On non-well-founded multisets: Scott Collapse in the Multiworld

Giovanna D'Agostino, Albert Visser

Abstract

We define the class of non well-founded multisets and provide three different descriptions of this class: as collapsed multigraphs, as trees, or as infinitary formulae. Our major tool in this task is *Scott-bisimulation*, which was originally conceived to give an axiomatization of non well-founded multisets. The natural generalization of Scott-bisimulation from graphs to multigraphs allows us to have a theory of multigraph decorations and a notion of collapse in complete analogy with the theory of graph decorations and collapse given by the non well-founded axiomatization of sets $ZFCA$. We also show how our approach to multisets fits in the framework developed by Barwise and Moss.

Contents

1 Introduction

Multisets are very natural objects: they can model a number of different situations in different contexts, like the store of a shop or the bag of a housewife. A multiset (a bag in Computer Science) is like a set, except that an element can have multiple occurrences in it. For example, a grocery shop with 50 apples, 150 pears, 100 banana, and 0 kiwi in store can be modeled by the multiset

$$[\![apple, apple, \cdots, pear, pear, \cdots, banana, banana, \cdots]\!] .$$

In proof theory sequents are often modeled as pairs of multisets (see e.g. [4]).

In general, a multiset of objects from a given domain D can be represented in two ways:

(a) as a partial function from D to $Card^+$, where $Card^+$ is the class of all strictly positive cardinals;

(b) as an equivalence class $[f]$ of functions from a cardinal k to D, where $[f]$ is the set of all function $g \in D^k$ such that there exists a bijection $h : k \to k$ with $f = g \circ h$.

For example, in the grocery shop above the domain D is given by the set $\{a, p, b, k\}$, and the shop is represented by:

(a) a function f with $f : D \to Card^+$, $f(a) = 50$, $f(p) = 150$, $f(b) = 100$, $f(k) =\uparrow$, or

(b) the equivalence class of the function $f : 300 \to D$, where

$$f(x) = \begin{cases} a & \text{if } x < 50 \\ p & \text{if } 50 \leq x < 150 \\ b & \text{if } 150 \leq x < 300. \end{cases}$$

Multiple occurrences of an object $d \in D$ in a multiset can be described by the family of relations $\mathcal{L}_{MS} = \{\in_k : k \in Card^+\}$, interpreted in ($a$) by $d \in_k f \Leftrightarrow f(d) \geq k$ and in (b) by $d \in_k [f] \Leftrightarrow |f^{-1}(d)| \geq k$. When possible, we use a notation like $[\![d, d, c, e, e]\!]$ to represent a multiset.

If the domain D is fixed, there is an obvious equivalence between representations (a) and (b): they both describe the multisets based on D, or D-multisets. In the following, we refer to them as the (a) or (b) representations of D-multisets.

As in the case of ordinary sets we want to represent *iterated* multisets, where multisets contain (various occurrences) of other multisets. In this case the domain D is made of multisets. By epsilon recursion it is easy to define the class of well-founded multisets, but here we are interested in circular situations, like in set theory: we want to have the possibility for a multiset x to be a member of itself, repeated any number of times, i.e. we want to guarantee the existence of multisets satisfying equations like $x = [\![x, x]\!]$.

Circular multisets can be modeled using the theory ZFC^-A (Zermelo-Fraenkel Theory with choice, with foundation replaced by the anti-foundation axiom AFA, see [1]). One possibility is to use the representation (a) above, and define the operator

$$\Delta(X) = \{f \text{ is a function} : dom(f) \subseteq X \wedge range(f) \subseteq Card^+\}.$$

Δ is a monotone operator; its least fixed point gives us the class of *well-founded multisets*. The same class is obtained by considering the operator Γ defined by using representation (b):

$$\Gamma(X) = \{[f] : f \in X^k \text{ for some cardinal } \ k\}.$$

Γ is a monotone operator; its least fixed point is isomorphic, as an $\mathcal{L}_{MS}$-structure, to the least fixed point of Δ. We denote it by $wfMS$. The operators Δ and Γ also have greatest fixed points, which in ZFC^-A contain strictly the well-founded multisets. However, we prove that the greatest fixed point Γ^* of Γ *is not* isomorphic to the greatest fixed point Δ^* of Δ as an $\mathcal{L}_{MS}$-structure: Γ^* has nice AFA-like properties connected with decorations of *multigraphs* which are not shared by Δ^*. Using these properties we find a correspondence between Γ^* and a certain class of *collapsed multigraphs*, similar to the correspondence between sets and collapsed graphs given by the anti-foundation axiom AFA ([1]). Defining the multiset-class via Γ^* allows us to propose an alternative (and easier to visualize) definition of multisets. This alternative is lost when we consider multisets defined via Δ.

Our major tool in this task is *Scott-bisimulation*. In his work [3], Scott proposed a theory of non-well-founded sets, which was later reconsidered by Aczel in [1]. Using the notion of Scott-bisimulation, Aczel compares Scott-theory ZFC^-S with ZFC^-A. Both can be obtained from ZFC^- (Zermelo-Fraenkel with choice, without foundation) by using a strengthening of the extensionality axiom, defined in terms of bisimulation: the maximal bisimulation for ZFC^-A, Scott-bisimulation for ZFC^-S. In this confrontation, the Scott axiomatization ZFC^-S seems to be less natural and manageable than ZFC^-A: in ZFC^-A any graph has a *unique* decoration, while in ZFC^-S graphs can have more than one decoration; in ZFC^-A any set can be represented by a collapsed graph, while a similar notion of collapse is not available in ZFC^-S.

In this work we claim that the Scott axiomatization looks like ZFC^-A more than it seems, but that this resemblance can be appreciated only by changing from the set to the multiset-context (defined via the Γ operator above). We show that Scott-bisimulation (in its generalization to multigraphs) gives us a way to define multiset-decorations of multigraphs, a criterion for deciding multisets equality, and a notion of collapsed multigraphs corresponding to multisets. In this way we show that the multiset-context is more natural than the set-context for Scott-bisimulation, because in the latter there is no natural notion of collapse or decoration one can work with.

Using Scott bisimulation we prove that the multiset-class admits three different descriptions, beside the greatest fixed point one. These representations are obtained by identifying multisets with rooted collapsed multigraphs, with rooted trees, or with a fragment of the class of infinitary graded modal formulae.

Summarizing, the interest of the work goes two ways. Starting from multisets, rooted collapsed multigraphs, rooted trees, and graded formulas are useful as representations. Conversely, multisets correspond directly to rooted collapsed multigraphs and rooted trees, providing standard representatives of multigraphs modulo Scott-bisimulation and of rooted trees modulo isomorphism, respectively.

This paper is organized as follows. In Section 2 we compare two possible definitions of the multiset-class, working out a theory of multigraph decoration. In Section 3 we give a notion of multi-bisimulation and the corresponding notion of collapse, and we identify multisets and rooted collapsed multigraphs. In Section 4 we compare our notion of multi-bisimulation with Scott-bisimulation, and this gives the second representation of multisets, via rooted trees. Finally, in Section 5 we deal with the identification of multisets with infinitary graded modal formulae.

Due to space limits, we only give examples and no proofs; to understand the paper the reader is supposed to be familiar with the results in [1] and [2].

We would like to thank Marc Bezem, Jelle Gerbrandy, Vincent van Oostrom, and Alberto Policriti for discussions on multisets.

2 Multisets and multigraph decorations

In this section we compare the greatest fixed points of the operators Δ and Γ defined in the introduction. In particular, we study their behavior w.r.t. decorations. We want to develop a theory of decorations which is the multi-analogue of the theory of graph decorations. In non-well-founded set theory, graphs are used as pictures of sets. To model multiplicity, we shall use *multigraphs*.

2.1 Multigraphs

A multigraph is nothing but a graph in which two nodes can be connected by more than one arrow. A graph can be determined by the set of nodes and a subset of successors $\mathsf{Succ}(w)$ for each node w, giving the nodes that are directly accessible from w. Analogously, we specify a multigraph giving a set of nodes $O_{\mathcal{G}}$ and an $O_{\mathcal{G}}$-multiset $\mathsf{MSucc}(w)$, for each node w.

Definition 2.1 *A* multigraph $\mathcal{G}$ *is a pair* $(O_{\mathcal{G}}, \mathsf{MSucc})$*, where* $O_{\mathcal{G}}$ *is a set of* objects *and* MSucc *is a function from* $O_{\mathcal{G}}$ *to the* $O_{\mathcal{G}}$*-multisets.*

If we use representation (a) of an $O_{\mathcal{G}}$-multiset, $\mathsf{MSucc}(w)$ is a partial function from $O_{\mathcal{G}}$ to the strictly positive cardinals, for each node w; if we use representation (b), $\mathsf{MSucc}(w)$ is an equivalence class of functions from a cardinal k to $O_{\mathcal{G}}$. By abuse of notation we denote the set of objects $O_{\mathcal{G}}$ of a multigraph $\mathcal{G}$ by the same symbol $\mathcal{G}$. For each node w, we denote by $\mathsf{Succ}(w)$ the set on which $\mathsf{MSucc}(w)$ is based, e.g. by using (a)-representation the set $\mathsf{Succ}(w)$ is the domain of the partial function $\mathsf{MSucc}(w)$. If $w, v \in \mathcal{G}$, we denote the multiplicity of v as an element in $\mathsf{MSucc}(w)$ by $m_{\mathcal{G}}(w, v)$: e.g. under the (a)-representation $m_{\mathcal{G}}(w, v)$ is equal to $\mathsf{MSucc}(w)(v)$, if v is in the domain of the partial function

$\mathsf{MSucc}(w)$, 0 otherwise. Notice that the function $m_{\mathcal{G}}$ determines the function MSucc, and we could as well define a multigraph as a pair $(O_{\mathcal{G}}, m_{\mathcal{G}})$, where $m_{\mathcal{G}}$ is a function from $\mathcal{G} \times \mathcal{G}$ to cardinals. In the following, we represent a multigraph by a picture in which two nodes w, v of $\mathcal{G}$ are linked by $m_{\mathcal{G}}(w, v)$ arrows. A *rooted multigraph* is a multigraph $\mathcal{G}$ with a distinguished object (the *root*) $w \in O_{\mathcal{G}}$, in which every object is accessible from the root: if $v \in O_{\mathcal{G}}$ then there is a sequence $v_0, \ldots, v_n$ of elements in $O_{\mathcal{G}}$ such that $v_0 = w$ $v_n = v$ and $m_{\mathcal{G}}(v_i, v_{i+1}) \geq 1$, for all $i < n$.

An *isomorphism* between two multigraphs $\mathcal{G}$, $\mathcal{H}$ is a bijective function ϕ between $\mathcal{G}$ and $\mathcal{H}$ such that for all $w, v \in \mathcal{G}$ it holds: $m_{\mathcal{G}}(w, v) = m_{\mathcal{G}'}(\phi(w), \phi(v))$. An isomorphism of rooted multigraphs $(\mathcal{G}, w)$, $(\mathcal{H}, v)$ is a isomorphism ϕ between $\mathcal{G}$ and $\mathcal{H}$ such that $\phi(w) = v$.

2.2 Multi-decorations

The following multigraph analogue of a decoration should not come as a surprise, for readers familiar with graph decorations: a multigraph decoration consists of a function from the set of objects of the multigraph to multisets, in such a way that the decoration of a node w *is* the multiset consisting of all decorations (counting multiplicity) of $Succ(w)$. The definition of multigraph decorations obviously depends on the choice of the class C we use to represent multisets: we speak about C-multi-decorations of multigraphs. For example, in fig.1 the multigraph $\mathcal{G}$ has a $wfMS$-multi-decoration defined by decorating v with $\emptyset$ and the root w by $[\![\emptyset, \emptyset]\!]$. As another example consider the multigraph $\mathcal{H}$ in fig.1. In this case, the leaves v_1, v_2 are both $wfMS$-multi-decorated by $\emptyset$ and the root w' by $[\![\emptyset, \emptyset]\!]$: hence multiple occurrences are not obtained exclusively because of multiple edges.

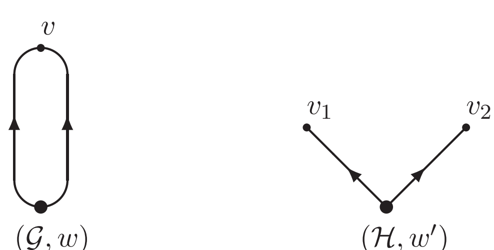

Fig.1

More formally, consider a class-structure C for the language $\mathcal{L}_{MS} = \{\in_k : k \in Card^+\}$, and denote by $\in_k^C$ the interpretation of $\in_k$ in the structure C.

Definition 2.2 *A C-*multi-decoration *of a multigraph $\mathcal{G}$ is a function D from $\mathcal{G}$ to multisets, in such a way that for any object $w \in \mathcal{G}$ it holds:*

$$z \in_k^C D(w) \Leftrightarrow \sum_{D(v)=z} m_{\mathcal{G}}(w, v) \geq k.$$

A multi-decoration of a rooted multigraph $(\mathcal{G}, w)$ is a multiset $D(w)$, where D is a multi-decoration of the multigraph $\mathcal{G}$.

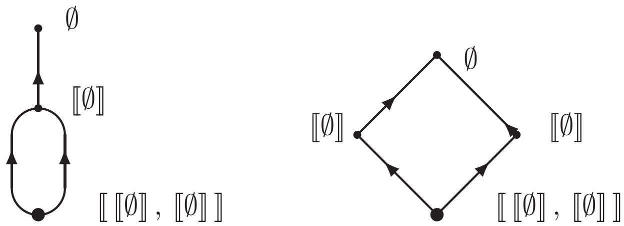

Fig.2

In the previous pictures, each object has been labeled by its $wfMS$-multi-decoration. Notice that the two rooted multigraphs have the same multi-decoration.

Can we find a class-structure C for $\mathcal{L}_{MS}$ having the property that any multigraph has a unique C-multi-decoration? Let us consider the two candidates Δ^* and Γ^* defined in the introduction. First we provide a multigraph which does not have any Δ^*-multi-decoration.

Counterexample.
Consider the following (multi)graph $\mathcal{G} = (\{w, v\}, m_{\mathcal{G}})$, where the function $m_{\mathcal{G}}$ is defined as follows: $m_{\mathcal{G}}(w, w) = m_{\mathcal{G}}(w, v) = m_{\mathcal{G}}(v, v) = 1$, and 0 in all other cases. We first prove that if D is a Δ^*-multi-decoration of $\mathcal{G}$, then $D(w)$ must be different from $D(v)$. Indeed, if $D(w) = D(v)$ then $\sum_{D(u)=D(v)} m_{\mathcal{G}}(w, u) = 2$ and $D(v) \in_2 D(w)$; but $\sum_{D(u)=D(v)} m_{\mathcal{G}}(v, u) = 1$ and $D(v) \notin_2 D(v)$, contradicting $D(w) = D(v)$. On the other hand, a Δ^*-multi-decoration of $\mathcal{G}$ with $D(w) \neq D(v)$ is a solution for the system of equations $x_w = \{\langle x_w, 1\rangle, \langle x_v, 1\rangle\}$; $\quad x_v = \{\langle x_v, 1\rangle\}$, where we denoted by $\langle x, y\rangle$ the ordered pair $\{\{x\}, \{x, y\}\}$. But then $D(v) = D(w) = \{\langle D(v), 1\rangle\}$ after all, by uniqueness of solutions.

Hence, the class Δ^* is not well-behaved w.r.t. multigraph decorations. Fortunately, the class Γ^* behaves much better. There is a general reason for this: Γ is monotone and proper (this and the following terminology is from [2]), a multigraph $\mathcal{G}$ can be seen as a flat Γ-coalgebra, and a Γ^*-multi-decoration of $\mathcal{G}$ is nothing but a solution for this Γ-coalgebra, with value in Γ^*. Then Proposition 16.2 in [2] tells us that any flat Γ-coalgebra has a unique solution in Γ^*, if Γ is monotone and proper. This gives:

Theorem 2.3 *Any multigraph has a unique Γ^*-multi-decoration.*

Things go wrong with the operator Δ because we cannot see a multigraph as a flat Γ-coalgebra any more, because the operator Δ does not commute with the action of substitutions.

A direct proof of Theorem 2.3 can be given by associating to each multigraph $\mathcal{G}$ a system of equations $Eq(\mathcal{G})$ in such a way that a function D with domain $\mathcal{G}$ is a Γ^*-multi-decoration of $\mathcal{G}$ iff it is a solution for $Eq(\mathcal{G})$. Then existence and uniqueness of Γ^*-multi-decorations follows directly from ZFC^-A.

Theorem 2.3 gives us a good reason for choosing Γ^* to represent the class of non well-founded multisets. We acknowledge this by a formal definition:

Definition 2.4 *The multiset-class MS is the class Γ^* endowed with the family of binary relations $\{\in_k : k \in Card^+\}$, defined by $[f] \in_k [g] \Leftrightarrow |g^{-1}([f])| \geq k$.*

¿From now on, a multiset is an element in Γ^*, and a multi-decoration is a Γ^*-multi-decoration.

3 Multi-bisimulation

We now give a multiset-free criterion for establishing whether two multigraphs $\mathcal{G}, \mathcal{H}$ have the same multi-decoration. In the set-context this was done via the notion of bisimulation. We generalize this notion in two ways. The first comes from the general theory of monotone operators of [2], while the second is more suited for working with graded modal formulae (see Section 5). Although the first notion is strictly stronger than the second, in the sense that there are relations satisfying the second but not the first definition, we prove that the maximal elements in both classes coincide, being both equal to the class $\{(w, w') \in \mathcal{G} \times \mathcal{H} : D_{\mathcal{G}}(w) = D_{\mathcal{H}}(w')\}$.

Let us start with the notion of Γ-bisimulation. By the general theory in [2] we know that any smooth operator Σ that preserves covers admits a notion of Σ-bisimulation for which it holds: two points x, y of a Σ-coalgebra are Σ-bisimilar iff they have the same value under the solution for the coalgebra. It is possible to prove that the operator Γ defining multisets is smooth and cover preserving, and, as we already pointed out in the previous section, a multigraph is nothing but a Γ-coalgebra, with multi-decorations corresponding to solutions. In our case, the notion of Γ-bisimulation amounts to the following:

Definition 3.1 *A Γ-bisimulation between two multigraphs $\mathcal{G}, \mathcal{H}$ is a relation $Z \subseteq \mathcal{G} \times \mathcal{H}$ such that if wZw' then $\mathsf{MSucc}(w)$ and $\mathsf{MSucc}(w')$ are projections of a Z-multiset, i.e. (using representation (b)): there are functions f, g such that $\mathsf{MSucc}(w) = [f], \mathsf{MSucc}(w') = [g], dom(f) = dom(g)$, and $f(\nu)Zg(\nu)$ for all $\nu \in dom(f)$.*

¿From [2] we have:

Proposition 3.2 *Two rooted multigraphs have the same decoration iff they are Γ-bisimilar.*

We shall now define a weaker notion of bisimulation between multigraphs, the *multi-bisimulation*.

Definition 3.3 *A* multi-bisimulation *between two multigraphs $\mathcal{G}$, $\mathcal{H}$ is a relation $Z \subseteq \mathcal{G} \times \mathcal{H}$ such that if wZw' then*

(a) if $X \subseteq Succ(w)$ and $Y = \{y \in Succ(w') : \exists x \in X \quad xZy\}$, then

$$\sum_{x \in X} m_{\mathcal{G}}(w, x) \leq \sum_{y \in Y} m_{\mathcal{H}}(w', y);$$

(b) if $Y \subseteq Succ(w')$ and $X = \{x \in \mathcal{G} : \exists y \in Y \quad xZy\}$, then

$$\sum_{y \in Y} m_{\mathcal{H}}(w', y) \leq \sum_{x \in X} m_{\mathcal{G}}(w, x).$$

A multi-bisimulation Z between two rooted multigraphs $(\mathcal{G}, w)$, $(\mathcal{H}, w')$ is a multi-bisimulation between $\mathcal{G}$ and $\mathcal{H}$ with wZw'. We write $(\mathcal{G}, w) \approx_m (\mathcal{H}, w')$ (or simply $w \approx_m w'$, if the multigraphs $\mathcal{G}$, $\mathcal{H}$ are uniquely determined from the context) if there is a multi-bisimulation between $(\mathcal{G}, w)$ and $(\mathcal{H}, w')$.

In the following, we often consider a multi-bisimulation Z between rooted multigraphs $(\mathcal{G}, w)$, $(\mathcal{H}, w')$ which is the restriction to $\mathcal{G} \times \mathcal{H}$ of an equivalence relation $\sim$ on the class of rooted multigraphs. This is the case of $\approx_m$, or, as we shall see later, of the relation *to satisfy the same graded formulae*. In this case it make sense to consider, for any node w in a multigraph, the classes of $\sim$ over w, that is, the sets $\{z \in Succ(w) : z \sim v\}$, for any $v \in Succ(w)$. The following remark proves that we can consider only this kind of *elementary* subsets to give an equivalent (and easier to visualize) definition of a multi-bisimulation.

Remark 3.4 If $Z \subseteq \mathcal{G} \times \mathcal{H}$ is the restriction to $\mathcal{G} \times \mathcal{H}$ of an equivalence relation $\sim$ on the class of rooted multigraphs, Definition 3.3 is equivalent to: if wZw' then

- (c) for all $v \in Succ(w)$ there exists a $v' \in Succ(w')$ such that vZv' and $\sum_{z \sim v} m_{\mathcal{G}}(w, z) = \sum_{z' \sim v'} m_{\mathcal{H}}(w', z')$;
- (d) for all $v' \in Succ(w')$ there exists a $v \in Succ(w)$ such that vZv'.

It is easy to check that a Γ-bisimulation (as in Definition 3.1) is a multi-bisimulation, but the converse is not true. However, the following lemma says that $\approx_m$ is the maximal multi-bisimulation between any pair of multigraphs, and Lemma 3.6 says that $\approx_m$ is a Γ-bisimulation as well.

Lemma 3.5 *The equivalence relation $\approx_m$, when restricted to multigraphs $\mathcal{G}$, $\mathcal{H}$, is a multi-bisimulation between $\mathcal{G}$ and $\mathcal{H}$. In particular, if $(\mathcal{G}, w) \approx_m (\mathcal{H}, w')$ and $u \in \mathsf{Succ}(w) \cup \mathsf{Succ}(w')$ then $\sum_{v \approx_m u} m_{\mathcal{G}}(w, v) = \sum_{v' \approx_m u} m_{\mathcal{H}}(w', v')$.*

Lemma 3.6 *The relation $\{(w, w') : (\mathcal{G}, w) \approx_m (\mathcal{H}, w')\}$ is a Γ-bisimulation between any pair of multigraphs $\mathcal{G}, \mathcal{H}$.*

From the above lemma and Proposition 3.2 it follows:

Lemma 3.7 $(\mathcal{G}, w) \approx_m (\mathcal{H}, w') \quad \Leftrightarrow \quad D_{\mathcal{G}}(w) = D_{\mathcal{H}}(w')$.

3.1 The Collapse

As in the set-context we use the notion of collapse to give canonical representatives for the classes of the equivalence relation *to have the same multi-decoration* between multigraphs. In other words, given a class of multigraphs with the same multi-decoration (that is, a class of multigraphs that are *pictures* of the same multiset), we give a way to select a unique representative in the class (modulo isomorphism). This allow us to identify collapsed multigraphs and multisets.

Definition 3.8 *The* Collapse *of a multigraph $\mathcal{G}$ is defined as the multigraph $\mathcal{G}^\star$ with set of objects equal to $\{[w] : w \in \mathcal{G}\}$ where $[w] = \{v \in \mathcal{G} : w \approx_m v\}$, and $m_{\mathcal{G}^\star}([w], [v]) = \sum_{z \approx_m v} m_{\mathcal{G}}(w, z)$.*

The collapse of a multigraph is well-defined by Lemma 3.5. We define the collapse of a rooted multigraph $(\mathcal{G}, w)$ as the rooted multigraph $(\mathcal{G}^\star, [w])$.

Notice that it is not possible to define $\mathcal{G}^\star$ by simply considering as edges between two classes σ, τ all $\mathcal{G}$-edges between objects in σ and objects in τ. If we were doing so, the collapse of the rooted multigraph $(\mathcal{G}, w)$ on the right of picture 2 would be the multigraph $\mathcal{K} = (O_\mathcal{K}, m_\mathcal{K})$, with $O_\mathcal{K} = \{w_1, w_2, w_3\}$, root w_1, and $m_\mathcal{K}(w_i, w_j) = 2$, if $j = i + 1$, and 0 otherwise. But the two rooted multigraphs do not have the same decoration (the root of the multigraph on picture 2 is decorated by $\llbracket \llbracket \emptyset \rrbracket , \llbracket \emptyset \rrbracket \rrbracket$, while the root w_1 of the one described here is decorated by $\llbracket \llbracket \emptyset, \emptyset \rrbracket , \llbracket \emptyset, \emptyset \rrbracket \rrbracket$).

A multigraph and its collapse belong to the same $\approx_m$-class:

Lemma 3.9 *The relation $Z = \{(w, [w]) : w \in \mathcal{G}\}$ is a multi-bisimulation between $\mathcal{G}$ and $\mathcal{G}^\star$.*

The following lemma says that a collapsed multigraph is an exact picture of its decoration, in the sense that different multisets correspond to different nodes.

Lemma 3.10 *The decoration $D_{\mathcal{G}^\star}$ of a collapsed multigraph $\mathcal{G}^\star$ is an injective function.*

Using the previous lemma we see that for a collapsed multigraph $\mathcal{G}^\star$ the multi-decoration $D_{\mathcal{G}^\star}$ satisfies

$$z \in_k D_{\mathcal{G}^\star}(w) \Leftrightarrow \exists v \in \mathsf{Succ}(w) \text{ with } \ D(v) = z \text{ and } \ m_{\mathcal{G}^\star}(w, v) \geq k,$$

and vice versa, any *injective* function satisfying this equivalence is a multi-decoration of $\mathcal{G}^\star$.

A rooted multigraph has the same decoration as its collapse:

Lemma 3.11 *$D_{\mathcal{G}^\star}([w]) = D_\mathcal{G}(w)$, for any multigraph $\mathcal{G}$.*

and multi-bisimilar rooted multigraphs have the same collapse, modulo isomorphism:

Theorem 3.12 *$(\mathcal{G}, w) \approx_m (\mathcal{H}, w') \Leftrightarrow (\mathcal{G}^\star, [w])$ is isomorphic to $(\mathcal{H}^\star, [w'])$.*

Theorem 3.12 and Lemma 3.7 allow us to identify multisets and collapsed multigraphs. Given a multiset x, define its multi-transitive closure mTC as the *set* containing all multisets that are *hereditarily* members of x (i.e. if $x = [f]$, where f is a function from a cardinal to multisets, then $\mathsf{mTC}(x)$ contains the multisets that are in the range of f, as well as the multisets that are in the range of a function representing an element in this range, and so on; $\mathsf{mTC}(x)$ can be defined by a simple recursion. Define the *canonical rooted multigraph* $\mathcal{G}(x)$ of a multiset x as the rooted multigraph having set of objects equal to $\mathsf{mTC}(x) \cup \{x\}$, as root the multiset x, and as MSucc the identity function. Starting from a multigraph $(\mathcal{G}, w)$, we can either perform its collapse $(\mathcal{G}^*, [w])$, or first consider its multi-decoration $D_\mathcal{G}(w)$ and then its canonical rooted multigraph

$\mathcal{G}(D_{\mathcal{G}}(w))$, ending with isomorphic rooted multigraphs. The isomorphism is given by sending $[v] \in \mathcal{G}^*$ to $D_{\mathcal{G}^*}([v]) = D_{\mathcal{G}}(v)$. Define a multigraph to be *extensional* if two different objects are never multi-bisimilar. Every collapsed multigraph is extensional and every extensional rooted multigraph $(\mathcal{G}, w)$ is isomorphic to $\mathcal{G}(D_{\mathcal{G}}(w))$. So the $\mathcal{G}(D_{\mathcal{G}}(w))$ provide standard representatives for the extensional-rooted-multigraph equivalence classes modulo isomorphism.

Finally, notice that the multi-membership relations are represented in the class of collapsed multigraphs by:

$$(\mathcal{G}, w) \in_k (\mathcal{H}, v) \Leftrightarrow \exists z \in \mathcal{H} \text{ with } (\mathcal{H}_{|_z}, z) \cong (\mathcal{G}, w) \text{ and } m_{\mathcal{H}}(v, z) \geq k,$$

where $\mathcal{H}_{|_z}$ is the restriction of $\mathcal{H}$ to the set of nodes that are reachable in $\mathcal{H}$ from z.

4 Scott-bisimulation and unravelings

In this section we show that the notion of multi-bisimulation between multigraphs is the natural generalization of the notion of Scott-bisimulation between graphs (see [1]). This allow us to identify multisets and rooted trees. With this aim, we generalize the notion of unraveling, from graphs to multigraphs. First, we summarize the relationship between bisimulation and unraveling in the context of graphs. Here, unravelings can be used to characterize both the maximal bisimulation and the Scott one. The unraveling of a graph produces a rooted tree, in which every node is copied once (in the simple unraveling) or k-times for a cardinal k (in the k-unraveling). It is possible to prove that two graphs are bisimilar if there exists a cardinal k such that the k-unravelings of the graphs are isomorphic, while two graphs are Scott-bisimilar if the simple unravelings are isomorphic.

In the following definition we generalize the notion of simple unraveling to multigraphs.

Definition 4.1 *The* Unraveling *of a rooted multigraph* $(\mathcal{G}, w)$ *is the rooted multigraph* $(\mathcal{G}, w)^U$ *defined as follows:*

(a) The set of objects is the set of finite sequences of type $v_0 k_1 v_1 \ldots k_n v_n$*, where* $v_0 = w$ *and the* k_i*'s are cardinal numbers satisfying:* $k_{i+1} < m_{\mathcal{G}}(v_i, v_{i+1})$*, for all* $i \in \{0, \ldots, n-1\}$ *(in particular, if* $v_0 k_1 v_1 \ldots k_n v_n \in (\mathcal{G}, w)^U$ *then* $m_{\mathcal{G}}(v_i, v_{i+1}) \geq 1$ *and* $v_{i+1} \in \mathsf{Succ}(v_i)$*, for all* $i < n$*).*

b) The root of $(\mathcal{G}, w)^U$ *is the sequence* w*.*

c) If $\sigma = v_0 k_1 v_1 \ldots k_n v_n$ *and* $\tau = \sigma k_{n+1} v_{n+1} \in (\mathcal{G}, w)^U$ *then* $m_{(\mathcal{G},w)^U}(\sigma, \tau) = 1$*, while in all other cases we have* $m_{(\mathcal{G},w)^U}(\sigma, \tau) = 0$*.*

For example, in fig.1 the unraveling of the multigraph $(\mathcal{G}, w)$ is the multigraph $(\mathcal{H}, w')$. Notice that the unraveling of a multigraph $\mathcal{G}$ is always a graph (i.e. $m_{\mathcal{G}}(w, v) \leq 1$, for $w, v \in \mathcal{G}$), and if $\mathcal{G}$ is a graph then the notion of multigraph unraveling coincides with the notion of graph unraveling. Scott-bisimulation (in the equivalent definition given in [1]) relates two rooted graphs

$(\mathcal{G}, w)$, $(\mathcal{H}, w')$ exactly when their unravelings are isomorphic. Theorem 4.2 below says that two multigraphs are multi-bisimilar if and only if their unravelings are isomorphic, thereby showing that we are working on a generalization of Scott-bisimulation.

Theorem 4.2 *If $\mathcal{G}, \mathcal{H}$ are multigraphs then $(\mathcal{G}, w) \approx_m (\mathcal{H}, v) \Leftrightarrow (\mathcal{G}, w)^U$ is isomorphic to $(\mathcal{H}, v)^U$.*

Theorem 4.2 and Lemma 3.7 allow us to identify multisets and rooted trees. Define the *canonical rooted tree* $\mathcal{T}(x)$ of a multiset x as $(\mathcal{G}(x))^U$, where the rooted multigraph $\mathcal{G}(x)$ has been defined at the end of Section 3.1. Staring from a rooted multigraph $(\mathcal{G}, w)$, we can either perform its unraveling $(\mathcal{G}, w)^U$, or first consider its multi-decoration $D_{\mathcal{G}}(w)$ and then its canonical rooted tree $\mathcal{T}(D_{\mathcal{G}}(w))$ ending with isomorphic rooted trees. The isomorphism is given by sending a sequence $\sigma = v_0 k_1 \dots k_n v_n$ to $D_{\mathcal{G}}(v_1) k_1 \dots k_n D_{\mathcal{G}}(v_n)$. Then as in the case of collpased multigraphs we have a way to select standard representatives for the rooted-tree equivalence classes modulo isomorphism.

The multi-membership relations are represented in the rooted-tree class by

$$(T, t) \in_k (S, s) \Leftrightarrow |\{u \in S : (S_{|_u}, u) \text{ is isomorphic to } (T, t)\}| \geq k,$$

where (T, t) and (S, s) are rooted trees, and $(S_{|_u}, u)$ denotes the subtree of S with root u.

5 Multisets and the logic of graded modalities

In this section we give a characterization of multi-bisimulation via logic. In the set-context, the appropriate logic for describing bisimulation between graphs was proved to be infinitary modal logic ([2]). In the multiset-context we shift to the *graded extension* of this logic by proving that two rooted multigraphs are multi-bisimilar iff they satisfy the same formulae of infinitary graded modal logic. More than this, we prove that any multigraph can be characterized by a single infinitary graded modal formula, and we isolate a class of formulae that correspond to multisets. In this way we have three alternative ways for modeling multisets: as collapsed multigraphs, as trees, or as infinitary formulae.

Let $\mathcal{L}_\infty^{grad}$ be the *infinitary logic of graded modalities*, that is, the logic obtained from infinitary modal logic by adding the unary operators $\diamond_h$, for all $h \in Card^+$. More formally, we define $\mathcal{L}_\infty^{grad}$ as the smallest class closed under infinitary conjunction (if $\Phi \subseteq \mathcal{L}_\infty^{grad}$ is a *set* then $\bigwedge \Phi \in \mathcal{L}_\infty^{grad}$), negation (if $\phi \in \mathcal{L}_\infty^{grad}$ then $\neg\phi \in \mathcal{L}_\infty^{grad}$), and graded diamonds (if h is a strictly positive cardinal and $\phi \in \mathcal{L}_\infty^{grad}$ then $\diamond_h \phi \in \mathcal{L}_\infty^{grad}$).

The *truth* of a formula ϕ of $\mathcal{L}_\infty^{grad}$ in a rooted multigraph $(\mathcal{G}, w)$ is obtained by adding the clause below to the inductive definition of truth in $\mathcal{L}_\infty$:

$$(\mathcal{G}, w) \models \diamond_h \phi \quad \Leftrightarrow \quad \sum_{(\mathcal{G}_{|_v}, v) \models \phi} m_{\mathcal{G}}(w, v) \geq h.$$

If we denote $\diamond_1$ by $\diamond$, we see that the logic $\mathcal{L}_\infty^{grad}$ is an extension of infinitary modal logic. We write $(\mathcal{G}, w) \equiv_g (\mathcal{H}, w')$ (or simply $w \equiv_g w'$) if $(\mathcal{G}, w), (\mathcal{H}, w')$ are two rooted multigraphs that satisfy the same $\mathcal{L}_\infty^{grad}$-formulae.

The following theorem show that $\mathcal{L}_\infty^{grad}$ is the appropriate language for characterizing multi-bisimulation.

Theorem 5.1 *For any rooted multigraph $(\mathcal{G}, w)$ there exists a formula $\phi_{(\mathcal{G},w)} \in \mathcal{L}_\infty^{grad}$ which characterizes $(\mathcal{G}, w)$ modulo multi-bisimulation, i.e., for any multigraph $\mathcal{H}$ it holds: $(\mathcal{H}, w') \models \phi_{(\mathcal{G},w)} \Leftrightarrow (\mathcal{G}, w) \approx_m (\mathcal{H}, w')$.*

This result suggests another representation of the class of multisets, in which the domain of the universe is a fragment of the class of infinitary graded modal formulae. First, we characterize the graded formulae of type $\phi_{(\mathcal{G},w)}$ for a rooted multigraph $(\mathcal{G}, w)$ (for a set-analogue, see [2]).

Definition 5.2 *Consider the preorder $\leq$ defined in $\mathcal{L}_\infty^{grad}$ by $\psi \leq \phi \Leftrightarrow \models \psi \to \phi$, where $\models \psi \to \phi$ stands for: any rooted multigraph that satisfies ψ, satisfies ϕ as well. Define the class $MS(\mathcal{L}_\infty^{grad})$ as the one containing, modulo equivalence, all satisfiable $\mathcal{L}_\infty^{grad}$-formulae which are minimal with respect to $\leq$ on satisfiable formulae, that is:*

$$\phi \in MS(\mathcal{L}_\infty^{grad})$$
$$\Updownarrow$$
$$\phi \text{ is satisfiable and for all satisfiable } \psi \text{ if } \psi \leq \phi \text{ then } \psi \text{ is equivalent to } \phi.$$

Lemma 5.3 *$\phi \in MS(\mathcal{L}_\infty^{grad}) \Leftrightarrow \exists\ (\mathcal{G}, w)$ with $\models (\phi_{(\mathcal{G},w)} \leftrightarrow \phi)$, for all $\phi \in \mathcal{L}_\infty^{grad}$.*

Then, we identify the class MS with the class $MS(\mathcal{L}_\infty^{grad})$, with $\in_k$ given by the relation $\{(\phi, \psi) \in MS(\mathcal{L}_\infty^{grad}) \times MS(\mathcal{L}_\infty^{grad}) : \models \phi \to \diamond_k \psi\}$.

References

[1] P. Aczel. *Non-well-founded sets.* CSLI, Lecture Notes 14, 1988.

[2] J. Barwise and L. Moss. *Vicious Circles.*, volume 60 of *CSLI Lecture Notes*, Stanford California, 1996.

[3] D. Scott. A different kind of model for set theory. Unpublished paper, given at the *1960 Stanford Congress of Logic, Methodology and Philosophy of Science.*

[4] A.S. Troelstra and H. Schwichtenberg. *Basic Proof Theory* Cambridge Tracts in Theoretical Computer Science 43, CUP, 1996.