Complexity of two-variable Dependence Logic and IF-Logic

Jonni Virtema
(joint work with Juha Kontinen, Antti Kuusisto, Peter Lohmann)

University of Tampere

June 23, 2011

Outline

1. Backround
2. Motivation
3. Introducing IF and D
4. Introducing other relevant logics
5. Undecidability of IF^2
6. Decidability of D^2
7. Summary

Background

- The semantics of first order logic can be defined game theoretically by a two player game with perfect information.
- In FO the order in which quantifiers are written determines dependence relations between variables, e.g., in
$$\forall x_1 \exists y_1 \forall x_2 \exists y_2 \psi(x_1, x_2, y_1, y_2)$$
the value chosen for y_1 depends on the value of x_1 and y_2 depends on both x_1 and x_2.
- A natural question that arises is what happens if we allow a richer structure of dependence.

- Leon Henkin (1959) introduced formulas (called Henkin quantifier or Branching quantifier) of the form

$$\begin{pmatrix} \forall x_1 & \exists y_1 \\ \forall x_2 & \exists y_2 \end{pmatrix} \psi(x_1, x_2, y_1, y_2) \tag{1}$$

where y_1 depends on x_1 and y_2 only depends on x_2. Formula (1) is equivalent to the formula

$$\exists f \exists g \forall x_1 \forall x_2 \psi(x_1, x_2, f(x_1), g(x_2))$$

of existential second-order logic ESO.

- It was soon observed that the expressive power of branching quantifiers goes beyond FO. Infact it's equi-expressive to the full existential second order logic.
- The idea of Henkin was developed further by Jaakko Hintikka and Gabriel Sandu (80's) with their *Independence Friendly Logic* (IF). In IF-logic the branching quantifier can be expressed as:

$$\forall x_1 \exists y_1 \forall x_2 \exists y_2 / \{x_1, y_1\} \psi(x_1, x_2, y_1, y_2),$$

  where $\exists y_2 / \{x_1, y_1\}$ means that the choice for the value of y_2 has to be independent of the values of x_1 and y_1.

- The semantics of IF-logic was first defined game theoretically with a two party game of imperfect information.
- In the 90's, Wilfrid Hodges gave a Tarski style truth definition for IF where the basic notion used to define satisfaction is not *assignment s satisfying a formula* as in FO, but *a set X of assignments satisfying a formula*.
- Dependence logic of Jouko Väänänen (2007) adds the concept of dependence to FO in terms of new atomic dependence formulas. In Dependence logic the branching quantifier can be expressed as

$$\forall x_1 \exists y_1 \forall x_2 \exists y_2 (=(x_2, y_2) \wedge \psi(x_1, x_2, y_1, y_2)).$$

Motivation

- Extensions of FO^2.
- Differences in D and IF.

Dependence Logic and IF-logic

Definition (IF-logic)

The syntax of IF extends the syntax of FO defined in negation normal form by adding quantifiers of the form

$$(\exists x/W)\phi$$
$$(\forall x/W)\phi$$

called slashed quantifiers. Here W is a finite set of first order variables.

Dependence logic

Definition

The syntax of D extends the syntax of FO defined in negation normal form by new atomic (dependence) formulas of the form

$$=(x_1, \ldots, x_n).$$

Team semantics

The semantics of D and IF are defined in terms of *Teams* (sets of assignments):

Definition

Let A be a set and $\{x_1, \ldots, x_k\}$ a set of variables. A *team* X of A with domain $\{x_1, \ldots, x_k\}$ is a set of assignments s from $\{x_1, \ldots, x_k\}$ into A.

Semantics for D and IF

Definition

Let $\mathfrak{A}$ be a model and X a team of A. The satisfaction relation $\mathfrak{A} \models_X \phi$ is defined as follows:

1. If ϕ is a first-order literal, then $\mathfrak{A} \models_X \phi$ iff for all $s \in X$: $\mathfrak{A}, s \models_{\mathrm{FO}} \phi$.
2. $\mathfrak{A} \models_X \psi \wedge \phi$ iff $\mathfrak{A} \models_X \psi$ and $\mathfrak{A} \models_X \phi$.
3. $\mathfrak{A} \models_X \psi \vee \phi$ iff there exist teams Y and Z such that $X = Y \cup Z$, $\mathfrak{A} \models_Y \psi$ and $\mathfrak{A} \models_Z \phi$.
4. $\mathfrak{A} \models_X \exists y \psi$ iff $\mathfrak{A} \models_{X(F/y)} \psi$ for some $F \colon X \to A$.
5. $\mathfrak{A} \models_X \forall y \psi$ iff $\mathfrak{A} \models_{X(A/y)} \psi$.

Here $X(F/y) = \{s(F(s)/y) \mid s \in X\}$ and $X(A/y) = \{s(a/y) \mid a \in A, s \in X\}$.

Semantics of D

Definition

6. $\mathfrak{A} \models_X =(x_1, \ldots, x_n)$ iff for all $s, s' \in X$ such that $s(x_1) = s'(x_1), \ldots, s(x_{n-1}) = s'(x_{n-1})$, we have that $s(x_n) = s'(x_n)$.
7. $\mathfrak{A} \models_X \neg{=}(x_1, \ldots, x_n)$ iff $X = \emptyset$.

Semantics of IF

Definition

8. $\mathfrak{A} \models_X \exists y/W\, \phi$ iff $\mathfrak{A} \models_{X(F/y)} \phi$ for some W-independent function $F : X \to A$.
9. $\mathfrak{A} \models_X \forall y/W\, \phi$ iff $\mathfrak{A} \models_{X(A/y)} \phi$.

We say that a function $F : X \to A$ is W-independent if for all $s, s' \in X$ with $s(x) = s'(x)$ for all $x \in \mathrm{dom}(X) \setminus W$ we have that $F(s) = F(s')$.

Properties expressible in D^2 and IF^2

Proposition

The following properties can be expressed in D^2*:*

1. *For unary relation symbols* P *and* Q, D^2 *can express* $|P| = |Q|$. *This shows that* $D^2 \not\leq FO$.
2. *If the vocabulary of* $\mathfrak{A}$ *contains a constant* c, *then* D^2 *can express that* A *is infinite.*
3. $|A| \leq k$ *can be already expressed in* D^1.

Comparison of D^2 and IF^2

Theorem

$D^2 \leq IF^2 \leq D^3$

Proof.

The claim follows by relatively straightforward translations. $\square$

Other relevant logics

1. FO^2, two-variable first order logic,
2. FOC^2, two-variable first order logic with counting,
3. $\mathrm{FO}^2(\mathrm{I})$, two-variable first order logic with the Härtig quantifier $\mathrm{I}xy(\phi(x), \psi(y))$
4. ESO, existential second order logic.

Satisfiability problem

Definition

Let $\mathcal{L}$ be a logic. The satisfiability problem SAT[$\mathcal{L}$] is the following problem:
Input: a sentence $\phi \in \mathcal{L}$.
Output: **Yes**, if there is a model $\mathfrak{A}$ such that $\mathfrak{A} \models \phi$, and **No** otherwise.

The finite satisfiability problem FINSAT[$\mathcal{L}$] is the version of the above question in which $\mathfrak{A}$ must also be finite.

Some complexity results

Logic	*Complexity of SAT / FINSAT*	*References*
FO, FO^3	Π^0_1 / Σ^0_1	[Chu36, Tur36]
ESO, D, IF	Π^0_1 / Σ^0_1	[Chu36, Tur36]
FO^2	*NEXPTIME*	[GKV97]
FOC^2	*NEXPTIME*	[PH05]
$FO^2(\mathrm{I})$	Σ^1_1-hard / in Σ^0_1	[GOR97]
D^2	*NEXPTIME*	[LICS 2011]
IF^2	Π^0_1 / Σ^0_1	[LICS 2011]

Complexity of SAT(IF^2) and FINSAT(IF^2)

Theorem (LICS 2011)

SAT(IF^2) is Π^0_1-complete.

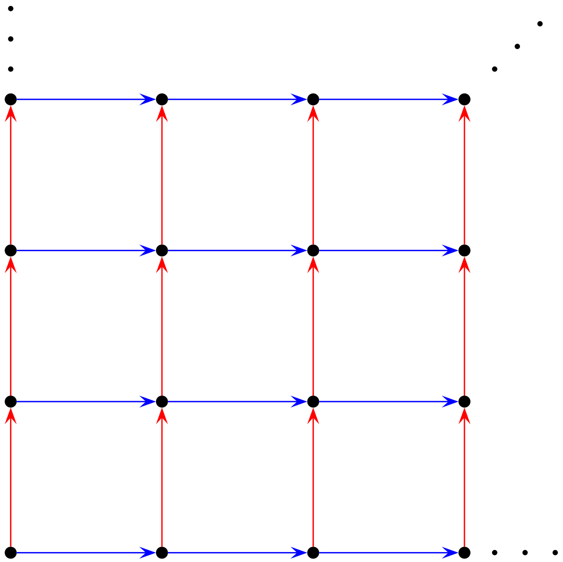
Theorem (LICS 2011)

FINSAT(IF^2) is Σ^0_1-complete.

Tiling

- A tile is a square whose each side is assigned a color, i.e., it is a square that has four colors (up, right, down, left).
- A set of tiles T can tile a model $\mathfrak{A} = (A, V, H)$ with two binary relations V and H if a tile can be placed on every point in the domain A s.t
  1. for all pairs of points $(a, b) \in H$ the right color of the tile on a is the same as the left color on the tile on b and
  2. for all pairs of points $(a, b) \in V$ the top color of the tile on a is the same as the bottom color on the tile on b.

The grid



where $H(\rightarrow)$, $V(\rightarrow)$.

Undecidability of tiling problems

Tiling problem for a fixed model $\mathfrak{A}$ is the following problem: Given a set of tiles T can T tile the model $\mathfrak{A}$. We denote this problem as Tiling($\mathfrak{A}$).

Theorem ([Ber66])

Tiling($\mathfrak{G}$), where $\mathfrak{G}$ is the $\mathbb{N} \times \mathbb{N}$ grid is Π^0_1-complete problem.

Hardness

Π_1^0 hardness follows from the following lemma:

Lemma

For every set of tiles T we have IF^2 *formula* γ_T *s.t* γ_T *is satisfiable iff T can tile the* $\mathbb{N} \times \mathbb{N}$ *grid.*

Expressing tiling (the easy part)

Given a set of tiles T it easy to write an FO^2 sentence ϕ_T s.t T tiles a model $\mathfrak{A} = (A, V, H)$ iff there exists $\mathfrak{A}^*$, an extension of $\mathfrak{A}$ with some unary relation symbols, s.t $\mathfrak{A}^* \models \phi_T$.

Expressing grid-likeness (a bit harder part)

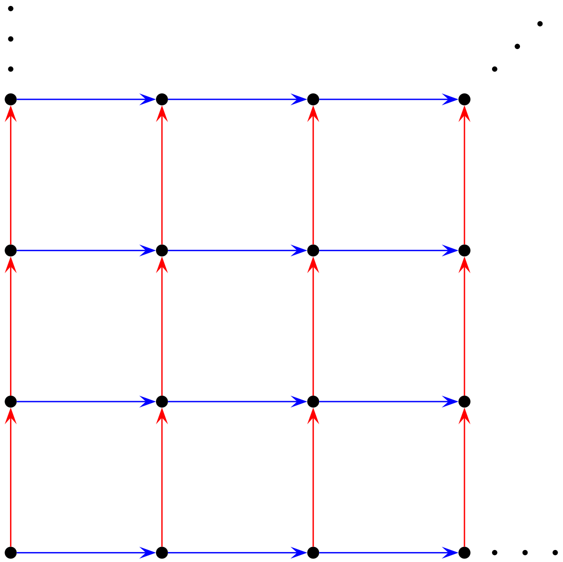
The problem lies in expressing that a model is an infinite grid or something close enough.

Grid-likeness

We use the following properties to say that a structure (A, V, H) is grid-like:

1. V and H are graphs of injective functions.
2. There exists a root of the grid.
3. $V \cap H = \emptyset$
4. Borders of the grid are constructed correctly.
5. Amalgamation property for V and H hold.
6. The grid is infinite.

The grid



where $H(\rightarrow)$, $V(\rightarrow)$.

Expressing amalgamation

In the formula ϕ_{grid} the key ingredient is to express the following property:

Property

For all points x there exists a point y s.t.

$$x(V \circ H)y \text{ and } x(H \circ V)y.$$

Sentence

We use the following IF^2 sentence to mimic the above property, note that they are not equivalent

$$\forall x \forall y \Big(\big(V(x,y) \vee H(x,y) \big) \rightarrow \exists x/\{y\} \, \big(V(y,x) \vee H(y,x) \big) \Big).$$

Lemma

For every set of tiles T we have IF^2 formula γ_T s.t γ_T is satisfiable iff T can tile the $\mathbb{N} \times \mathbb{N}$ grid.

Theorem

SAT(IF^2) is Π^0_1-hard.

In Π^0_1

Since SAT(ESO) is Π^0_1-complete and there is a polynomial translation from IF into ESO, it follows that SAT(IF^2) is in Π^0_1.

SAT/FINSAT(D^2) is decidable

Theorem ([GKV97])

SAT(FO^2) and FINSAT(FO^2) are NEXPTIME-complete.

Theorem ([PH05])

SAT(FOC^2) and FINSAT(FOC^2) are NEXPTIME-complete.

Hence SAT(Σ_1^1 (FOC^2)) and FINSAT(Σ_1^1 (FOC^2)) are *NEXPTIME*-complete.

SAT/FINSAT(D^2) is decidable

Since D^2 is a conservative extension of FO^2. And there exists a polynomial translation from D^2 to Σ_1^1 (FOC^2) [LICS 2011] it follows.

Theorem (LICS 2011)

SAT(D^2) and FINSAT(D^2) are NEXPTIME-complete.

Expressing functionality

The key ingredient in translating D^2 into Σ_1^1 (FOC^2) is to express the dependence atom with two variables using counting quantifiers. We use the following translation:

$$=(x,y) \longmapsto \forall x \exists^{\leq 1} y R(x,y)$$

where R is binary relation correspoding to a team.

Separation of D^2 and IF^2

As a by-product of the complexity results we obtain the following result concerning expressivity of the finite variable logics:

Theorem (LICS 2011)

$D^2 < IF^2 \leq D^3$

Clonclusion

Open questions:

- Complexity of the validity problem for the logics D^2 and IF^2.
- Is it possible to define NP-complete problems in D^2 or IF^2?

Clonclusion

Open questions:

- Complexity of the validity problem for the logics D^2 and IF^2.
- Is it possible to define NP-complete problems in D^2 or IF^2?
  - Yes, the dominating set problem is quite easy to express already in D^2.

Robert Berger.
The undecidability of the domino problem.
Number 66 in Memoirs of the American Mathematical Society. American Mathematical Society, 1966.

Alonzo Church.
A note on the entscheidungsproblem.
Journal of Symbolic Logic, 1(1):40–41, 1936.

Erich Grädel, Phokion G. Kolaitis, and Moshe Y. Vardi.
On the decision problem for two-variable first-order logic.
The Bulletin of Symbolic Logic, 3(1):53–69, 1997.

Erich Grädel, Martin Otto, and Eric Rosen.
Undecidability results on two-variable logics.
In *STACS '97: Proceedings of the 14th Annual Symposium on Theoretical Aspects of Computer Science*, pages 249–260, London, UK, 1997. Springer-Verlag.

Ian Pratt-Hartmann.
Complexity of the two-variable fragment with counting quantifiers.
J. of Logic, Lang. and Inf., 14:369–395, June 2005.

Alan Turing.
On computable numbers, with an application to the entscheidungsproblem.
Proceedings of the London Mathematical Society, Series 2, 42:230–265, 1936.