexity

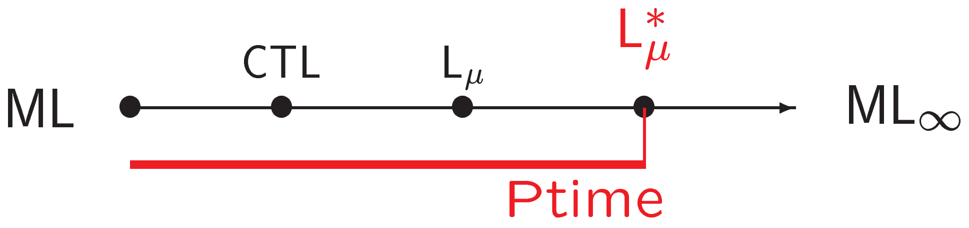
Ptime fragments

semantic completeness for Ptime

capturing results for

Ptime/$\sim$ and Ptime/$\simeq^2$ (O_95,O_97)

e.g.



L^*_μ L_μ with fixed-points in all arities respecting $\sim$ componentwise

undecidable for SAT

NB: orthogonal to capturing Ptime/$\simeq$ over restricted classes

proof via game invariants

replace $\mathfrak{A}$ by $I_{\approx}(\mathfrak{A})$, "its quotient w.r.t. $\approx$"
a structure representing $[\mathfrak{A}]^{\approx}$

to get $\mathrm{Ptime}/\!\approx\; \equiv\; \mathrm{Ptime} \circ \mathbf{I}_{\approx} \;\equiv\; \mathrm{FP}(\mathbf{I}_{\approx})$

e.g. in the modal domain the following coincide

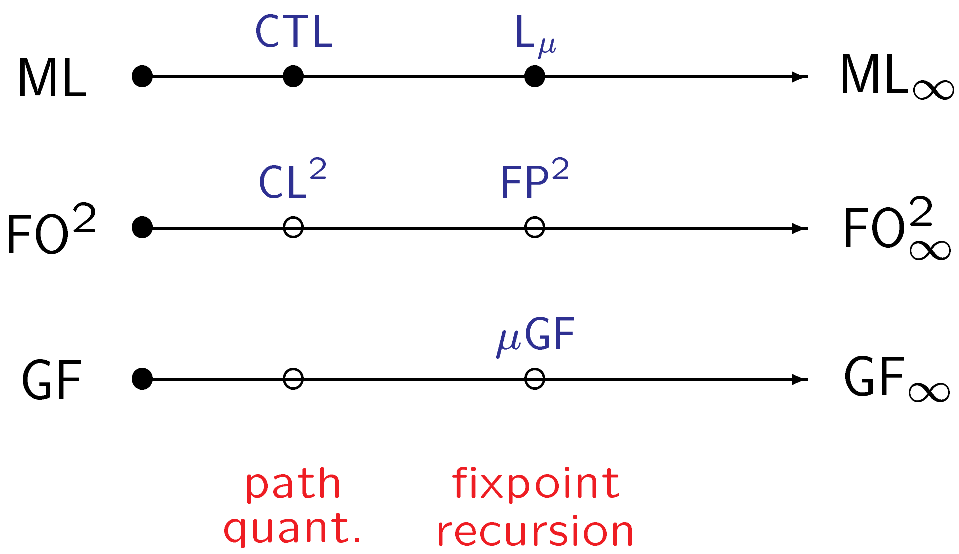
- Ptime/$\sim$
- Ptime on quotients w.r.t. $\sim$
- FP on quotients w.r.t. $\sim$
- L^{*}_{μ} (appropriate syntax)

generalisation to $\sim^{\mathrm{G}}$ unproblematic

limitations for $k > 2$ variables $\rightarrow$ Grohe

satisfiability: SAT and FINSAT

explore borderline in related domains



- understand borderline of decidability
- understand modal robustness (Vardi)

main results

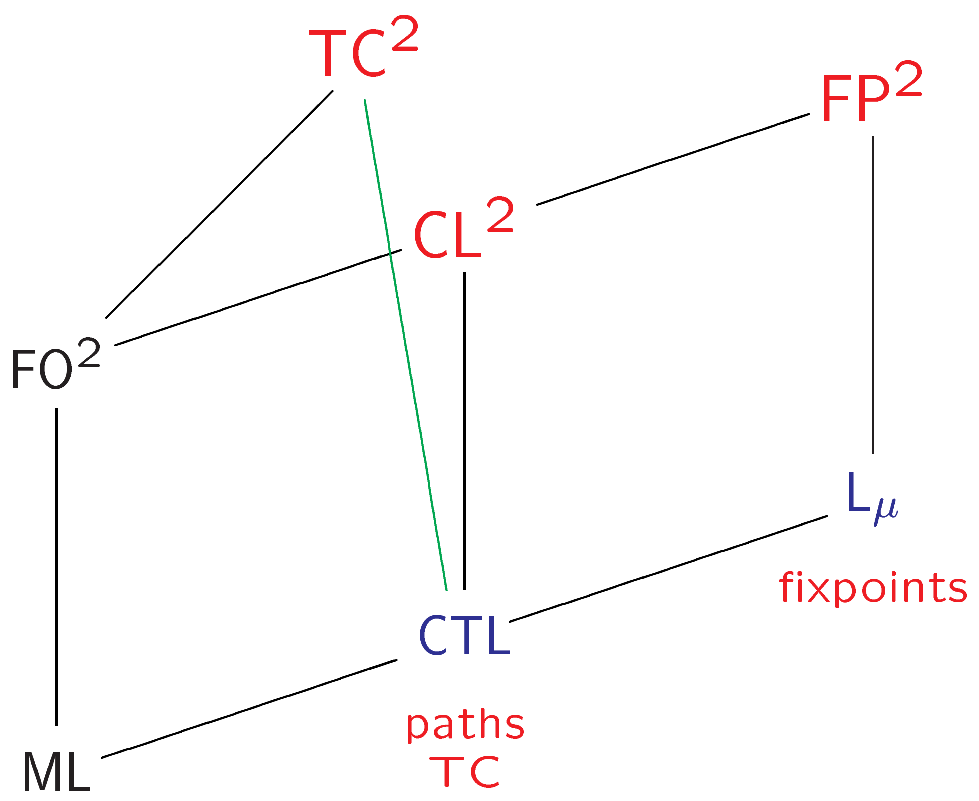
modal robustness not inherited in 2-variable domain:

CL^2 undecidable

inherited by guarded domain:

μGF decidable $\rightarrow$ Grädel, Walukiewicz'99

candidates



FP^2: two-variable fixed-point logic

TC^2: two-variable transitive closure logic

CL^2: two-variable path closure logic

two-variable fixed-point logic $\mathbf{FP^2}$

FP^2 : FO^2 + monadic least fixed points

$\mu_{X,x}\varphi(X,x)$, φ positive in X:

the least P s.t. $P = \{a\colon \varphi(P,a)\}$

two-variable transitive closure logic $\mathbf{TC^2}$

TC^2 : FO^2 + transitive closures $\mathrm{TC}\varphi(x,y)$

$\mathrm{TC}\varphi(x,y)$:

"there is a φ-path from x to y"

two-variable path closure logic $\mathbf{CL^2}$

CL^2 : FO^2 + reachability $\langle\varphi(x,y)\rangle\psi(y)$
and 'dually' $[\varphi(x,y)]\psi(y)$

$\langle\varphi(x,y)\rangle\psi(y)$:

"there is a φ-path from x to a y s.t. $\psi(y)$"

expressiveness

- $\mathrm{CTL} \subseteq \mathrm{TC}^2$ (FIN) (Immerman,Vardi)

  TC^2 advocated for model checking

  not a two-variable logic

  not $\mathrm{CTL} \subseteq \mathrm{TC}^2$ in general

- $\mathrm{CTL} \subsetneq \mathrm{CL}^2 \subsetneq \mathrm{FP}^2$

  $\mathrm{CL}^2 \subsetneq \mathrm{TC}^2$ (FIN)

  CL^2 the least common extn of FO^2 and CTL

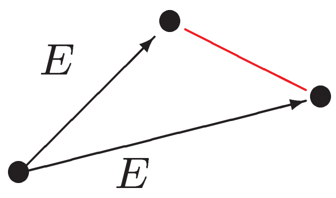
- $\mathrm{L}_\mu \subsetneq \mathrm{FP}^2$

  FP^2 the least common extn of FO^2 and L_μ

expressiveness: examples

matchings in TC^2 (functionality/injectivity):

$\mathrm{dom}(E) \cap \mathrm{im}(E) = \emptyset$ $E \cup E^{-1} \cup (=)$ transitive	$\Rightarrow$	E the graph of a 1-1 partial function



pre-wellorderings in FP^2 (transitivity):

$\prec$ a pre-wellordering with
associated equivalence $\sim$ iff

$$\left.\begin{array}{l} x \sim y \leftrightarrow (\neg x \prec y \wedge \neg y \prec x) \\ \sim \ \text{reflexive} \end{array}\right\} \text{in FO}^2$$

$$\sim \circ \prec \circ \sim \ \text{well-founded}$$

---

well-founded support of $R_1 \circ R_2$ in FP^2:

$$\mu_X \Big[\forall y \Big(R_2 yx \to \forall x (R_1 xy \to Xx) \Big) \Big]$$

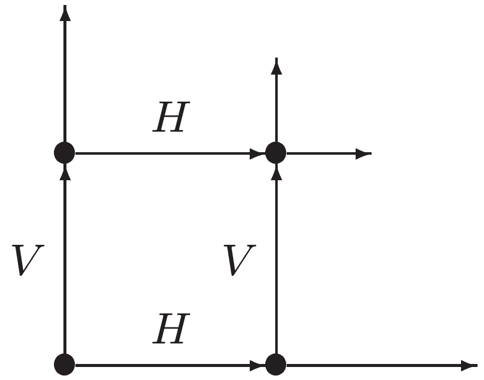
undecidability via dominoes

$\mathcal{D} = (D, R_h, R_v)$ D finite, $R_{h/v} \subseteq D \times D$
domino system hor./vert. adjacency

homomorphism $\mathfrak{G}_{\mathbb{N}} \xrightarrow{\text{hom}} \mathcal{D}$

$$\mathfrak{G}_{\mathbb{N}} = \big(\mathbb{N}\times\mathbb{N}, H, V\big)$$

$\mathcal{D}$-tiling



a constraint satisfaction problem

described by a colouring of $\mathfrak{G}_{\mathbb{N}}$ with $(P_d)_{d\in D}$

homomorphism condition: $\psi_{\mathcal{D}}$ in FO^2
(local) configuration of $\mathfrak{G}_{\mathbb{N}}$: not in FO^2

$$\psi_{\mathcal{D}} = \forall x \dot{\bigvee}_{d\in D} P_d x$$
$$\wedge \forall x \forall y \Big(Hxy \rightarrow \bigvee_{(d,d')\in R_h} (P_d x \wedge P_{d'} y) \Big) \wedge \ \ldots$$

Theorem (Berger'66,Gurevich,Koryakov'72)

D1 : Ex. of $\mathcal{D}$-tilings
D2 : Ex. of finite, periodic $\mathcal{D}$-tilings

undecidable

reduction idea

IF $\psi^*_{\text{grid}} \in$ L forces local grid structure

THEN $\mathcal{D} \longmapsto \psi^*_{\text{grid}} \wedge \psi_{\mathcal{D}}$ reduces

D1/2 $\longrightarrow$ SAT/FINSAT(L)

$\Rightarrow$ undecidability

crux of the reduction : ψ^*_{grid}

$\left.\begin{array}{l}\text{functionality}\\ \text{commutativity}\end{array}\right\}$ of H and V not 2-variable

alternative reduction pattern

towards a domain specific view

---

for suitable classes $\mathcal{C}$:

$\psi^*_{\mathcal{C}} \in \mathsf{L}$	axiomatizing $\mathcal{C}$
$\psi_{\mathsf{grid}} \in \mathsf{FO}^2$	interpreting grids over $\mathcal{C}$

get two reductions:

$$\mathcal{D} \longmapsto \psi^*_{\mathcal{C}} \wedge \psi_{\mathsf{grid}} \wedge \psi_{\mathcal{D}} \qquad \in \mathsf{L}$$

D1/2 $\longrightarrow$ SAT/FINSAT(L)

$$\mathcal{D} \longmapsto \psi_{\mathsf{grid}} \wedge \psi_{\mathcal{D}} \qquad \in \mathsf{FO}^2$$

D1/2 $\longrightarrow$ **SAT/FINSAT(FO^2) over $\mathcal{C}$**

results cf. modal frame conditions

SAT/FINSAT(FO^2) undecidable over $\mathcal{C}$ for the classes $\mathcal{C}$ of $\left(A, E_1, \ldots, E_k, \bar{E}, \bar{P}\right)$

with E_i $\left\{\begin{array}{ll} \text{equivalence relations} & FP^2,\ TC^2 \\ \text{well-founded} & FP^2,\ CL^2 \\ \text{well-orderings} & FP^2,\ CL^2 \\ \text{transitive} & TC^2 \end{array}\right.$

$\Rightarrow$
TC^2, FP^2, CL^2 undecidable for SAT and FINSAT

O_'97;Grädel,O_,Rosen'97;O_'98

SAT/FINSAT(FO^2)
decidable over functional E_i:

$\Rightarrow$ decidability of C^2 Grädel,O_,Rosen'97

recursion mechanisms — the boundedness problem

eliminability of fixed-point recursion in the spectra

$$\mathsf{ML} \bullet\!\!\longrightarrow \overset{\mathsf{L}_\mu}{\bullet} \!\!\longrightarrow \mathsf{ML}_\infty$$

$$\mathsf{FO}^2 \bullet\!\!\longrightarrow \overset{\mathsf{FP}^2}{\bullet} \!\!\longrightarrow \mathsf{FO}^2_\infty$$

inductive generation of least fixed-points

$\varphi(X, x)$ positive in X has least fixed-point $\mu_X \varphi$

$$\mu_X \varphi = \bigcup_\alpha \varphi^\alpha \qquad \varphi^\alpha = \Big\{ \boldsymbol{x} : \varphi\Big[\bigcup_{\beta<\alpha} \varphi^\beta, \boldsymbol{x}\Big]\Big\}$$

inductive stages

$\gamma_{\mathfrak{A}} = \min_\alpha(\varphi^\alpha[\mathfrak{A}] = \varphi^{\alpha+1}[\mathfrak{A}])$ closure ordinal

$\varphi(X, \boldsymbol{x})$ bounded iff uniformly $\gamma_{\mathfrak{A}} \leqslant n$ for some $n \in \mathbb{N}$

Barwise-Moschovakis-Theorem

$\varphi(X, \boldsymbol{x}) \in$ FO bdd $\quad \Leftrightarrow \quad \mu_X\varphi$ FO-definable

$\Rightarrow$: $\quad \varphi^n \in$ FO

$\Leftarrow$: $\quad$ classical model theory: compactness / ω-sat. ($\gamma_{\mathfrak{A}} \leqslant \omega$)

variations for fragments L $\subseteq$ FO

$$\boxed{\varphi(X, \boldsymbol{x}) \in \mathsf{L} \text{ bdd} \;\Leftrightarrow\; \mu_X\varphi \ \mathsf{L}\text{-definable}}$$

holds for $\quad \exists$-FO, $\forall$-FO, FO^k, ML, GF

$$\boxed{\varphi(X, \boldsymbol{x}) \in \mathsf{L} \text{ bdd} \;\Leftrightarrow\; \mu_X\varphi \ \mathsf{L}\text{-def.} \qquad \text{(FIN)}}$$

holds for $\exists$-FO, FO^2, ML, GF
not for $\forall$-FO, FO^3

$\varphi(X, x) =$ "$<$ lin. ord." $\wedge \forall y(y < x \rightarrow Xy)$

unbdd, but $\quad \mu_X\varphi \equiv$ "$<$ lin. ord." (FIN)

boundedness as a decision problem

BD(L)

given: $\varphi(X, \boldsymbol{x}) \in$ L, positive in X
decide: is φ bounded?

BD and SAT

for $\wedge$-closed L with non-trivial BD(L):

SAT(L) $\leqslant$ BD(L)

if $\varphi_0(X, x) \in$ L unbdd, then :
$\psi \wedge \varphi_0(X, x)$ unbdd $\Leftrightarrow$ ψ has infinite models

---

major results concern

BD(Datalog) (query optimization)

Cosmadakis, Gaifman, Hillebrand, Kanellakis, Mairson, Sagiv, Vardi

not candidates like $\exists$-FO, $\forall$-FO, FO^2

undecidability of boundedness in the 2-variable domain

Theorem (Kolaitis,O_'98)

BD(FO^2) undecidable

monadic, no $\neg$, $=$; w.r.t. universal recursion

proof: reduction of the domino tiling problem

König's lemma:

$\mathcal{D}$ tiles $\mathbb{N}\times\mathbb{N}$ $\quad\Leftrightarrow\quad$ $\mathcal{D}$ tiles arbitrarily large $n\times n$

IF stages of $\varphi(X,x)$ cover increasing $n\times n$ grid domains

THEN $\varphi(X,x)\wedge\psi_{\mathcal{D}}$ / $\varphi(X,x)\vee\neg\psi_{\mathcal{D}}$ bdd

$\Leftrightarrow$ $\varphi(X,x)$ bdd in models of $\psi_{\mathcal{D}}$

$\Leftrightarrow$ $\mathcal{D}$ does not tile large $n\times n$ / $\mathbb{N}\times\mathbb{N}$

can use $\varphi(X,x)$ based on well-foundedness recursion

for pre-wellordreings w.r.t. rows and columns

decidability of boundedness in the modal domain

recursion-free μ-calculus

given: $\psi(x) \in \mathsf{L}_\mu$
decide: ex. $\psi_0 \in$ ML, $\psi \equiv \psi_0$?

decidable in Exptime (O_, STACS'99)

proof: tree model property (bisimulation-invariance)
$\rightarrow$ automata theory / S2S

B-M-Thm: $\varphi(X, x) \in$ ML bdd $\Leftrightarrow$ $\mu_X \varphi$ ML-definable

Corollary

BD(ML) decidable

---

open: BD(GF)

summary

robustness of modal domain

tree model property (bisimulation invariance)

generalizes to guarded domain

but not to 2-variable domain

for SAT and BD: **model existence**

$\rightarrow$ automata theory / S2S

some good algorithmic behaviour

shared by all these restrictive locality-oriented domains

in particular: **capturing model checking complexity**

no apparent penalty in modal $\rightarrow$ 2-variable, e.g.

MC(ML) and MC(FO^2) in $\mathcal{O}(|M||\varphi|)$

or, w.r.t. data complexity,

CTL and TC^2 in NLogspace