

# On the expressive power of IF-logic with classical negation

Santiago Figueira<sup>\*1,3</sup>, Daniel Gorín<sup>1</sup>, and Rafael Grimson<sup>2,3</sup>

<sup>1</sup> Dto. Computación, FCEN, Universidad de Buenos Aires, Argentina

<sup>2</sup> Dto. Matemática, FCEN, Universidad de Buenos Aires, Argentina

<sup>3</sup> CONICET, Argentina

**Abstract.** It is well-known that Independence Friendly (IF) logic is equivalent to existential second-order logic ( $\Sigma_1^1$ ) and, therefore, is not closed under classical negation. The boolean closure of IF sentences, called Extended IF-logic, on the other hand, corresponds to a proper fragment of  $\Delta_2^1$ . In this paper we consider IF-logic extended with Hodges' flattening operator, which allows classical negation to occur also under the scope of IF quantifiers. We show that, nevertheless, the expressive power of this logic does not go beyond  $\Delta_2^1$ . As part of the proof, we give a prenex normal form result and introduce a non-trivial syntactic fragment of full second-order logic that we show to be contained in  $\Delta_2^1$ .

## 1 Introduction

*Independence Friendly* logic [4] (IF, for short) is an extension of first-order logic (FO) where each disjunction and each existential quantifier may be decorated with denotations of universally-quantified variables, as in:

$$\forall x \forall y \exists z \bigvee_{| \forall y} \exists w \bigvee_{| \forall y} [y = z \vee \bigvee_{| \forall x, \forall y} w = y] \quad (1)$$

The standard interpretation of IF is through a variation of the classical game-theoretical semantics for FO: Eloïse's strategy function for a position of the form  $\exists x \bigvee_{| \forall y, \forall z} \psi$  or  $\psi \bigvee_{| \forall y, \forall z} \chi$ , under valuation  $v$ , cannot depend on neither  $v(y)$  nor  $v(z)$ . Thus, we say that a sentence  $\varphi$  is *true* in model  $\mathcal{A}$  (notation,  $\mathcal{A} \models^+ \varphi$ ) if Eloïse has a winning strategy on the associated game; and that it is *false* (notation,  $\mathcal{A} \models^- \varphi$ ) whenever Abélard has a winning strategy.

Now, the fact that Eloïse's strategy may not take into account all the available information turns the game into one of "imperfect information". In fact, it is a non-determined game, i.e, for certain instances of the game, no winning strategy exists. As an example of non-determinacy, consider this formula:

$$\chi_1 := \forall x \exists y \bigvee_{| \forall x} x \neq y \quad (2)$$

It is not hard to see that if  $\mathcal{A}$  is a model with at least two elements, then  $\mathcal{A} \not\models^+ \chi_1$  and  $\mathcal{A} \not\models^- \chi_1$ . That is,  $\chi_1$  is neither *true* nor *false* in  $\mathcal{A}$ .

---

\* S. Figueira was partially supported by CONICET Argentina (grant PIP 370).

In game-theoretical semantics, negation is interpreted as a switch of roles, i.e., Abélard plays on Eloïse’s former positions and vice versa. We use  $\sim$  to denote this form of negation and we refer to it as *game negation*. For any IF-formula  $\psi$  and any model  $\mathcal{A}$ ,  $\mathcal{A} \models^+ \psi$  iff  $\mathcal{A} \models^- \sim\psi$ . (i.e., Eloïse has a winning strategy for  $\psi$  on  $\mathcal{A}$  iff Abélard has one for  $\sim\psi$  on  $\mathcal{A}$ ). However, observe that  $\psi \vee \sim\psi$  is not in general a valid IF-formula (e.g., take  $\psi$  to be  $\chi_1$  in (2)). This means that game negation in IF is not equivalent to *classical negation*, which will be denoted with  $\neg$  and is characterized by:

$$\mathcal{A} \models^+ \neg\psi \text{ iff } \mathcal{A} \not\models^+ \psi \tag{3}$$

Since the expressive power of IF corresponds to that of existential second-order logic ( $\Sigma_1^1$ ) [4,5] and  $\Sigma_1^1$  is not closed by (classical) negation, it is clear that classical negation cannot be defined in IF.

Classical negation plays an important role in Hintikka’s original programme. In [4], he claims that “virtually all off classical mathematics can in principle be done in extended IF first-order logic” (in a way that is ultimately “reducible” to plain IF logic). What he calls “(truth-functionally) extended IF logic” is the closure of the set of IF-sentences with operators  $\neg$ ,  $\wedge$  and  $\vee$ . Clearly, extended IF logic corresponds in expressive power to the boolean closure of  $\Sigma_1^1$ , which is known to be a proper fragment of  $\Delta_2^1$  [2,9].

In this paper we will consider the extension of IF logic where classical negation can be combined with IF-operators without restrictions (in contrast to Extended IF, where formulas with  $\neg$  are obtained from boolean combination of IF sentences). One might suspect the resulting logic to be extremely expressive: freely combining classical negation with second order existential quantifiers leads to full second-order logic (SO). This suspicion is true for the logic with Henkin quantifiers when the set of available quantifiers is closed by a dualization operator [10]. Indeed, this logic has the same expressive power as SO. We will show that, in our context, this is not the case: IF with unrestricted classical negation is not more expressive than  $\Delta_2^1$ .

We take as a basis Hodges’ treatment of classical negation [6]. He defines an extension of IF, called *slash logic* (SL), in which independence restrictions can occur in any connective (and not only  $\exists$  and  $\vee$ ) and provide a compositional semantics for it. To support classical negation, the *flattening* operator  $\downarrow$  is introduced. Intuitively, this operator satisfies:

$$\mathcal{A} \models^- \downarrow\varphi \text{ iff } \mathcal{A} \not\models^+ \varphi \tag{4}$$

It is easy to see that classical negation can be expressed combining  $\sim$  and  $\downarrow$ :

$$\mathcal{A} \models^+ \neg\varphi \text{ iff } \mathcal{A} \models^+ \sim\downarrow\varphi \text{ iff } \mathcal{A} \models^- \downarrow\varphi \text{ iff } \mathcal{A} \not\models^+ \varphi \tag{5}$$

Let  $\text{SL}(\downarrow)$  denote the extension of SL with the flattening operator. We will use the game-theoretical semantics introduced in [3] (equivalent to Hodges’ compositional semantics), which we present in §2. From this we derive a prenex normal form result in §3 from which a Skolem form for  $\text{SL}(\downarrow)$ -formulas is obtained in §4. By analyzing the syntactic fragment of SO in which the Skolem form falls, we ultimately arrive to the  $\Delta_2^1$  upper-bound.

## 2 Syntax and semantics of $\text{SL}(\downarrow)$

We assume a fixed first-order relational language  $\mathcal{L}$ , as well as a collection  $X$  of first-order variables, which we will denote  $x, y, z$ , perhaps with subindices. Formulas of  $\text{SL}(\downarrow)$ , in negation normal form, correspond to the following grammar:

$$\varphi ::= l(x_1, \dots, x_k) \mid \exists x_{i|\rho} \varphi \mid \forall x_{i|\rho} \varphi \mid \downarrow \varphi \mid \uparrow \varphi \mid \varphi \vee_{|\rho} \varphi \mid \varphi \wedge_{|\rho} \varphi \quad (6)$$

where  $\rho$  denotes a (possibly empty) finite set of variables and  $l(x_1, \dots, x_k)$  is any first-order literal (i.e., an atom or a negated atom). We will typically use  $\exists x_i, \forall x_i, \vee$  and  $\wedge$  instead of  $\exists x_{i|\emptyset}, \forall x_{i|\emptyset}, \vee_{|\emptyset}$  and  $\wedge_{|\emptyset}$ . Since we are working in negation normal form, (game) negation  $\sim$  will be a mapping on functions satisfying  $\sim \forall x_{i|\rho} \varphi = \exists x_{i|\rho} \sim \varphi$ ;  $\sim \downarrow \varphi = \uparrow \sim \varphi$ , etc. Finally,  $\neg \varphi$  will be short for  $\sim \downarrow \varphi$ .  $\text{Fv}(\varphi)$  denotes the set of free variables of  $\varphi$  (see [3] for a formal definition). A *sentence* is a formula with no free variables.

*Remark 1.* For the sake of simplicity we will impose a further restriction on formulas: there can be no nested bindings of the same variable (e.g.,  $\exists x \exists x \varphi$ ) nor a variable that occurs both free and bound in a formula (e.g.,  $x \vee \exists x \varphi$ ). This is called the *regular fragment* of  $\text{SL}(\downarrow)$  and it has simpler formal semantics. The results in this paper apply to the whole language under the proviso that *history-preserving valuations* are used instead of standard ones (cf. [3] for details).

We interpret a  $\text{SL}(\downarrow)$ -formulas using first-order models  $\mathcal{A}$  with domain  $|\mathcal{A}|$ . We use sets of *finite* valuations to account for free variables; the domains of these valuations must be large enough to interpret them all (but they can be larger).

**Definition 1.** *Given  $\varphi$  and  $\mathcal{A}$ , we say that,  $V$ , a set of finite valuations over  $\mathcal{A}$ , is suitable for  $\varphi$  iff there is a finite set  $D \supseteq \text{Fv}(\varphi)$  such that  $V \subseteq |\mathcal{A}|^D$ .*

We define now the game  $\text{G}(\mathcal{A}, \varphi, V)$ , where  $\mathcal{A}$  is a model and  $V$  is a set of finite valuations over  $\mathcal{A}$  suitable for  $\varphi$ . Furthermore, let  $\tilde{D} \supseteq \text{Fv}(\varphi)$  be such that  $V \subseteq |\mathcal{A}|^{\tilde{D}}$ . As is customary, this game is played between two opponents: Eloïse and Abélard (sometimes called *Verifier* and *Falsifier*). There is also a third agent, called Nature, which acts either as a generator of random choices or as a referee.

*The board.* Game  $\text{G}(\mathcal{A}, \varphi, V)$  is played over the syntactic tree of  $\varphi$ . There is, additionally, a placeholder for a set  $D$  and a valuation  $v : D \rightarrow |\mathcal{A}|$ . Initially,  $D$  is set to  $\tilde{D}$  and  $v$  is empty. In the syntactic tree of  $\varphi$ , all the  $\exists, \vee$  and  $\downarrow$ -nodes of the tree *belong* to Eloïse; while the  $\forall, \wedge$  and  $\uparrow$ -nodes belong to Abélard. Moreover,  $\exists, \forall, \vee$  and  $\wedge$ -nodes will be (repeatedly) decorated with functions during the game; the first two admit any function  $f : |\mathcal{A}|^D \rightarrow |\mathcal{A}|$ ; the last two, only functions  $f : |\mathcal{A}|^D \rightarrow \{L, R\}$ . Initially, nodes have no decoration.

*The turns.* At any point of the game, the remaining number of turns is bounded by the maximum number of nested occurrences of  $\downarrow$ -nodes and  $\uparrow$ -nodes in the game-board.

- *The opening turn.* The first turn is different from the rest. It is composed of two clearly distinguished phases. In the first phase, both players decorate all their nodes with suitable functions. The order in which they tag their nodes is not important as long as they do not get to see their opponent’s choices in advance. For simplicity, we will assume they both play simultaneously. In the second phase, Nature picks a valuation from  $V$  and puts it in the placeholder  $v$  and finally *evaluates* the outcome of the turn, as described below.
- *The subsequent turns.* In all but the first turn, the formula tree is of the form  $\downarrow\psi$  or  $\uparrow\psi$  (see next). In these turns, both players get to redecorate their nodes, one after the other; Eloïse goes first when the formula tree is of the form  $\downarrow\psi$  and Abélard does so on  $\uparrow\psi$ . Finally, Nature replaces the tree with  $\psi$  and proceeds to evaluate.

The recursive evaluation procedure used by Nature is the following:

- R1** If the tree is of the form  $\psi_1 \vee_{|y_1, \dots, y_k} \psi_2$  or  $\psi_1 \wedge_{|y_1, \dots, y_k} \psi_2$ , then its root must have been with a function  $f : |\mathcal{A}|^D \rightarrow \{L, R\}$ . Nature picks elements  $a_1 \dots a_k$  from  $|\mathcal{A}|$  and evaluates  $f(v[y_1 \mapsto a_1, \dots, y_k \mapsto a_k])$  —by construction  $\{y_1, \dots, y_k\} \subseteq D$ . That is, the values the player was not supposed to consider are randomly replaced prior to evaluating the function provided. The tree is then updated with  $\psi_1$ , if the result is  $L$ , and with  $\psi_2$ , otherwise.  $D$  and  $v$  remain unchanged and evaluation proceeds.
- R2** If the tree is of the form  $\exists x_{|y_1, \dots, y_k} \psi$  or  $\forall x_{|y_1, \dots, y_k} \psi$ , then it must have been decorated with a function  $f : |\mathcal{A}|^D \rightarrow |\mathcal{A}|$ . Again, Nature picks  $a_1 \dots a_k$ , evaluates  $b := f(v[y_1 \mapsto a_1, \dots, y_k \mapsto a_k])$  and records this choice by replacing the placeholder with  $D := D \cup \{x\}$  and  $v := v \cup \{x \mapsto b\}$ . Finally, the tree is updated with  $\psi$  and evaluation proceeds.
- R3** If the tree is of the form  $\downarrow\psi$  or  $\uparrow\psi$ , the evaluation —and, thus, the turn— ends.
- R4** Finally, if the root of the tree is a literal  $l(x_1, \dots, x_k)$ , the game ends. Eloïse is declared the winner if  $\mathcal{A} \models l(x_1, \dots, x_k)[v]$ ; otherwise, Abélard wins.

Nodes may get redecorated during the game but only by its owner, that is fixed. Hence it is equivalent to assume that players decorate only those nodes that are not under nested  $\downarrow$  or  $\uparrow$ . This way, each node gets decorated only once.

*Winning strategies.* We will not go into a formal description of what a strategy for  $G(\mathcal{A}, \varphi, V)$  is. We simply take it to be a form of oracle that tells the player how to proceed in each turn. As usual, a strategy is said to be *winning* for a player if it guarantees that the he or she will win every match of the game, regardless the strategy of the opponent and the choices made by Nature.

We are now ready to give our game-semantic notion of *truth* and *falsity*.

**Definition 2.** *Let  $V$  be a set of finite valuations suitable for  $\varphi$ . We define:*

- $\mathcal{A} \models^+ \varphi[V]$  iff Eloïse has a winning strategy for the game  $G(\mathcal{A}, \varphi, V)$ ;
- $\mathcal{A} \models^- \varphi[V]$  iff Abélard has a winning strategy for the game  $G(\mathcal{A}, \varphi, V)$ .

When  $V = \{v\}$  we may alternatively write  $\mathcal{A} \models^+ \varphi[v]$  and  $\mathcal{A} \models^- \varphi[v]$ . Also, for a *sentence*  $\varphi$  we may write  $\mathcal{A} \models^+ \varphi$  and  $\mathcal{A} \models^- \varphi$  meaning  $\mathcal{A} \models^+ \varphi[\emptyset]$  and  $\mathcal{A} \models^- \varphi[\emptyset]$ , respectively, where  $\emptyset$  is the empty valuation (i.e.  $\text{dom } \emptyset = \emptyset$ ). Unlike classical logics, from  $\mathcal{A} \models^+ \varphi[\emptyset]$  we cannot infer  $\mathcal{A} \models^+ \varphi[v]$  for every suitable  $v$ . This is due to *signaling*: the value of a variable a player is supposed not to know is available through the value of another one (cf. [7,8]).

**Definition 3.** We write  $\varphi_1 \equiv \varphi_2$  whenever, for every  $\mathcal{A}$  and every set  $V$  suitable for  $\varphi_1$  and  $\varphi_2$ ,  $\mathcal{A} \models^+ \varphi_1[V]$  iff  $\mathcal{A} \models^+ \varphi_2[V]$  and  $\mathcal{A} \models^- \varphi_1[V]$  iff  $\mathcal{A} \models^- \varphi_2[V]$ .

When interested in winning strategies for, say, Eloïse in  $\mathbf{G}(\mathcal{A}, \varphi, \{v\})$ , it makes no difference whether in the opening turn both players play simultaneously or if Eloïse goes first. This is not true, though, in  $\mathbf{G}(\mathcal{A}, \varphi, V)$  for a non-singleton  $V$ .

**Proposition 1.** Given  $v$ , a finite valuation over  $\mathcal{A}$  suitable for  $\varphi$ , we have that  $\mathcal{A} \models^+ \varphi[v]$  iff  $\mathcal{A} \models^+ \downarrow\varphi[v]$ , and that  $\mathcal{A} \models^- \varphi[v]$  iff  $\mathcal{A} \models^- \uparrow\varphi[v]$ .

*Remark 2.* Whenever one is interested in whether  $\mathcal{A} \models^+ \varphi[V]$  holds or not, it is clearly equivalent (and sometimes convenient, too) to consider a simplified version of  $\mathbf{G}(\mathcal{A}, \downarrow\varphi, V)$  in which Eloïse plays functions and Abélard plays *elements* (until the game reaches a  $\uparrow$ , where the situation gets reversed). This resembles the perfect-information game for IF given by Väänänen in [11].

Operator  $\downarrow$  turns a formula that may lead to a non-determined game, into one that always leads to a determined one. This suggests the following notion.

**Definition 4.** We say that  $\varphi$  is determined whenever, for every model  $\mathcal{A}$ , and every set  $V$  suitable for  $\varphi$ ,  $\mathcal{A} \not\models^+ \varphi[V]$  iff  $\mathcal{A} \models^- \varphi[V]$ .

Intuitively, determined formulas are those that have a well-defined truth-value on every structure. Plain first-order formulas (i.e., those with no independence restrictions) are determined, but one can give more general conditions.

**Proposition 2.** The following hold:

1. Every literal is a determined formula.
2.  $\downarrow\psi$  and  $\uparrow\psi$  are determined formulas.
3. If  $\varphi$  and  $\psi$  are determined formulas, so are  $\varphi \wedge_{|\emptyset} \psi$ ,  $\varphi \vee_{|\emptyset} \psi$ ,  $\exists x_{|\emptyset} \varphi$  and  $\forall x_{|\emptyset} \varphi$ .

### 3 Normal forms for $\mathbf{SL}(\downarrow)$

Normal forms in the context of SL were initially investigated in [1]. Later, Janssen [7] observed some anomalies which cast doubt on the correctness of these results. However, it was shown in [3] that only the formal apparatus employed in [1] was defective, and not the results per se.

In this section we revisit the prenex normal form results of [1] and extend them to account for  $\downarrow$  and  $\uparrow$ . For this, bound variables will be tacitly renamed when necessary<sup>4</sup> and the following formula manipulation tools will be employed.

<sup>4</sup> While this assumption was considered problematic in the context of [1], it is safe here since we are using regular formulas. Moreover, this can also be assumed for arbitrary formulas under an adequate formalization (cf. Remark 1).

**Definition 5.** Let  $x_1 \dots x_n$  be variables not occurring in  $\varphi$ ; we denote with  $\varphi_{|x_1 \dots x_n}$  the formula obtained by adding  $x_1 \dots x_n$  as restrictions to every quantifier, every conjunction and every disjunction in  $\varphi$ . Also, we write  $\varphi^c$  for the formula obtained by replacing all independence restrictions in  $\varphi$  by  $\emptyset$ .

Notice that  $\varphi^c$  is essentially a FO formula. As is observed in [1], independence restrictions on boolean connectives can be removed by introducing additional quantifications. It is not hard to extend this result to  $\text{SL}(\downarrow)$  (in what follows, we shall use, for emphasis,  $\vee_{|\emptyset}$  instead of  $\vee$ ,  $\wedge_{|\emptyset}$  instead of  $\wedge$ , etc.).

**Theorem 1.** For every formula  $\varphi$ , there exists a  $\varphi'$  such that  $\varphi \equiv \varphi'$  and every disjunction (resp. conjunction) in  $\varphi'$  is of the form  $\psi_1 \vee_{|\emptyset} \psi_2$  (resp.  $\psi_1 \wedge_{|\emptyset} \psi_2$ ).

*Proof.* When restricted to models with at least two elements, a simple inductive argument gives us the desired formula. The important step is that, given a formula  $\psi := \psi_1 \vee_{|x_1 \dots x_k} \psi_2$  and given  $y_1, y_2$  fresh in  $\psi$ , we can define

$$\psi^* := \exists y_{1|x_1 \dots x_k} \exists y_{2|y_1, x_1 \dots x_n} [(y_1 = y_2 \wedge \psi_{1|y_1, y_2}) \vee (y_1 \neq y_2 \wedge \psi_{2|y_1, y_2})] \quad (7)$$

On models with at least two elements, we have  $\psi \equiv \psi^*$ . By successively applying this truth-preserving transformation in a top-down manner, one can obtain, for any given  $\varphi$ , a formula  $\varphi^*$  that is equivalent on models with at least two elements.

Now, observe that on models with exactly one element, restrictions are meaningless. Therefore, for any given  $\varphi$  we can define the equivalent formula:

$$\varphi' := (\forall x \forall y [x = y] \wedge \varphi^c) \vee (\exists x \exists y [x \neq y] \wedge \varphi^*) \quad (8)$$

Formula (7) in the above proof was taken from [1], with the proviso that independences on  $y_1$  and  $y_2$  are added to  $\psi_1$  and  $\psi_2$  in order to avoid unwanted signaling [7,8,3]. This was most probably an involuntary omission in [1].

The prenex normal form result for SL in [1] uses a lemma that we will need.

**Lemma 1 ([1]).** If  $x$  does not occur in  $\psi$ , then the following equivalences hold:

1.  $\exists x_{|\rho} [\varphi] \vee_{|\emptyset} \psi \equiv \exists x_{|\rho} [\varphi \vee_{|\emptyset} \psi_{|x}]$ .
2.  $\exists x_{|\rho} [\varphi] \wedge_{|\emptyset} \psi \equiv \exists x_{|\rho} [\varphi \wedge_{|\emptyset} \psi_{|x}]$ .
3.  $\forall x_{|\rho} [\varphi] \vee_{|\emptyset} \psi \equiv \forall x_{|\rho} [\varphi \vee_{|\emptyset} \psi_{|x}]$ .
4.  $\forall x_{|\rho} [\varphi] \wedge_{|\emptyset} \psi \equiv \forall x_{|\rho} [\varphi \wedge_{|\emptyset} \psi_{|x}]$ .

For  $\text{SL}(\downarrow)$ , we need to show how to extract  $\downarrow$  and  $\uparrow$  from arbitrary formulas.

**Lemma 2.** The following hold:

1. If  $\psi$  is determined, then  $\downarrow \psi \equiv \uparrow \psi \equiv \psi$ .
2.  $\downarrow(\varphi \wedge_{|\emptyset} \psi) \equiv \downarrow \varphi \wedge_{|\emptyset} \downarrow \psi$  and  $\uparrow(\varphi \wedge_{|\emptyset} \psi) \equiv \uparrow \varphi \wedge_{|\emptyset} \uparrow \psi$ .
3.  $\downarrow(\varphi \vee_{|\emptyset} \psi) \equiv \downarrow \varphi \vee_{|\emptyset} \downarrow \psi$  and  $\uparrow(\varphi \vee_{|\emptyset} \psi) \equiv \uparrow \varphi \vee_{|\emptyset} \uparrow \psi$ .

*Proof.* We shall only discuss  $\downarrow(\varphi \wedge_{|\emptyset} \psi) \equiv \downarrow \varphi \wedge_{|\emptyset} \downarrow \psi$ ; the remaining equivalences are similar. Suppose, then, that  $\mathcal{A} \models^+ \downarrow(\varphi \wedge_{|\emptyset} \psi)[V]$ ; this means that Eloïse has a way of decorating both  $\varphi$  and  $\psi$  that guarantees she wins both games. Therefore, we have  $\mathcal{A} \models^+ \downarrow \varphi[V]$  and  $\mathcal{A} \models^+ \downarrow \psi[V]$  which implies  $\mathcal{A} \models^+ (\downarrow \varphi \wedge_{|\emptyset} \downarrow \psi)[V]$ . The right to left direction is analogous, and one establishes that  $\mathcal{A} \models^+ \downarrow(\varphi \wedge_{|\emptyset} \psi)[V]$  iff  $\mathcal{A} \models^+ (\downarrow \varphi \wedge_{|\emptyset} \downarrow \psi)[V]$ . Since  $\downarrow(\varphi \wedge_{|\emptyset} \psi)$  and  $(\downarrow \varphi \wedge_{|\emptyset} \downarrow \psi)$  are determined formulas (Proposition 2), this implies  $\mathcal{A} \models^- \downarrow(\varphi \wedge_{|\emptyset} \psi)[V]$  iff  $\mathcal{A} \models^- (\downarrow \varphi \wedge_{|\emptyset} \downarrow \psi)[V]$ .

**Definition 6.** A  $SL(\downarrow)$ -formula is said to be in prenex normal form if it is of the form  $Q_0^* \downarrow_1 Q_1^* \downarrow_2 \dots Q_{n-1}^* \downarrow_{n-1} Q_n^* \varphi$  with  $n \geq 0$ , where each  $Q_i^*$  is a (perhaps empty) chain of quantifiers,  $\downarrow_i \in \{\downarrow, \uparrow\}$  and  $\varphi$  contains only literals,  $\wedge_{|\emptyset}$  and  $\vee_{|\emptyset}$ .

**Theorem 2.** For every  $\varphi$ , there exists a  $\varphi^*$  in prenex normal form with  $\varphi \equiv \varphi^*$ .

*Proof.* By Theorem 1 we can obtain a  $\varphi'$  such that  $\varphi' \equiv \varphi$  and no boolean connective in it contains independences. We proceed now by induction on  $\varphi'$ . If  $\varphi'$  is a literal,  $\varphi^* = \varphi'$ . If  $\varphi' = \exists x_{|y_1 \dots y_k} \psi$ , we have  $\varphi^* = \exists x_{|y_1 \dots y_k} \psi^*$  and the cases for  $\varphi' = \forall x_{|y_1 \dots y_k} \psi$ ,  $\varphi' = \downarrow \psi$  and  $\varphi' = \uparrow \psi$  are analogous. We analyze now the case for  $\varphi' = \psi \vee \chi$ ; the one for  $\varphi' = \psi \wedge \chi$  is symmetrical.

We need to show that there exists a  $\varphi^* \equiv (\psi^* \vee \chi^*)$ , in prenex normal form. We do it by induction on the sum of the lengths of the prefixes of  $\psi^*$  and  $\chi^*$ . The base case is trivial; for the inductive case we show that one can always “extract” the outermost operator of either  $\psi^*$  or  $\chi^*$ .

The first thing to note is that if  $\psi^* = Qx_{|y_1 \dots y_k} \psi'$  ( $Q \in \{\forall, \exists\}$ ), then using Lemma 1 (renaming variables, if necessary) we have  $\varphi^* := Qx_{|y_1 \dots y_k} (\psi' \vee \chi^*)^*$  and the same applies to the case  $\chi^* = Qx_{|y_1 \dots y_k} \chi'$ . So suppose now that neither  $\psi^*$  nor  $\chi^*$  has a quantifier as outermost operator. In that case, they start with one of  $\downarrow$  or  $\uparrow$ , or they contain only  $\wedge_{|\emptyset}$ ,  $\vee_{|\emptyset}$  and literals. In either case, they are both determined and at least one of them starts with  $\downarrow$  or  $\uparrow$  (or we would be in the base case). If we assume that  $\psi^* = \downarrow \psi'$ , using Proposition 2 repeatedly, we have  $(\downarrow \psi' \vee \chi^*) \equiv (\downarrow \psi' \vee \downarrow \chi^*) \equiv \downarrow (\psi' \vee \downarrow \chi^*) \equiv \downarrow (\psi' \vee \chi^*)$ , and we can apply the inductive hypothesis. The remaining cases are analogous.

## 4 Skolem forms, dependencies and $\Delta_2^1$ -expressivity

We finally turn our attention to the expressive power of  $SL(\downarrow)$ . We shall see that every  $SL(\downarrow)$ -formula is equivalent to both  $\Sigma_2^1$  and  $\Pi_2^1$ -formulas, and therefore, is included in  $\Delta_2^1$ . We reserve letters  $f$ ,  $g$  and  $h$  (probably with subindices) to denote second-order functional variables; arities will be left implicit. We identify first-order variables with 0-ary second-order variables; letters  $x$ ,  $y$  and  $z$  (with subindices) are to be interpreted always as 0-ary functions ( $f$ ,  $g$ , etc. could be 0-ary too, unless stated). We also assume that functional symbols occur always fully-saturated (e.g.,  $f = g$  is not a valid formula, but  $\forall x[f(x) = g(x)]$  is).

### 4.1 Skolem form of $SL(\downarrow)$ sentences

Based on the game-theoretical semantics of §2 we show that every *sentence* of  $SL(\downarrow)$  can be turned into an equivalent SO formula. We will motivate this sort of Skolem form with a short example. For this, let  $\varphi$  be quantifier-free, with variables among  $\{x_1, x_2, x_3, y_1, y_2, y_3\}$  and consider:

$$\chi_2 := \downarrow \forall y_1 \forall y_2 \exists x_1 |_{y_2} \uparrow \exists x_2 |_{y_2} \exists x_3 \forall y_3 |_{x_3} [\varphi] \quad (9)$$

Assume that  $\mathcal{A} \models^+ \chi_2$ , i.e., that Eloïse has a winning strategy for  $\mathbf{G}(\mathcal{A}, \chi_2, \{\emptyset\})$ . Using the simplification of Remark 2 this is the case if and only if  $\mathcal{A} \models \chi'_2$ , where

$$\chi'_2 := \exists f \forall y_1 \forall y_2 \forall z_1 \forall g \exists x_2 \exists x_3 \exists z_2 [\varphi \sigma_1] \quad (10)$$

and  $\sigma_1 = \{x_1 \mapsto f(y_1, z_1), y_3 \mapsto g(y_1, y_2, f(y_1, z_1), x_2, z_2)\}$  is a substitution of variables by terms<sup>5</sup>. Notice that  $z_1$  and  $z_2$  represent the random choices made by Nature during the evaluation phases; e.g.,  $f(y_1, z_1)$  expresses that Nature replaced the value of  $y_2$  by a randomly picked  $z_1$  when evaluating  $x_1$ . Since  $z_1$  and  $z_2$  do not occur in  $\varphi$  and  $y_1$  and  $y_2$  occur universally quantified, just as  $g$ , we have that  $\chi'_2$  is equivalent to  $\chi''_2$ , where

$$\chi''_2 := \exists f \forall y_1 \forall y_2 \forall g \exists x_2 \exists x_3 [\varphi \sigma_2] \quad (11)$$

and  $\sigma_2 = \{x_1 \mapsto f(y_1), y_3 \mapsto g(x_2, f(y_1))\}$ . Of course, one could simplify further and replace  $g(x_2, f(y_2))$  by  $g(x_2)$ , but this will be discussed in more detail in §4.2.

In order to formalize this transformation, we will use some conventions. First,  $\lambda$  denotes an empty sequence (of quantifiers, of variables, etc.). When describing  $\text{SL}(\downarrow)$  prefixes we shall use patterns such as  $\downarrow \forall y_1 |_{\tau_1} \exists x_1 |_{\rho_1} \dots \forall y_k |_{\tau_k} \exists x_k |_{\rho_k} \uparrow Q$ ; it must be understood that not necessarily all the  $x_i$  and  $y_i$  are present in the prefix, and that either  $\uparrow Q = \lambda$  or else  $\uparrow \in \{\downarrow, \uparrow\}$  and  $Q$  is a (possibly empty)  $\text{SL}(\downarrow)$ -prefix.

**Definition 7.** *Given  $Q\psi$  in prenex normal form ( $\psi$  quantifier-free),  $T(Q\psi)$  is the SO-formula  $\pi(\downarrow Q)[\psi\sigma]$ , where  $\sigma = \tau_\lambda^\lambda(\{\}, \downarrow Q)$ ,  $\pi(\lambda) = \lambda$ ,  $\tau_\beta^\alpha(\sigma, \lambda) = \sigma$  and*

$$\begin{aligned} \pi(\downarrow \forall y_1 |_{\tau_1} \exists x_1 |_{\rho_1} \dots \forall y_k |_{\tau_k} \exists x_k |_{\rho_k} \uparrow Q) &= \exists f_1 \dots \exists f_k \forall y_1 \dots \forall y_k \pi(\uparrow Q) \\ \pi(\uparrow \exists x_1 |_{\rho_1} \forall y_1 |_{\tau_1} \dots \exists x_k |_{\rho_k} \forall y_k |_{\tau_k} \uparrow Q) &= \forall g_1 \dots \forall g_k \exists x_1 \dots \exists x_k \pi(\uparrow Q) \\ \tau_\beta^\alpha(\sigma, \downarrow \forall y_1 |_{\tau_1} \exists x_1 |_{\rho_1} \dots \forall y_k |_{\tau_k} \exists x_k |_{\rho_k} \uparrow Q) &= \tau_{\beta'}^{\alpha'}(\sigma \cup \{x_i \mapsto f_{x_i}(\alpha' \setminus \rho_i)\sigma\}, \uparrow Q) \\ \tau_\beta^\alpha(\sigma, \uparrow \exists x_1 |_{\rho_1} \forall y_1 |_{\tau_1} \dots \exists x_k |_{\rho_k} \forall y_k |_{\tau_k} \uparrow Q) &= \tau_{\beta'}^{\alpha'}(\sigma \cup \{y_i \mapsto g_{y_i}(\beta' \setminus \tau_i)\sigma\}, \uparrow Q) \end{aligned}$$

Here, we assumed  $\alpha' := \alpha, y_1 \dots y_k$  and  $\beta' := \beta, x_1 \dots x_k$ .

The reader should verify that, modulo variable names,  $T(\chi_2) = \chi''_2$ . In particular, substitution application in  $f_{x_i}(\alpha' \setminus \rho_i)\sigma$  and  $g_{y_i}(\beta' \setminus \tau_i)\sigma$  account for the introduction of nested terms like  $g(x_2, f(y_1))$  in (11).

**Lemma 3.** *For every sentence  $\varphi$  in prenex normal form and every model  $\mathcal{A}$ ,  $\mathcal{A} \models^+ \varphi$  iff  $\mathcal{A} \models T(\varphi)$ .*

*Proof.* First, observe that  $\mathcal{A} \models^+ \varphi$  iff  $\mathcal{A} \models^+ \downarrow \varphi$  (Proposition 1). One then can show that, for every  $\psi$  in prenex normal form (perhaps with free variables) and every suitable  $v$ ,  $\mathcal{A} \models^+ \downarrow \psi[v]$  iff  $\mathcal{A} \models T(\downarrow \psi)[v]$  by induction on the number of turns in  $\mathbf{G}(\mathcal{A}, \downarrow \psi, \{v\})$  (i.e., in the number of  $\downarrow$  and  $\uparrow$  occurring in  $\downarrow \psi$ ).

It follows directly from Definition 7 that  $T(\varphi)$  is a SO formula in prenex normal form (i.e., all quantifiers, either first or second-order, are at the beginning).

<sup>5</sup> As is customary, we use postfix notation for substitution application.

## 4.2 First-order dependencies and the $\Delta_2^1$ -class

If  $\varphi$ , in prenex normal form, has  $d$  occurrences of  $\downarrow$  or  $\uparrow$ , then  $T(\varphi)$  is a formula in  $\Sigma_{2d+1}^1$ . But we will see that by repeatedly applying a truth-preserving quantifier-moving transformation, one can turn any such  $T(\varphi)$  into an equivalent formula in  $\Sigma_2^1$ . Since  $\text{SL}(\downarrow)$  is closed by negation, this gives us the desired  $\Delta_2^1$ -bound. We begin with a motivational example. Let  $\varphi$  be quantifier-free and consider:

$$\chi_3 := \downarrow \forall y_1 \exists x_1 \uparrow \exists x_2 \forall y_2 |_{x_1} \downarrow \forall y_3 \exists x_3 |_{y_1, y_2} \uparrow \exists x_4 \forall y_4 |_{x_1 \dots x_3} \varphi \quad (12)$$

$$\chi'_3 := T(\chi_3) = \exists f_1 \forall y_1 \forall g_1 \exists x_2 \exists f_2 \forall y_3 \forall g_2 \exists x_4 [\varphi \sigma_3] \quad (13)$$

where  $\sigma_3 = \{x_1 \mapsto f_1(y_1), y_2 \mapsto g_1(x_2), x_3 \mapsto f_2(y_3), y_4 \mapsto g_2(x_4)\}$ . The key intuition to show that  $\chi'_3$  is a  $\Sigma_2^1$ -formula is that the choice for  $f_2$  does not actually depend on the chosen value for  $g_1$  but only on that of  $g_1(x_2)$ , which is the only occurrence of  $g_1$  in  $\varphi \sigma_3$ . Thus,  $\chi'_3$  is equivalent to:

$$\chi_4 := \exists f_1 \exists \tilde{f}_2 \forall y_1 \forall g_1 \forall y_3 \exists x_2 \forall g_2 \exists x_4 [\varphi \sigma_4] \quad (14)$$

where  $\sigma_4 \{x_1 \mapsto f_1(y_1), y_2 \mapsto g_1(x_2), x_3 \mapsto \tilde{f}_2(y_1, g_1(x_2), y_3), y_4 \mapsto g_2(x_4)\}$ . Finally, one can “exchange”  $g_2$  and  $x_2$  in an analogous way, to obtain a  $\Sigma_2^1$ -formula.

We identify next the fragment of SO where this “weak-dependency” between functions occurs and show that, in it, the procedure sketched above can always be performed. In what follows we assume, without loss of generality, that if a variable occurs free in a SO-formula, it does not also appear bound.

**Definition 8.** *We say that a functional symbol  $f$  is strongly free in a SO-formula  $\varphi$  whenever  $f$  is free in  $\varphi$  and, if  $f(\dots g(\dots) \dots)$  occurs in  $\varphi$ ,  $g$  is strongly free in  $\varphi$  too.*

As an example,  $g$  is strongly free in  $\exists x[f(x) = g(y, g(z, y))]$ , while  $f$  is not (for  $x$  is not, either). Every free first-order variable is also strongly free.

**Lemma 4.** *Let  $\varphi$  be a SO-formula and let  $g_1 \dots g_k$  be strongly free in  $\varphi$ . Moreover, let  $v_1$  and  $v_2$  be interpretations of functional variables in  $\mathcal{A}$  such that (i)  $v_1(f) = v_2(f)$  for every  $f \notin \{g_1, \dots, g_k\}$ , and (ii) for every  $g_i(t_1, \dots, t_m)$  occurring in  $\varphi$ ,  $v_1(g_i(t_1, \dots, t_m)) = v_2(g_i(t_1, \dots, t_m))$ . Then,  $\mathcal{A} \models \varphi[v_1]$  iff  $\mathcal{A} \models \varphi[v_2]$ .*

*Proof.* Follows by a straightforward induction on  $\varphi$ . In the base case, (i) and (ii) guarantee the equivalence; for the case  $\varphi = \exists f \psi$ , removing the existential preserves the hypothesis of the lemma (needed for the inductive hypothesis).

It is well-known that  $\forall x_1 \dots x_n \exists f \varphi$  is equivalent to  $\exists \tilde{f} \forall x_1 \dots x_n \tilde{\varphi}$ , where  $\tilde{\varphi}$  is obtained by replacing every occurrence of a term of the form  $f(t_1, \dots, t_k)$  in  $\varphi$  by  $\tilde{f}(t_1, \dots, t_k, x_1, \dots, x_n)$ . The following result is a generalization of this idea, with strongly free second-order variables used instead of first-order ones.

**Theorem 3.** *Let  $g_1 \dots g_n$  be strongly free in  $\varphi$  and let  $h$ , free in  $\varphi$ , be such that for no  $i$ ,  $g_i(\dots h(\dots) \dots)$  occurs in  $\varphi$ . Then, for every  $f_1 \dots f_m$  free in  $\varphi$ , there exists  $\tilde{\varphi}$  such that  $g_1 \dots g_n$  are strongly free in  $\tilde{\varphi}$ ;  $f_1 \dots f_m$  are free in  $\tilde{\varphi}$  and  $\forall g_1 \dots \forall g_n \exists h \exists f_1 \dots \exists f_m \varphi \equiv \exists h \forall g_1 \dots \forall g_n \exists f_1 \dots \exists f_m \tilde{\varphi}$ .*

*Proof.* One can obtain  $\tilde{\varphi}$  by replacing every occurrence of  $h(t_1, \dots, t_k)$  in  $\varphi$  by  $\tilde{h}(t_1, \dots, t_k, s_1, \dots, s_l)$ , where  $s_1 \dots s_l$  are the terms of the form  $g_i(\dots)$  occurring in  $\varphi$ . Since  $g_i(\dots h(\dots) \dots)$  does not occur in  $\varphi$ , no occurrence of  $h$  is left in  $\tilde{\varphi}$ .

In a way, what Theorem 3 says is that a strongly free second-order variable corresponds, in terms of *information*, to a finite number of first-order terms. This motivates the following definition.

**Definition 9.** *We say that a SO-formula in prenex normal form has first-order dependencies if in every subformula of the form  $\forall g_1 \dots \forall g_n \exists f_1 \dots \exists f_m \varphi$  (with  $\varphi \neq \exists h\psi$ ) or  $\exists g_1 \dots \exists g_n \forall f_1 \dots \forall f_m \varphi$  ( $\varphi \neq \forall h\psi$ ),  $g_1 \dots g_n$  are strongly free in  $\varphi$ .*

**Theorem 4.** *The class of SO-formulas with first-order dependencies is in  $\Delta_2^1$ .*

*Proof.* We shall only see that if  $\varphi$  has first-order dependencies then  $\varphi$  is in  $\Sigma_2^1$ . This suffices since  $\varphi \equiv \neg\neg\varphi$  and  $\neg\varphi$  has first-order dependencies (and is therefore in  $\Sigma_2^1$ ). To see that any such  $\varphi$  can be turned to  $\Sigma_2^1$ -form, we will show how quantifiers can be reordered, one at a time, in a top-down manner.

Suppose, then, that  $\varphi$  is not in  $\Sigma_2^1$ -form; and let us consider the leftmost quantifier that is misplaced. There are two cases we need to consider. First, suppose  $\varphi = \exists h_1 \dots \exists h_l \forall g_1 \dots \forall g_n \exists x_1 \dots \exists x_k \exists h \exists f_1 \dots \exists f_m \psi$  ( $k, l, m \geq 0, n \geq 1$ ) for an  $h$  of *non-zero arity*. We can assume without loss of generality that  $g_i(\dots h(\dots) \dots)$  does not occur in  $\psi$ , e.g., if  $g_i(h(x_j))$  occurs in  $\psi$ , then  $\varphi$  is equivalent to  $\exists h_1 \dots \exists h_l \forall g_1 \dots \forall g_n \exists x_1 \dots \exists x_k \exists h \exists z \exists f_1 \dots \exists f_m [z = h(x_j) \wedge \chi]$  where  $\chi$  is like  $\psi$  with  $h(x_j)$  replaced by  $z$ . By Theorem 3 we can relocate the misplaced  $\exists h$  and obtain the equivalent formula  $\exists h_1 \dots \exists h_l \exists \tilde{h} \forall g_1 \dots \forall g_n \exists x_1 \dots \exists x_k \exists f_1 \dots \exists f_m \tilde{\psi}$ .

The second case is when every variable in the first block of misplaced quantifiers is first-order. Here we cannot always circumvent the condition “ $g_i(\dots x_j \dots)$  does not occur in  $\psi$ ” so we must handle it in a different way. Therefore, suppose  $\varphi = \exists h_1 \dots \exists h_l \forall g_1 \dots \forall g_n \exists x_1 \dots \exists x_k \forall f_1 \dots \forall f_m \psi$ , ( $l \geq 0, k, m, n \geq 1$ ) where  $\psi$  does not start with  $\forall$ . Using Theorem 3 ( $x_1 \dots x_k$  are strongly free in  $\psi$ ), we obtain the equivalent  $\exists h_1 \dots \exists h_l \forall g_1 \dots \forall g_n \forall \tilde{f}_1 \dots \forall \tilde{f}_m \exists x \dots \exists x_k \tilde{\psi}$ . It is important to notice that  $\forall g_1 \dots \forall g_n \forall \tilde{f}_1 \dots \forall \tilde{f}_m \exists x \dots \exists x_k \tilde{\psi}$  has first-order dependencies.

**Corollary 1.**  *$T(\varphi)$  has first-order dependencies, for every  $SL(\downarrow)$ -formula  $\varphi$  in prenex normal form. Therefore,  $SL(\downarrow)$  is not more expressive than  $\Delta_2^1$ .*

## 5 Conclusions

We have given an upper bound for the expressive power of  $SL(\downarrow)$  and, therefore, for IF with full-fledged classical negation. This logic contains Extended-IF, which is known to correspond to a proper fragment of  $\Delta_2^1$ ; hence, our  $\Delta_2^1$ -upper bound for  $SL(\downarrow)$  is quite tight. This result is mounted on three additional results: *i*) a game-theoretical semantics for  $SL(\downarrow)$ , *ii*) a prenex normal form result for  $SL(\downarrow)$  and, *iii*) a characterization of a new (to the best of our knowledge) syntactic fragment of SO (formulas with first-order dependencies) which we have shown to be contained in  $\Delta_2^1$ . The last two are given in this article (for the first one, cf. [3]); the prenex normal form uses ideas from [1]. The question of which of the containments are strict remains open.

## References

1. Xavier Caicedo and Michal Krynicki. Quantifiers for reasoning with imperfect information and  $\Sigma_1^1$ -logic. *Contemporary Mathematics*, 235:17–31, 1999.
2. Herbert Enderton. Finite partially ordered quantifiers. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 16:393–397, 1970.
3. Santiago Figueira, Daniel Gorín, and Rafael Grimson. On the formal semantics of IF-like logics. *Journal of Computer and System Sciences*, 76(5):333–346, 2009.
4. Jaako Hintikka. *The Principles of Mathematics Revisited*. Cambridge University Press, 1996.
5. Jaako Hintikka and Gabriel Sandu. Game-theoretical semantics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of logic and language*, chapter 6. The MIT press, 1997.
6. Wilfrid Hodges. Compositional semantics for a language of imperfect information. *Logic Journal of the IGPL*, 5(4), 1997.
7. Theo M. V. Janssen. Independent choices and the interpretation of IF logic. *Journal of Logic, Language and Information*, 11(3):367–387, 2002.
8. Theo M. V. Janssen and Francien Dechesne. Signalling in IF games: a tricky business. In *The age of alternative logics*, chapter 15, pages 221–241. Springer, 2006.
9. Marcin Mostowski. Arithmetic with the Henkin quantifier and its generalizations. In F. Gaillard and D. Richard, editors, *Séminaire du Laboratoire Logique, Algorithmique et Informatique Clermontoise*, volume 2, pages 1–25, 1991.
10. Leszek A. Kołodziejczyk. The expressive power of Henkin quantifiers with dualization. Master’s thesis, Institute of Philosophy, Warsaw University, Poland, 2002.
11. Jouko A. Väänänen. On the semantics of informational independence. *Logic Journal of the IGPL*, 10(3):339–352, 2002.