Bisimulations for generalized Veltman semantics

Sebastijan Horvat

Math Department, Faculty of Science of the University of Zagreb

Introduction

R. M. Solovay proved in 1976. that system **Gödel-Löb** (GL) is provability logic of Peano arithmetics. That means that the set of all modal formulas such that their every arithmetical interpretation is theorem of Peano arithmetic is precisely the set of all theorem in system GL. Since provability logic GL cannot distuinguish some properties that different first-order theories have, many extensions of provability logics have been considered. One of them is interpretability logic (IL) defined by **A. Visser** in 1990. System IL is natural from the modal point of view, but arithmetically incomplete - it doesn't prove all of the formulas which are valid in every adequate theory. Various extenstions of IL are obtained by adding some new axioms - for instance, Feferman's principle F which is a modal description of Feferman's theorem - a generalization of Gödel's second incompleteness theorem.

The basic semantics for interpretability logic are Veltman models. **V. Švejdar** in 1991. proved some independence results using Veltman models. They were suitable to distinguish some non-equivalent formulas of interpretability. **D. de Jongh** tried to generalize Švejdar's arguments and came up with the notion of generalized Veltman semantics. **R. Verbrugge** worked this out in an unpublished note.

Although this semantics has so far been called generalized Veltman semantics, after discussion in a wide circle of experts in interpretability logics, it was agreed that the name **Verbrugge's semantics** would be used in her honor.

Fig. 1: Kurt Gödel (1906 - 1978)

(en.wikipedia.org)

Fig. 2: Rineke Verbrugge

(rinekeverbrugge.nl)

Verbrugge frame and Verbrugge model

An ordered triple $(W, R, \{S_w : w \in W\})$ is called a **Verbrugge frame** if it satisfies the following conditions:

a) W is a non-empty set and R is transitive and reverse well-founded relation on W;

b) For every $w \in W$ set S_w is subset of $R[w] \times \mathcal{P}(R[w]) \setminus \{\emptyset\}$;

c) If wRu then $uS_w\{u\}$;

d) If uS_wV and $(\forall v \in V)(vS_wZ_v)$ then $uS_w \bigcup_{v\in V} Z_v$;

e) If $wRuRv$ then $uS_w\{v\}$;

f) If uS_wV and $V \subseteq Z \subseteq R[w]$ then uS_wZ.

An ordered quadruple $(W, R, \{S_w : w \in W\}, \Vdash)$ is called a **Verbrugge model** if it satisfies the following conditions:

1) $(W, R, \{S_w : w \in W\})$ is a Verbrugge frame;

2) $\Vdash$ is a forcing relation. We emphasize only the definition

$$w \Vdash A \rhd B \text{ iff } \forall u((wRu \;\&\; u \Vdash A) \Rightarrow \exists V(uS_wV \& (\forall v \in V)(v \Vdash B))).$$

w-bisimulations and their finite approximations - n-w-bisimulations

A **n-w-bisimulation** between two Verbrugge models $\mathfrak{M} = (W, R, S, \Vdash)$ and $\mathfrak{M}' = (W', R', S', \Vdash')$ is a decreasing sequence of relations $Z_n \subseteq Z_{n-1} \subseteq \cdots \subseteq Z_1 \subseteq Z_0 \subseteq W \times W'$, that possesses the following properties:

(at) If wZ_0w' then $w \Vdash p$ if and only if $w' \Vdash' p$, for all propositional letters p;

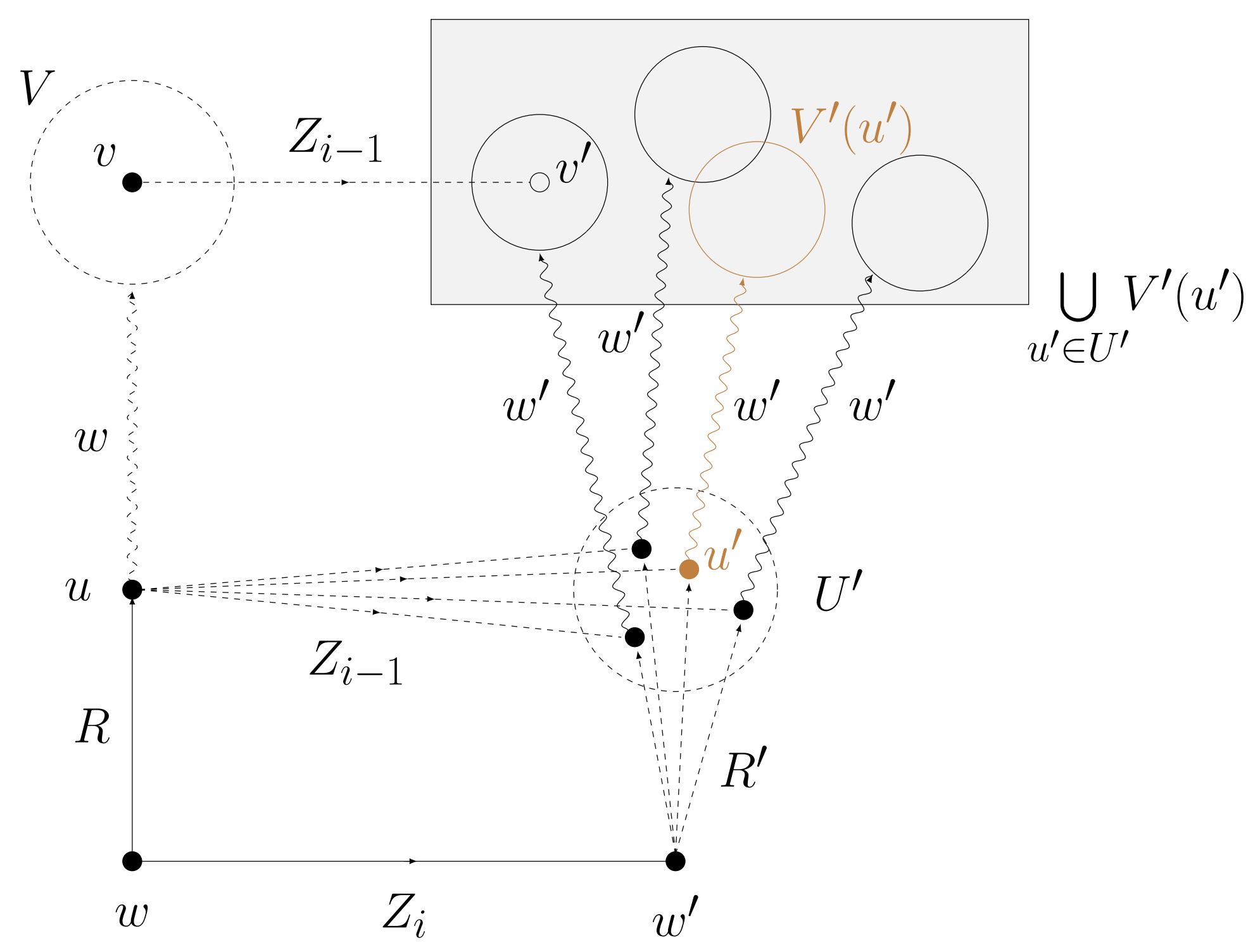
(w-forth) For every i from 1 to n, if wZ_iw' and wRu, then there exists a nonempty set $U' \subseteq W'$ such that for all $u' \in U'$, $uZ_{i-1}u'$ and $w'R'u'$, and for each function $V' : U' \to \mathcal{P}(W')$ such that for all $u' \in U'$, $u'S'_{w'}V'(u')$, there exists set V with uS_wV and for all $v \in V$ there exists $v' \in \bigcup_{u'\in U'} V'(u')$ with $vZ_{i-1}v'$;

(w-back) If wZ_iw' and $w'R'u'$, then there exists a nonempty set $U \subseteq W$ such that for all $u \in U$, $uZ_{i-1}u'$ and wRu, and for each function $V : U \to \mathcal{P}(W)$ such that for all $u \in U$, $uS_wV(u)$, there exists set V' with $u'S'_{w'}V'$ and for all $v' \in V'$ there exists $v \in \bigcup_{u\in U} V(u)$ with $vZ_{i-1}v'$.

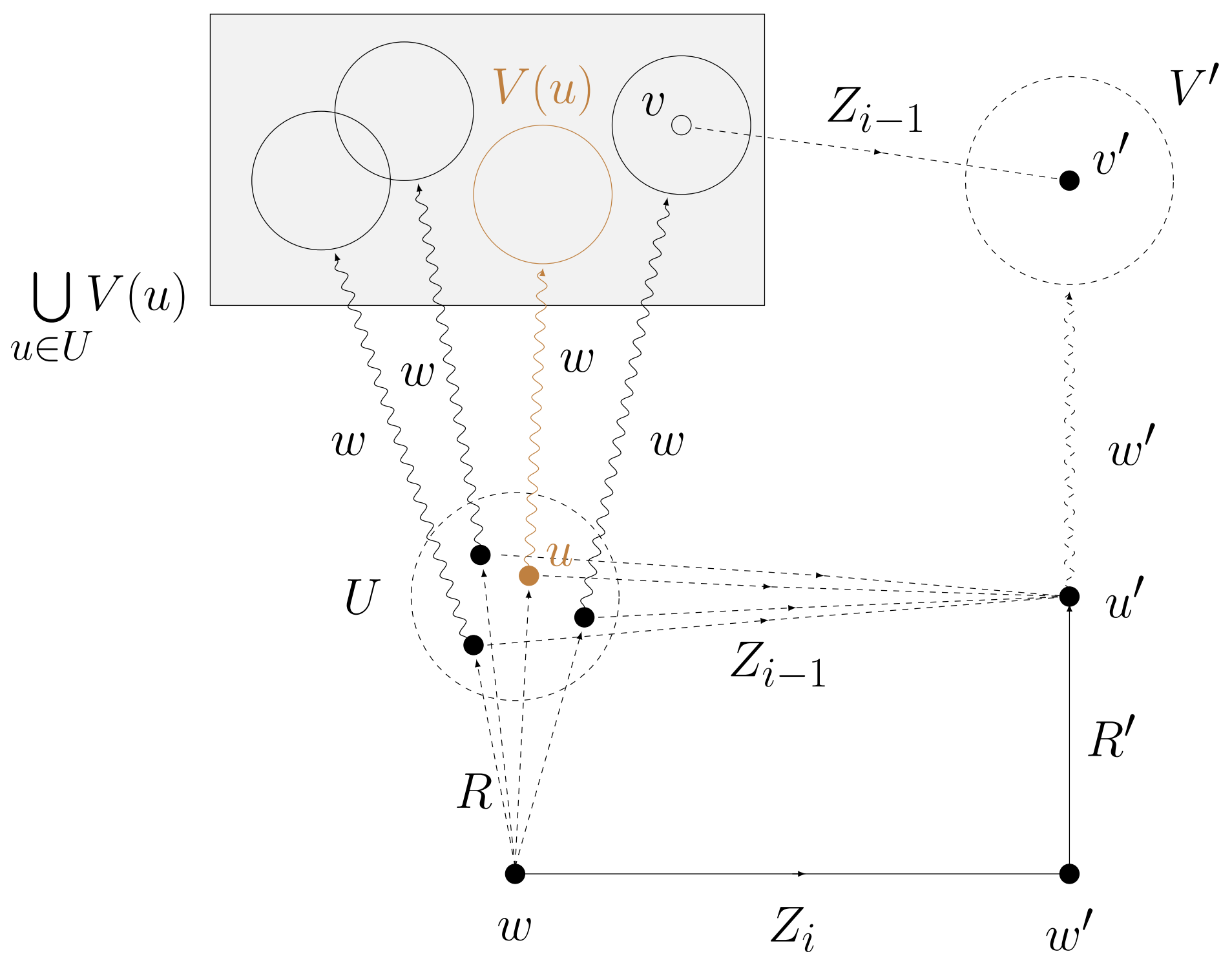
A **w-bisimulation** between $\mathfrak{M}$ and $\mathfrak{M}'$ is a single relation $Z \subseteq W \times W'$, that has the properties (at), (W-forth) and (W-back), with all Z_i being equal to Z. When $Z_n \subseteq Z_{n-1} \subseteq \cdots \subseteq Z_1 \subseteq Z_0$ is a n-W-bisimulation linking the nodes $w \in W$ and $w' \in W'$ we say that w and w' are **n-w-bisimilar**. When Z is a w-bisimulation linking the nodes $w \in W$ and $w' \in W'$ we say that w and w' are **w-bisimilar**.

Note that (w-back) is the same property as (w-forth), just with $\mathfrak{M}$ and $\mathfrak{M}'$ interchanged!

Ilustration of w-forth condition



Ilustration of w-back condition



Basic properties of w-bisimulations

Let $\mathfrak{M}$, $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be Verbrugge models.

a) The relation $Z = \{(w, w) : w \in \mathfrak{M}\}$ is a w-bisimulation. So, w is w-bisimilar to itself.

b) If Z is a w-bisimulation between models $\mathfrak{M}_1$ and $\mathfrak{M}_2$ then $Z^{-1} = \{(w', w) : wZw'\}$ is a w-bisimulation between models $\mathfrak{M}_2$ and $\mathfrak{M}_1$.

c) If Z is a w-bisimulation between models $\mathfrak{M}$ and $\mathfrak{M}_1$, and Z' is a w-bisimulation between models $\mathfrak{M}_1$ and $\mathfrak{M}_2$, then $Z \circ Z'$ is a w-bisimulation between models $\mathfrak{M}$ and $\mathfrak{M}_2$.

d) If $\{Z_i : i \in I\}$ is a set of w-bismulations between models $\mathfrak{M}_1$ and $\mathfrak{M}_2$ then the union $\cup Z_i$ is a w-bisimulation. There is a maximum of all w-bisimulations between models $\mathfrak{M}_1$ and $\mathfrak{M}_2$.

This means that two w-bisimilar worlds are indistinguishable by modal languages!

w-bisimulations preserve modal truth

If $\mathfrak{M}$ and $\mathfrak{M}'$ are two Verbrugge models and $w \in \mathfrak{M}$, $w' \in \mathfrak{M}'$ we say that w and w' are **modally equivalent** if they satisfy the same formula. We have proved the following: if Z is a w-bisimulation between models $\mathfrak{M}$ and $\mathfrak{M}'$ such that $(w, w') \in Z$, then w and w' are modally equivalent.

w-bisimulation games

Let $\mathfrak{M}_i = \left(W_i, R_i, \{S_w^{(i)} : w \in W_i\}, \Vdash\right)$, $i \in \{0, 1\}$, be two Verbrugge models. The **w-bisimulation game** is played by two players, challenger and defender, who move from one configuration to another in a series of consecutive rounds. A configuration is simply a 4-tuple $(\mathfrak{M}_0, w_0, \mathfrak{M}_1, w_1)$, where $w_0 \in W_0$ and $w_1 \in W_1$. A single round, starting with that configuration, is played as follows:

1. Challenger chooses $i \in \{0, 1\}$, index of one Verbrugge model. We denote $j := 1-i$, the index of another model.
2. Challenger picks $u_i \in W_i$ such that $w_iR_iu_i$. If there are no such worlds, the defender wins and game is over.
3. Defender picks a non-empty set of worlds $U_j \subseteq W_j$ such that for all $u_j \in U_j$, $w_jR_ju_j$. If there are no such sets, the challenger wins and game is over.
4. Challenger picks a function $V_j : U_j \to \mathcal{P}(W_j)$ such that $u_jS_{w_j}^{(j)}V_j(u_j)$, for all $u_j \in U_j$.
5. Defender picks a set $V_i \subseteq W_i$ such that $u_iS_{w_i}^{(i)}V_i$.

The configuration from which the next round is played is determined as follows:

(i) Challenger picks $u_j \in U_j$ or $v_i \in V_i$.

(ii) In case the challenger has chosen the world u_j, the configuration with which the next round starts is $(\mathfrak{M}_0, u_0, \mathfrak{M}_1, u_1)$. If the challenger has chosen $v_i \in V_i$, then the defender chooses a world $v_j \in \bigcup_{u_j\in U_j} V_j(u_j)$, and the configuration with which the next round starts is $(\mathfrak{M}_0, v_0, \mathfrak{M}_1, v_1)$.

At the beginning of a game, it is checked that w_0 and w_1 satisfy exactly the same propositional variables. Also, this is checked after each round for worlds u_0 and u_1, and for worlds v_0 and v_1. If any of these checks fail, challenger wins.

It can be proved that every w-game ends (i.e., there are no infinite games)!

w-games and w-bisimulations

An n**-w-game** is a w-game with the following rule added: if n rounds have been played, and challenger hasn't won, then defender wins and the game ends.

A **winning strategy** for a player is a tactic for picking worlds in response to opponent's move, such that the player following it always wins by above rules.

Theorem

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be two Verbrugge models and $w \in \mathfrak{M}$, $w' \in \mathfrak{M}'$ be worlds in them, respectively. Defender has a winning strategy in an n-w-game with a starting configuration $(\mathfrak{M}, w, \mathfrak{M}', w')$ if and only if w and w' are n-w-bisimilar.

Future work

Expressivity issues of modal logics were studied by van Benthem, who developed the subject now known as **correspondence theory**. He proved the fundamental **Characterisation Theorem**: basic modal languages are the bisimulation invariant fragment of the corresponding first-order language. We wish to prove the analogue of this theorem for w-bisimulations and interpretability logics.

Fig. 3: Johan van Benthem

(picture from en.wikipedia.org)

Supported by Croatian Science Foundation (HRZZ), projects UIP-2017-05-9219 and IP-01-2018-7459.