



HAL
open science

Axiomatizations of betweenness in order-theoretic trees

Bruno Courcelle

► **To cite this version:**

| Bruno Courcelle. Axiomatizations of betweenness in order-theoretic trees. 2020. hal-02205829v2

HAL Id: hal-02205829

<https://hal.science/hal-02205829v2>

Preprint submitted on 21 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Axiomatizations of betweenness in order-theoretic trees

Bruno COURCELLE
LaBRI, CNRS
and Bordeaux University
courcell@labri.fr

April 21, 2020

Abstract

The ternary *betweenness relation* of a tree, $B(x, y, z)$, indicates that y is on the unique path between x and z . This notion can be extended to *order-theoretic trees* defined as partial orders such that the set of nodes greater than any node is linearly ordered. In such generalized trees, the unique "path" between two nodes can have infinitely many nodes.

We generalize some results obtained in a previous article for the betweenness of *join-trees*. Join-trees are order-theoretic trees such that any two nodes have a least upper-bound. The motivation was to define conveniently the rank-width of a countable graph. We called *quasi-tree* the structure based on the betweenness relation of a join-tree. We proved that quasi-trees are axiomatized by a first-order sentence.

Here, we obtain a monadic second-order axiomatization of betweenness in order-theoretic trees. We also define and compare several *induced betweenness relations*, *i.e.*, restrictions to sets of nodes of the betweenness relations in generalized trees of different kinds. We prove that induced betweenness in quasi-trees is characterized by a first-order sentence. The proof uses order-theoretic trees.

All trees and related structures are finite or countably infinite.

Keywords : Betweenness, order-theoretic tree, join-tree, first-order logic, monadic second-order logic, quasi-tree.

Introduction

The *rank-width* $rdw(G)$ of a finite graph G defined by Oum in [13], is a complexity measure based on ternary trees whose leaves hold the vertices. If H is an induced subgraph of G , then $rdw(H) \leq rdw(G)$. In order to define the rank-width of a countable graph in such a way that it be the least upper-bound of those of its finite induced subgraphs, we have defined in [4] certain generalized (undirected) trees called *quasi-trees* such that the unique "path" between any two nodes can have infinitely many nodes. In particular, it can have the order-type of an interval of the set \mathbb{Q} of rational numbers. As no notion of adjacency can be used, we have defined them in terms of a notion of betweenness. The *betweenness relation* of a tree is the ternary relation B , such that $B(x, y, z)$ holds if and only if x, y, z are distinct and y is on the unique path between x and z . It can be extended to *order-theoretic trees* defined as partial orders such that the set of elements greater than any element is linearly ordered. A *join-tree* is an order-theoretic tree such that any two nodes have a *least upper-bound*, equivalently in this case, a *least common ancestor*. A join-tree may have no root, *i.e.*, no largest element. A quasi-tree is defined abstractly as a ternary structure $S = (N, B)$ satisfying finitely many first-order *betweenness axioms*. But quasi-trees are equivalently characterized as the betweenness relations of join-trees [4].

In this article we axiomatize in monadic second-order logic betweenness in order-theoretic trees¹. We also define and study several *induced betweenness relations*, *i.e.*, the restrictions to sets of nodes of betweenness relations in generalized trees of different kinds. An induced betweenness relation in a quasi-tree need not be that of a quasi-tree. However, induced betweenness relations in quasi-trees are also axiomatized by a single *first-order sentence*. This fact does not follow immediately from the first-order characterization of quasi-trees by a general logical argument. The proof that this axiomatization is valid uses order-theoretic trees.

We define actually four types of betweenness structures $S = (N, B)$ for which we prove that the inclusions that follow easily from the definitions are proper. For each type of betweenness, a structure S is defined from an order-theoretic tree T . Except for the case of induced betweenness in order-theoretic trees, some defining tree T can be described in S by monadic second-order formulas. In technical words, T is defined from S by a *monadic second-order transduction*, a notion thoroughly studied in [8]. The construction of a monadic second-order transduction for induced betweenness in quasi-trees is not straightforward. It is based on a notion of *structuring* of order-theoretic trees already used in [3, 4, 5]. In these articles, we obtained algebraic characterizations of the join-trees and quasi-trees that are the unique countable models of monadic-second order sentences².

¹All trees and related structures (except lines in the plane in the definition of topological trees) are finite or countably infinite.

²This type of characterization will be extended to order-theoretic trees in a future work.

In order to provide a concrete view of our generalized trees, we embed them into *topological trees*, defined as unions of possibly unbounded segments of straight lines in the plane that are connected but have no subset homeomorphic to a circle. Induced betweenness relations in topological trees and in quasi-trees are the same.

We study finite and countably infinite structures. Our main results concern the class **IBQT** of induced betweenness relations in quasi-trees and are the following ones:

- this class is first-order axiomatizable (Theorem 3.1),
- a join-tree witnessing that a ternary structure S is in **IBQT** can be constructed in S by monadic second-order formulas (Theorem 3.25),
- induced betweenness relations in topological trees and in quasi-trees are the same (Theorem 4.4).

Betweenness in other structures has been studied. In *partial orders* it is axiomatized by an infinite set of first-order sentences in [12], that cannot be replaced by a finite one [7]. In this article, we axiomatize betweenness in partial orders by a *single monadic second-order* sentence. Although we define generalized trees as partial orders, we do not use here this notion of betweenness. Several notions of betweenness in *graphs* have also been investigated and axiomatized. We only refer to the survey [1] that contains a rich bibliography. Another reference is [2].

Summary: We review definitions and notation in Section 1. We define four different notions of betweenness in order-theoretic trees in Section 2. We establish in Section 3 the first-order and monadic second-order axiomatizations presented above. The case of induced betweenness in order-theoretic trees is left as a conjecture. We also examine whether monadic second-order transductions can produce witnessing trees from given betweenness structures. In Section 4, we describe embeddings of join-trees into topological trees. In an appendix we give an example of a first-order class of relational structures (actually of labelled graphs) whose induced substructures do not form a first-order (and even a monadic second-order) axiomatizable class.

1 Definitions and basic facts

All trees, graphs and logical structures are countable, which means, finite or countably infinite. We will not repeat this hypothesis in our statements, except for emphasis.

In some cases, we denote by $X \uplus Y$ the union of sets X and Y to insist that they are disjoint. Isomorphism of ordered sets, trees, graphs and other structures is denoted by \simeq . We denote by $[n]$ the set of integers $\{1, \dots, n\}$.

The *arity* of a relation R is $\rho(R)$. The restriction of a relation R defined on a set V to a subset X of V is denoted by $R[X]$. If S is an $\{R_1, \dots, R_k\}$ -structure (N, R_1, \dots, R_k) , then $S[X] := (X, R_1[X], \dots, R_k[X])$.

The *Gaifman graph* of $S = (N, R_1, \dots, R_k)$ has vertex set N and an edge between x and $y \neq x$ if and only if x and y belong to a same tuple of some relation R_i . We say that S is *connected* if its Gaifman graph is connected. If it is not, S is the disjoint union of connected structures, each of them corresponding to a connected component of its Gaifman graph.

A class of relational structures is *first-order* (resp. *monadic second-order*) *axiomatizable* or *definable* if there exists a single first-order (resp. monadic second-order) sentence whose countable models form this class. See Section 1.4 for details.

1.1 Partial orders

For partial orders $\leq, \preceq, \sqsubseteq, \dots$ we denote respectively by $<, \prec, \sqsubset, \dots$ the corresponding strict partial orders. We write $x \perp y$ if x and y are incomparable for the considered order.

Let (V, \leq) be a partial order. For $X, Y \subseteq V$, the notation $X < Y$ means that $x < y$ for every $x \in X$ and $y \in Y$. We write $X < y$ instead of $X < \{y\}$ and similarly for $x < Y$. We use similar notation for \leq and \perp . The least upper-bound of x and y is denoted by $x \sqcup y$ if it exists and is called their *join*.

An *interval* X of (V, \leq) is a *convex subset*, i.e., $y \in X$ if $x < y < z$ and $x, z \in X$. If $X \subseteq V$, then $N_{\leq}(X) := \{y \in V \mid y \leq x \text{ for every } x \in X\}$ (hence $N_{\leq}(X) \leq X$) and $\downarrow(X) := \{y \in V \mid y \leq x \text{ for some } x \in X\}$.

Let (N, \leq) and (N', \leq') be partial orders. An *embedding* $j : (N, \leq) \rightarrow (N', \leq')$ is an injective mapping such that $x \leq y$ if and only if $j(x) \leq' j(y)$; in this case, (N, \leq) is isomorphic by j to $(j(N), \leq'')$, where \leq'' is the restriction of \leq' to $j(N)$; we will write more simply $(j(N), \leq')$. We say that j is a *join-embedding* if, furthermore, $j(x \sqcup y) = j(x) \sqcup' j(y)$ whenever $x \sqcup y$ is defined.

1.2 Trees

A *tree* is a possibly empty, undirected graph that is connected and has no cycles. Hence, it has neither loops nor multiple edges³. The set of nodes of a tree T is denoted by N_T .

A *rooted tree* is a nonempty tree equipped with a distinguished node called its *root*. We define on N_T the partial order \leq_T such that $x \leq_T y$ if and only if y is on the unique path between x and the root. The least upper-bound of x and y , denoted by $x \sqcup_T y$ is their least common ancestor. The minimal nodes are the *leaves*, and the root is the greatest node.

We will specify a rooted tree T by (N_T, \leq_T) and we will omit the index T when the considered tree is clear.

³No two edges with same ends.

Fact : A partial order (N, \leq) is (N_T, \leq_T) for some rooted tree T if and only if it has a largest element and, for each $x \in N$, the set $L_{\geq}(x) := \{y \in N \mid y \geq x\}$ is finite and linearly ordered. These conditions imply that any two nodes have a join.

1.3 Order-theoretic forests and trees

Definition 1.1 : *O-forests and O-trees.*

In order to have a simple terminology, we will use the prefix O- to mean *order-theoretic* and to distinguish these generalized trees from those of [5].

(a) An *O-forest* is a pair $F = (N, \leq)$ such that:

1) N is a possibly empty set called the set of *nodes*,

2) \leq is a partial order on N such that, for every node x , the set $L_{\geq}(x)$ is linearly ordered.

It is called an *O-tree* if furthermore:

3) every two nodes x and y have an upper-bound.

An O-forest is thus the union of the disjoint O-trees that are its connected components with respect to its Gaifman graph. Two nodes are in a same composing O-tree if and only if they have an upper-bound.

The *leaves* are the minimal elements. If N has a largest element r ($x \leq r$ for all $x \in N$) then F is a *rooted O-tree* and r is its *root*.

(b) A *line* in an O-forest (N, \leq) is a linearly ordered subset L of N that is *convex*, i.e., such that $y \in L$ if $x, z \in L$ and $x < y < z$. A subset X of N is *upwards closed* (resp. *downwards closed*) if $y \in X$ whenever $y > x$ (resp. $y < x$) for some $x \in X$. In an O-forest, the set $L_{>}(X)$ of *strict upper-bounds* of a nonempty set $X \subseteq N$, defined as $\{y \in N \mid y > X\}$ is an upwards closed line L .

(c) An O-tree T is a *join-tree*⁴ if every two nodes x and y have a least upper-bound denoted by $x \sqcup_T y$ and called their *join* (cf. Subsection 1.1). In a join-tree, every finite set has a least upper-bound, but an infinite one may have none.

(d) Let $J = (N, \leq)$ be an O-forest and $X \subseteq N$. Then $J[X] := (X, \leq)$ is an O-forest⁵. It is the *sub-O-forest* of J *induced on* X . Two elements x, y having a join in J may have no join in $J[X]$, or they may have a different join. If J is an O-tree, $J[X]$ may not be an O-tree. \square

Examples 1.2 :

(1) If T is a rooted tree, then (N_T, \leq_T) is a join-tree. Every finite O-tree is a join-tree of this form.

(2) Every linear order is a join-tree.

(3) Let $S := \mathbb{N} \cup \{a, b, c\}$ be ordered⁶ by $<_S$ such that $a <_S b, c <_S b$ and $b <_S i <_S j$ for all $i, j \in \mathbb{N}$ such that $j < i$, and a and c are incomparable. Then

⁴An *ordered tree* is a rooted tree such that the set of sons of any node is linearly ordered. This notion is extended in [5] to join-trees. Ordered join-trees should not be confused with order-theoretic trees, that we call O-trees for simplicity.

⁵We recall from Subsection 1.1 that the notation \leq is used for the restriction of \leq to X .

⁶The notation $<_S$ indicates that we define a strict partial order. The corresponding partial order will be denoted by \leq_S .

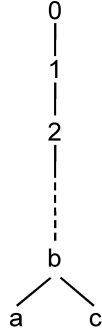


Figure 1: The join-tree of Example 1.2(3) and 3.4(a).

$T := (S, \leq_S)$ is a join-tree, see Figure 1. In particular $a \sqcup_S c = b$. The relation \leq_S is not the partial order associated with any rooted tree. We can consider $\mathbb{N} \cup \{a, b\}$ as forming a "path" in the join-tree T between a and 0 , the largest element. A formal definition of paths will be given.

If $S' := S - \{b\}$ is ordered by $<_S$, we have an O-tree with set of nodes S' . It is not a join-tree because a and c have no join.

(4) Fraïssé has defined in [9] (Section 10.5.3) a binary join-tree $T := (Seq_+(\mathbb{Q}), \preceq)$ where $Seq_+(\mathbb{Q})$ is the set of finite nonempty sequences of rational numbers partially ordered by \preceq as follows :

$$(x_n, \dots, x_0) \preceq (y_m, \dots, y_0) \text{ if and only if} \\ n \geq m, (x_{m-1}, \dots, x_0) = (y_{m-1}, \dots, y_0) \text{ and } x_m \leq y_m.$$

In particular, for all x_n, \dots, x_0, y we have $(x_n, \dots, x_0) \prec (x_{n-1}, \dots, x_0)$ and $(y, x_{n-1}, \dots, x_0) \prec (x_n, \dots, x_0)$ if and only if $y < x_n$. The strict partial order \prec is generated by transitivity from these particular relations.

The join of incomparable nodes (x_n, \dots, x_0) and (y_m, \dots, y_0) is (x'_p, \dots, x_0) such that $p \leq n, p < m, (y_p, \dots, y_0) = (x'_p, \dots, x_0)$ and $x_p < x'_p$. Examples of lines are $\{(x) \mid x \in \mathbb{Q}\}$, $\{(x, x_0) \mid x \in \mathbb{Q}\}$ and, for each $x_0 \in \mathbb{Q}$, $\{(x, x_0), (y) \mid y, x \in \mathbb{Q}, y \geq x_0\}$. See also Examples 3.6 and 3.28.

Every O-tree (N, \leq) is isomorphic to $T[X]$ for some subset X of $Seq_+(\mathbb{Q})$. \square

Definitions 1.3: *Extending and completing an O-forest.*

Let $F = (N, \leq)$ be an O-forest.

(a) Let \mathcal{C} a countable family of downwards closed nonempty subsets of N that is *nonoverlapping* : if two sets have a nonempty intersection, then, one is included in the other. We define $F(\mathcal{C}) := (\mathcal{C}, \subseteq)$. It is an O-forest.

Let $j : N \rightarrow \mathcal{P}(N)$ be such that $j(x) := N_{\leq}(x)$ where $N_{\leq}(x)$ denotes $\{y \in N \mid y \leq x\}$. The family of sets $j(x)$, denoted by $j(N)$, is countable, nonoverlapping and its elements are downwards closed in F . The mapping j is an isomorphism: $F \rightarrow F(j(N))$. If a family \mathcal{C} as above is nonoverlapping and contains $j(N)$, then j is an embedding : $F \rightarrow F(\mathcal{C})$. Hence, \mathcal{C} defines an

extension of F . The joins are not necessarily preserved by j . We will use this construction to "add" joins to O-trees.

(b) For every two, possibly equal, nodes x, y , we let $U(x, y) := N_{\leq}(L_{\geq}(x, y))$. It is the set of nodes z such that $z \leq u$ for every $u \geq \{x, y\}$. We have $\{x, y\} \subseteq U(x, y)$. If $x \leq y$, then $U(x, y) = N_{\leq}(y)$. If $x \sqcup y$ is defined, then $U(x, y) = N_{\leq}(x \sqcup y)$. If x and y have no upper-bound, then $U(x, y) = N_{\leq}(\emptyset) = N$.

The family \mathcal{U} of sets $U(x, y)$ is countable. It is nonoverlapping: if $z \in U(x, y) \cap U(x', y')$ then $L_{\geq}(x, y) \subseteq L_{\geq}(x', y')$ or vice-versa ; if $L_{\geq}(x, y) = L_{\geq}(x', y')$ then $U(x, y) = U(x', y')$ and if $L_{\geq}(x, y) \subset L_{\geq}(x', y')$ there is w in $L_{\geq}(x', y') - L_{\geq}(x, y)$ and we have $U(x', y') \subseteq N_{\leq}(w) = U(w, w) \subseteq U(x, y)$. Hence $F(\mathcal{U})$ is an O-tree. It is even a join-tree : if $x \sqcup y$ is defined, then, $N_{\leq}(x \sqcup y)$ identified with $x \sqcup y$, is $x \sqcup_{F(\mathcal{U})} y$; otherwise, $x \sqcup_{F(\mathcal{U})} y = U(x, y)$. This fact is easy to check, as is the nonoverlapping condition.

We call $F(\mathcal{U})$ the *join-completion* of F . We denote it by \widehat{F} . Its construction adds to F the "missing joins". The existing joins are preserved. It follows that every O-forest F with set of nodes N is $\widehat{F}[N]$ where \widehat{F} is a join-tree.

1.4 Monadic second-order logic

We will express properties of relational structures by first-order (FO in short) and monadic second-order (MSO) formulas and sentences. Logical structures are relational (they have only relation symbols) and countable.

Definitions 1.4 : *Quick review of terminology and notation.*

Monadic second-order logic extends first-order logic by the use of *set variables* $X, Y, Z \dots$ denoting subsets of the domain of the considered logical structure. The atomic formula $x \in X$ expresses the membership of x in X . We call *first-order* a formula where set variables are not quantified. For example, a first-order formula can express that $X \subseteq Y$. A *sentence* is a formula without free variables.

A property P of \mathcal{R} -structures where \mathcal{R} is a finite set of relation symbols, is *first-order* or *monadic second-order expressible* (*FO* or *MSO expressible*) if it is equivalent to the validity, in every \mathcal{R} -structure S , of a first-order or monadic second-order sentence φ . The validity of φ in S is denoted by $S \models \varphi$. We say that a property of tuples of subsets X_1, \dots, X_n of the domains of structures in a class \mathcal{C} is *FO* or *MSO definable* if it is equivalent to $S \models \varphi(X_1, \dots, X_n)$ in every \mathcal{R} -structure S in \mathcal{C} , where φ is a fixed FO or MSO formula with n free set variables. A class of structures is thus *FO* or *MSO definable* or *axiomatizable* if it is characterized by an FO or MSO sentence

Transitive closures and choices of sets, typically in graph coloring problems, are MSO but not FO expressible. See [8] for a detailed study of MSO expressible graph properties. Other comprehensive books are [10, 11].

Examples 1.5 : *Partial orders and graphs.*

(1) A simple undirected graph G can be identified with the $\{edg\}$ -structure (V_G, edg_G) where V_G is its vertex set and $edg_G(x, y)$ means that there is an edge

between x and y if G . For example, 3-colorability is expressed by the MSO sentence :

$$\exists X, Y [X \cap Y = \emptyset \wedge \neg \exists u, v (edg(u, v) \wedge [(u \in X \wedge v \in X) \vee (u \in Y \wedge v \in Y) \wedge (u \notin X \cup Y \wedge v \notin X \cup Y)])].$$

(2) We now consider partial orders (N, \leq) . The FO formula $Lin(X)$ defined as $\forall x, y. [(x \in X \wedge y \in X) \implies (x \leq y \vee y \leq x)]$ expresses that a subset X of N , partially ordered by \leq , is linearly ordered. The MSO formula

$$Lin(X) \wedge \exists a, b. [Min(X, a) \wedge Max(X, b) \wedge \theta(X, a, b)]$$

expresses that X is linearly ordered and finite, where $Min(X, a)$ and $Max(X, b)$ are FO formulas expressing respectively that X has a least element a and a largest one b , and $\theta(X, a, b)$ is an MSO formula expressing that :

- (i) each element x of X except b has a successor c in X (*i.e.*, c is the least element of $\{y \in X \mid y > x\}$), and
- (ii) $(a, b) \in Suc^*$, where Suc is the above defined successor relation (depending on X) and Suc^* is its reflexive and transitive closure.

Assertion (ii) is expressed by the MSO formula:

$$\forall U [U \subseteq X \wedge a \in U \wedge \forall x, y. ((x \in U \wedge (x, y) \in Suc) \implies y \in U) \implies b \in U].$$

First-order formulas expressing $U \subseteq X$, $(x, y) \in Suc$ and Property (i) are easy to write. The finiteness of a linear order is not FO expressible⁷. Without a linear order, the finiteness of a set X is not MSO expressible.

Definitions 1.6 : *Transformations of relational structures.*

As in [8], we call *transduction* a transformation of relational structures specified by logical formulas⁸. We will try to be not too formal but nevertheless precise.

(a) The basic type of transduction τ is as follows. A structure $S' = (D', R'_1, \dots, R'_m)$ is defined from a structure $S = (D, R_1, \dots, R_n)$ and a p -tuple (X_1, \dots, X_p) of subsets of D called *parameters* by means of formulas $\chi, \delta, \theta_{R'_1}, \dots, \theta_{R'_m}$ used as follows:

$$\begin{aligned} \tau(S, (X_1, \dots, X_p)) = S' \text{ is defined if and only if } S \models \chi(X_1, \dots, X_p), \\ S' = (D', R'_1, \dots, R'_m) \text{ has domain } D' \subseteq D \text{ such that } d \in D' \text{ if and} \\ \text{only if } S \models \delta(X_1, \dots, X_p, d), \\ R'_i \text{ is the set of tuples } (d_1, \dots, d_s) \in D'^s, s = \rho(R'_i), \text{ such that } S \models \\ \theta_{R'_i}(X_1, \dots, X_n, d_1, \dots, d_s). \end{aligned}$$

⁷Follows from the Compactness Theorem for FO logic [10].

⁸The usual terminology of *interpretation* is inconvenient as it is frequently unclear what is defined from what. The term *transduction* is borrowed to formal language theory that is concerned with transformations of words, trees and terms. There are deep links between monadic second-order definable transductions and tree transducers [8].

We call τ an FO or an MSO transduction if the formulas that define it are, respectively, first-order or monadic second-order ones.

As an example, the mapping from a graph $G = (V, \text{edg})$ to the connected component $(V', \text{edg}[V'])$ containing a vertex u is defined by χ, δ and θ_{edg} where $\chi(X)$ expresses that X is a singleton $\{u\}$, $\delta(X, d)$ expresses that there is a path between d and the vertex in X , and $\theta_{\text{edg}}(x, y)$ is the formula always **true**, say, $x = x$. It is an MSO transduction as path properties are expressible by monadic second-order formulas.

(b) Transductions of the general type may enlarge the domain of the input structure. A structure $S' = (D', R'_1, \dots, R'_m)$ is defined from $S = (D, R_1, \dots, R_n)$ and a p -tuple (X_1, \dots, X_p) of parameters as above by means of formulas $\chi, \delta_1, \dots, \delta_k$ and others, $\theta_{R'_i, i_1, \dots, i_s}$, used as follows:

$$\begin{aligned} \tau(S, (X_1, \dots, X_p)) = S' \text{ is defined if and only if } S \models \chi(X_1, \dots, X_p), \\ S' = (D', R'_1, \dots, R'_m) \text{ has domain } D' \subseteq (D \times \{1\}) \uplus \dots \uplus (D \times \{k\}) \\ \text{such that } (d, i) \in D' \text{ if and only if } S \models \delta_i(X_1, \dots, X_p, d), \\ R'_i \text{ is the set of tuples } ((d_1, i_1), \dots, (d_s, i_s)) \in D'^s, s = \rho(R'_i), \text{ such} \\ \text{that} \end{aligned}$$

$$S \models \theta_{R'_i, i_1, \dots, i_s}(X_1, \dots, X_p, d_1, \dots, d_s).$$

If D is finite, then $|D| \leq k|D'|$.

An easy example consists in the *duplication* of a graph $G = (V, \text{edg})$ into the graph $H := G \oplus G$, that is G together with a disjoint copy of it. We get a graph H up to isomorphism, because of the use of disjoint isomorphic copies. To define a transduction, we take $k = 2, p = 0$ (no parameter is needed), χ, δ_1, δ_2 always **true**, $\theta_{\text{edg}, i, j}(x, y)$ always **false** if $i \neq j$, and equal to $\text{edg}(x, y)$ if $i = j$, where $i, j \in [2]$.

Another more complicated example is the transformation of an O-forest $F = (N, \leq)$ into its join-completion \widehat{F} . We define concretely the set of nodes of \widehat{F} as $(N \times \{1\}) \uplus (M \times \{2\})$ where M is a subset of N in bijection with the set of sets $U(x, y)$ such that x and y have no join, cf. Example 1.3. This bijection can be made MSO definable, and so is the order relation of \widehat{F} . Defining M is not straightforward because the sets $U(x, y)$ are not pairwise disjoint. We can use the notion of *structuring of an O-tree*, see Remark 3.35.

2 Quasi-trees and betweenness in O-trees

In this section, we define a *betweenness relation* in O-trees, and compare it with the *betweenness relation induced* by sets of nodes in join-trees or O-trees. We generalize the notion of quasi-tree defined and studied in [4] and [5].

For a ternary relation B on a set N and $x, y \in N$, we define $[x, y]_B := \{x, y\} \cup \{z \in N \mid (x, z, y) \in B\}$. If $n > 2$, then the notation $\neq (x_1, x_2, \dots, x_n)$ means that x_1, x_2, \dots, x_n are pairwise distinct (hence abbreviates an FO formula).

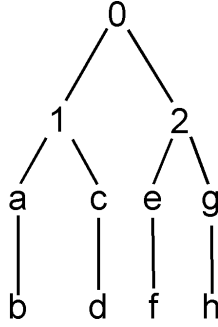


Figure 2: A rooted tree R cf. Example 2.2.

2.1 Betweenness in trees and quasi-trees

Definition 2.1 : *Betweenness in linear orders and in trees.*

(a) Let $L = (X, \leq)$ be a linear order. Its *betweenness relation*⁹ B_L is the ternary relation on X defined by :

$$B_L(x, y, z) :\iff x < y < z \text{ or } z < y < x.$$

(b) If T is a tree or a forest, its *betweenness relation* B_T is the ternary relation on N_T defined by :

$$B_T(x, y, z) :\iff x, y, z \text{ are pairwise distinct and } y \text{ is on the unique path between } x \text{ and } z.$$

If $R = (N, \leq_R)$ is a rooted tree, we define its *betweenness relation* B_R as $B_{Und(R)}$ where $Und(R)$ is the tree obtained from R by forgetting its root and its edge directions. For all $x, y, z \in N$, we have :

$$B_R(x, y, z) \iff x, y, z \text{ are pairwise distinct, } x \text{ and } z \text{ have a join } x \sqcup_R z \text{ and } x <_R y \leq_R x \sqcup_R z \text{ or } z <_R y \leq_R x \sqcup_R z. \square$$

Example 2.2 : Figure 2 shows a rooted tree R with root 0. For illustrating the above observation, we have $B_R(b, a, 0)$ and $b < a < 0 = b \sqcup 0$, and also $B_R(b, a, c)$ and $b < a < 1 = b \sqcup c$. \square

With a ternary relation B on a set X , we associate the ternary relation A on X : $A(x, y, z) :\iff B(x, y, z) \vee B(x, z, y) \vee B(y, x, z)$, to be read : x, y, z

⁹This definition can be used for partial orders. The corresponding notion of betweenness is axiomatized in [12, 7]. We will *not* use it for defining betweenness in order-theoretic trees, although these trees are partial orders.

are *aligned*. If $n \geq 3$, then $B^+(x_1, x_2, \dots, x_n)$ stands for the conjunction of the conditions $B(x_i, x_j, x_k)$ for all $1 \leq i < j < k \leq n$ and all $1 \leq k < j < i \leq n$.

The following is Proposition 5.2 in [5] or Proposition 9.1 in [8].

Proposition 2.3 : (a) The betweenness relation B of a linear order (X, \leq) satisfies the following properties for all $x, y, z, u \in X$.

$$A1 : B(x, y, z) \Rightarrow \neq(x, y, z).$$

$$A2 : B(x, y, z) \Rightarrow B(z, y, x).$$

$$A3 : B(x, y, z) \Rightarrow \neg B(x, z, y).$$

$$A4 : B(x, y, z) \wedge B(y, z, u) \Rightarrow B(x, y, u) \wedge B(x, z, u).$$

$$A5 : B(x, y, z) \wedge B(x, u, y) \Rightarrow B(x, u, z) \wedge B(u, y, z).$$

$$A6 : B(x, y, z) \wedge B(x, u, z) \Rightarrow y = u \vee [B(x, u, y) \wedge B(u, y, z)] \\ \vee [B(x, y, u) \wedge B(y, u, z)].$$

$$A7' : \neq(x, y, z) \Rightarrow A(x, y, z).$$

(b) The betweenness relation B of a tree T satisfies the properties A1-A6 for all x, y, z, u in N_T together with the following weakening of A7':

$$A7 : \neq(x, y, z) \Rightarrow A(x, y, z) \vee \exists w. (B(x, w, y) \wedge B(y, w, z) \wedge B(x, w, z)).$$

Remarks 2.4.

(1) Property A7' says that if x, y, z are three elements in a linear order, then, one of them is between the two others. Properties A1-A5 belong to the axiomatization of betweenness in partial orders given in [7, 12]. Property A6 is actually a consequence of Properties A1-A5 and A7', as one proves easily.

(2) Property A7 says that, in a tree T , if x, y, z are three nodes not on a same path, some node w is between any two of them. In this case, we have :

$$\{w\} = P_{x,y} \cap P_{y,z} \cap P_{x,z} \text{ where } P_{x,y} \text{ is the set of nodes on the path} \\ \text{between } x \text{ and } y,$$

so that we have $B(x, w, y) \wedge B(y, w, z) \wedge B(x, w, z)$.

If T is a rooted tree, and x, y, z are not on a path from a leaf to the root, then w is the join (the least common ancestor) of two nodes among x, y, z . In the rooted tree R of Figure 2, we have, for example, $w = 1$ if $x = a, y = d$ and $z = e$.

Property A6 is a consequence of Properties A1-A5 and A7.

(3) Properties A1-A6 (for an arbitrary structure $S = (N, B)$) imply that the two cases of the conclusion of A7 are exclusive¹⁰ and that, in the second one, there is a unique node w satisfying $B(x, w, y) \wedge B(y, w, z) \wedge B(x, w, z)$ (by Lemma 11 of [4]), that is denoted by $M_S(x, y, z)$. \square

¹⁰The three cases of $A(x, y, z)$ are exclusive by A2 and A3.

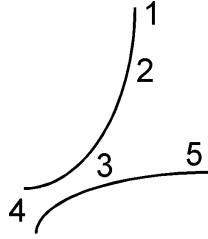


Figure 3: Structure S of Example 2.6(1)

Definitions 2.5 : *Other betweenness properties.*

The letter B and its variants, B_T, B_1 , etc. will denote ternary relations.

We define the following properties of a structure $S = (N, B)$:

$$\text{A8} : \forall u, x, y, z. [\neq(u, x, y, z) \wedge B(x, y, z) \Rightarrow \\ B(u, x, y) \vee B(u, y, z) \vee B(x, y, u) \vee B(y, z, u)].$$

$$\text{A8}' : \forall u, x, y, z. [\neq(u, x, y, z) \wedge B(x, y, z) \wedge \neg A(y, z, u) \Rightarrow B(x, y, u)].$$

If S satisfies A1-A6, the four cases of the conclusion of A8 are not exclusive : $B(u, x, y)$ implies $B(u, y, z)$ (because of $B(x, y, z)$ and A4). \square

Example and remark 2.6 :

(1) Properties A1-A6 do not imply A8'. Consider $S := ([5], B)$ where B satisfies (only) $B^+(1, 2, 3, 4) \wedge B^+(5, 3, 4)$ illustrated in Figure 3. (There is no curve line going through 1,2,5 because $B(1, 2, 5)$ is not assumed to be valid). Conditions A1-A6 hold but A8' does not, because we have $\neg A(2, 3, 5) \wedge B(1, 2, 3)$. Then, A8' would imply $B(1, 2, 5)$ that is not assumed. By the next lemma, A1-A6 do not imply A8 either.

(2) Properties A1-A5 and A8' imply A6.

Assume we have $B(x, y, z) \wedge B(x, u, z) \wedge y \neq u$. If $\neg A(y, z, u)$, then by A8' we have $B(x, y, u)$ and so, $B^+(x, y, u, z)$ but then we have $B(y, u, z)$ hence $A(y, z, u)$. Hence, we have $A(y, z, u)$, that is, $B(y, z, u)$ or $B(z, y, u)$ or $B(y, u, z)$.

If $B(y, z, u)$ holds, then we have $B^+(x, y, z, u)$ hence $B(x, z, u)$ that contradicts $B(x, u, z)$. If $B(z, y, u)$ holds, then we have $B^+(x, u, y, z)$ hence $B(x, u, y)$, one case of the conclusion. The last case is $B(y, u, z)$, the other case of the conclusion.

We keep Property A6 in our axiomatization for its clarity and to shorten proofs. \square

In the following proofs and discussions about a structure (N, B) , we will always assume (unless otherwise specified) that Properties A1-A6 hold, and we will not make their use explicit. We say that (N, B) is *trivial* if $B = \emptyset$. In this case, Properties A1-A6, A8 and A8' hold.

Lemma 2.7 : Let $S = (N, B)$ satisfy A1-A6.

(1) A8 is equivalent to A8'.

(2) A7 implies A8, and thus, A8'.

(3) If A8 holds, then the Gaifman graph¹¹ of S is either edgeless (if $B = \emptyset$) or connected.

Proof: (1) If $B = \emptyset$, then A8 and A8' both holds. Otherwise, assume that A8 holds and that we have $\neq (u, x, y, z) \wedge B(x, y, z) \wedge \neg A(u, y, z)$. Then, A8 yields the following possibilities :

(1.1) $B(u, x, y)$: we have $B^+(u, x, y, z)$, which implies $B(u, y, z)$, and thus $A(u, y, z)$,

(1.2) $B(u, y, z)$, which implies $A(u, y, z)$,

(1.3) $B(y, z, u)$, which implies $A(u, y, z)$.

These three cases cannot hold since we assume $\neg A(u, y, z)$. The only remaining case is :

(1.4) $B(x, y, u)$: this is the desired conclusion.

Hence, A8' is valid.

Conversely, assume that A8' holds and we have $\neq (u, x, y, z) \wedge B(x, y, z)$.

If $A(u, y, z)$ holds, we have $B(y, u, z) \vee B(u, y, z) \vee B(y, z, u)$. Because of $B(x, y, z)$, $B(y, u, z)$ implies $B(x, y, u)$. Hence we have $B(x, y, u) \vee B(u, y, z) \vee B(y, z, u)$. If $\neg A(u, y, z)$, then A8' yields $B(x, y, u)$, and the desired fact.

(2) We prove that A7 entails A8'. Assume we have $\neq (u, x, y, z) \wedge B(x, y, z) \wedge \neg A(u, y, z)$. There is w such that $B(u, w, y) \wedge B(y, w, z) \wedge B(u, w, z)$. With $B(x, y, z)$, we get : $B^+(x, y, w, z)$, hence, $B(x, y, w)$. With $B(y, w, u)$, we get $B(x, y, u)$, as desired.

(3) Clear from definitions. \square

Definition 2.8 : *Quasi-trees* [4].

(a) A *quasi-tree* is a structure $S = (N, B)$ such that B is a ternary relation on a set N , called the set of *nodes*, that satisfies conditions A1-A7. To avoid uninteresting special cases, we also require that $|N| \geq 3$. We say that S is *discrete* if $[x, y]_B := \{x, y\} \cup \{z \in N \mid B(x, z, y)\}$ is finite for all x, y .

(b) From a join-tree $J = (N, \leq)$, we define a ternary relation B_J on N by :

$$B_J(x, y, z) : \iff \neq (x, y, z) \wedge ([x < y \leq x \sqcup z] \vee [z < y \leq x \sqcup z]),$$

called its *betweenness relation*. As a definition, we use here the observation made for rooted trees in Definition 2.1. The join $x \sqcup z$ is always defined.

(c) In a quasi-tree $S = (N, B)$, the *path* that links x and y is the set $[x, y]_B$. It is linearly ordered with least element x and largest one y . It may be infinite in this case, an element may have no successor or no predecessor. \square

Figure 4 shows a quasi-tree, where the dashed lines represent infinite paths in the above sense. In such a structure, no adjacency notion is available. The ternary notion of betweenness is an alternative.

Theorem 2.9 [Proposition 5.6 of [5]] :

¹¹Defined in Section 1.

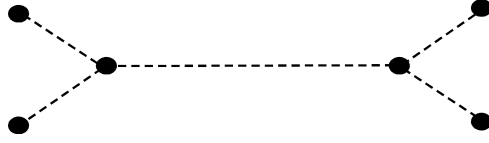


Figure 4: A quasi-tree.

(1) The structure $qt(J) := (N, B_J)$ associated with a join-tree $J = (N, \leq)$ with at least 3 nodes is a quasi-tree. Conversely, every quasi-tree S is $qt(J)$ for some join-tree J .

(2) A quasi-tree is discrete if and only if it is $qt(J)$ for the join-tree $J := (N_R, \leq_R)$ where R is a rooted tree.

This theorem shows that one can specify a quasi-tree by a binary relation, actually a partial order. However, this is inconvenient because choosing a partial order breaks the symmetry. This motivates the use of a ternary relation. Similarly, betweenness can formalize the notion of a linear order, *up to reversal*.

2.2 Other betweenness structures

Definition 2.10 : *Induced betweenness in a quasi-tree*

If $Q = (N, B)$ is a quasi-tree, $X \subseteq N$, we say that $Q[X] := (X, B[X])$ is an *induced betweenness relation in Q* . It is *induced on X* . \square

Remark and example 2.11: The structure $Q[X]$ need not be a quasi-tree because A7 does not hold for a triple (x, y, z) such that $M_Q(x, y, z)$ is not in X (cf. Proposition 2.3).

Figure 5 shows a tree T to the left, with $N_T = [7]$. Its betweenness relation B_T is expressed in a short way by the properties $B_T^+(1, 2, 7, 3, 4)$, $B_T^+(1, 2, 7, 5, 6)$ and $B_T^+(6, 5, 7, 3, 4)$. Let $Q := (N_T, B_T)$. The induced betweenness $S_1 := Q[6]$ is illustrated on the right, where the curve lines represent the facts $B_T^+(1, 2, 3, 4)$, $B_T^+(1, 2, 5, 6)$ and $B_T^+(6, 5, 3, 4)$. It is not a quasi-tree because $7 = M_Q(1, 4, 6)$ has been not kept in N_{S_1} . \square

Our objective is to axiomatize induced betweenness relations in quasi-trees (equivalently in join-trees), similarly as betweenness relations in join-trees¹² are by A1-A7 in Theorem 2.9(1).

Proposition 2.12 : An induced betweenness relation in a quasi-tree satisfies properties A1-A6 and A8.

¹²As in [4], we have defined quasi-trees (DEfinition 2.8) as the ternary structures that satisfy A1-A7. In the sequel, we will rather consider them as the betweenness relations of join-trees, and A1-A7 as their axiomatization.

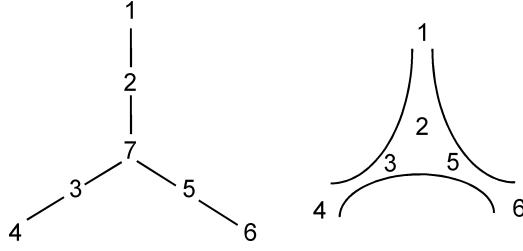


Figure 5: An induced betweenness in a quasi-tree, cf. Example 2.11.

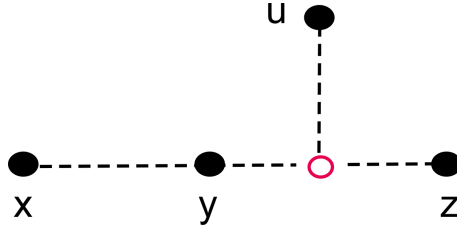


Figure 6: Illustration of Property A8'.

Proof: The sentences expressing A1-A6 and A8 are universal, that is, are of the form $\forall x, y, \dots, z. \varphi(x, y, \dots, z)$ where φ is quantifier-free. The validity of such sentences is preserved under taking induced substructures (we are dealing with relational structures). The result follows from Theorem 2.9 and Lemma 2.7(2) showing that a quasi-tree satisfies A8. \square

Our objective is to prove that a ternary relation is an induced betweenness in a quasi-tree if and only if it satisfies Properties A1-A6 and A8. Our proof will use O-trees.

Figure 6 illustrates Property A8' which says: $B(x, y, z) \wedge \neg A(y, z, u) \Rightarrow B(x, y, u)$. The white circle between y and z represents the node $M_Q(y, z, u)$ of a quasi-tree Q that has been deleted, so that Property A7 does not hold in the structure $Q[N - \{M_Q(y, z, u)\}]$.

Definition 2.13 : *Betweenness in O-forests.*

(a) The *betweenness relation* of an O-forest $F = (N, \leq)$ is the ternary relation B_F on N such that :

$$B_F(x, y, z) : \Leftrightarrow \neq (x, y, z) \wedge [(x < y \leq x \sqcup z) \vee (z < y \leq x \sqcup z)].$$

(b) If $F = (N, \leq)$ is an O-forest and $X \subseteq N$, then $(X, B_F[X])$ is an *induced betweenness relation* in F .

The difference with Definition 2.8(b) is that if x and z have no least upper-bound (*i.e.*, if $x \sqcup z$ is undefined, which implies that x and z are incomparable,

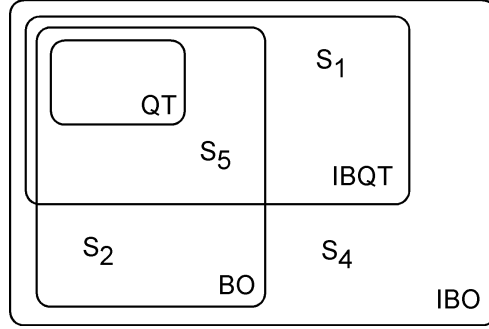


Figure 7: Proper inclusions of classes proved in Proposition 2.15.

denoted by $x \perp z$), then B_F contains no triple of the form (x, y, z) . If F is a finite O-tree, it is a join-tree and thus, (N, B_F) is a quasi-tree. \square

We have four classes of betweenness relations $S = (N, B)$: quasi-trees, induced betweenness relations in quasi-trees, betweenness and induced betweenness relations in O-forests, denoted respectively by **QT**, **IBQT**, **BO** and **IBO**.

Remarks 2.14 : (1) The induced betweenness (X, B) on a set X of leaves of a tree is *trivial*, which means that $B = \emptyset$.

(2) The Gaifman graph of a betweenness structure S is connected in the following cases : $S \in \mathbf{IBQT}$ and is not trivial (in particular $S \in \mathbf{QT}$), or S is the betweenness structure of an infinite O-tree. It may be not connected in the other cases.

(3) If S is an induced betweenness in an O-forest consisting of several disjoint O-trees, then two nodes in the different O-trees cannot belong to a same triple, hence, cannot be linked by a path in the Gaifman graph of S . Hence, a structure (N, B) is the betweenness of an O-forest, or an induced betweenness in an O-forest if and only if each of its connected components is so in an O-tree. We will only consider betweenness of O-trees (class **BO**) and induced betweenness in O-trees (class **IBO**).

Proposition 2.15 : We have the following proper inclusions :

$$\mathbf{QT} \subset \mathbf{IBQT}, \mathbf{QT} \subset \mathbf{BO} \subset \mathbf{IBO} \text{ and } \mathbf{QT} \subset \mathbf{IBQT} \cap \mathbf{BO}.$$

The classes **IBQT** and **BO** are incomparable. For finite structures, we have **QT** = **BO**. \square

These inclusions are illustrated in Figure 7. Structures S_1, S_2, S_4 and S_5 witnessing proper inclusions are described in the proof.

Proof: All inclusions are clear from the definitions. We give examples to prove that the inclusions are proper. We recall that $S[X] := (X, B[X])$ if $S = (N, B)$ and $X \subseteq N$.

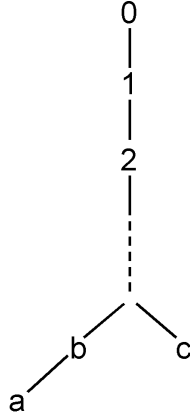


Figure 8: The O-tree T_2 used in the proof of Proposition 2.15, Parts (2) and (4).

(1) The structure S_1 of Example 2.11, shown in Figure 5, is in **IBQT** but not in **QT**. It is not in **BO** either, because otherwise, it would be a quasi-tree as it is finite.

(2) We consider $N_2 := \mathbb{N} \cup \{a, b, c\}$, the O-tree $T_2 := (N_2, \preceq)$ in Figure 8 such that $a \prec b \prec i \prec j$ and $c \prec i \prec j$ for all i, j in \mathbb{N} such that $j < i$. Its betweenness structure $S_2 := (N_2, B_2)$ is described by the properties $B_2^+(a, b, i, j, k)$ and $B_2^+(c, i, j, k)$ for all i, j, k in \mathbb{N} such that $k < j < i$. Since b and c have no least upper-bound in T_2 , we do not have $B_{T_2}(a, b, c)$. Hence, S_2 is in **BO** but not in **IBQT**, as it does not satisfy A8': we have $\neg A_{T_2}(0, b, c) \wedge B_{T_2}(a, b, 0)$ but not $B_{T_2}(a, b, c)$. The classes **IBQT** and **BO** are incomparable.

However, if we take c as new root, then we obtain a join-tree $U = (N_2, \preceq')$ where $a \prec' b \prec' c$ and $0 \prec' 1 \prec' 2 \dots \prec' i \prec' \dots \prec' c$. Clearly $B_U \neq B_{T_2}$. Hence, betweenness in O-trees depends some kind of orientation, that can be specified by a root, or by a line (cf. the notion of structuring used below). To the opposite, in the case of quasi-trees and induced betweenness in quasi-trees, any node can be taken as root in the constructions of the relevant join-trees (cf. [5] for quasi-trees, and the proof of Theorem 3.1 and Remark 3.4(d)).

(3) To prove that the inclusion of **BO** in **IBO** is proper, we consider $S_3 := (N_3, B_{T_3})$, $N_3 := \{a, b, c, d\} \cup \mathbb{Z}$ and the O-tree $T_3 := (N_3, \prec)$ with order: $a \prec b \prec i \prec j$ and $d \prec c \prec i \prec j$ for all $i, j \in \mathbb{N}$ such that $j < i$ and $i' \prec j' \prec i \prec j$ if $i, j \in \mathbb{N}$, $i', j' < 0$, $j < i$ and $j' < i'$. It is shown in Figure 9(a). We let then $S_4 := S_3[\{a, b, c, d, -1, 0, 1\}]$ with corresponding O-tree T_4 (Figure 9(b)). The structure S_4 is in **IBO** but not in **BO**. Otherwise, as it is finite, it would be a quasi-tree. But S_4 does not satisfy A8': we have $\neg A_{T_3}(0, b, c) \wedge B_{T_3}(a, b, 0)$ but $(a, b, c) \notin B_{T_3}$. For this reason, S_4 is not in **IBQT** either.

Note that S_4 in **IBO** is finite but is not the induced betweenness relation

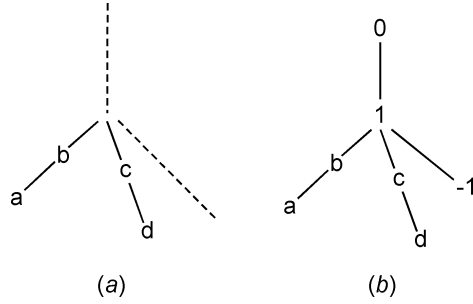


Figure 9: Part (a) shows T_3 and (b) shows T_4 of the proof of Proposition 2.15, Part (3), and Example (3.6).

of a *finite* O-tree. Otherwise, it would be in **IBQT** because a finite O-tree is a join-tree.

(4) Let T_5 be the O-tree $T_2[N_5]$ where $N_5 := \mathbb{N} \cup \{b, c\}$ and $S_5 := (N_5, B_{T_5})$. (Figure 8 shows T_2). It is in **BO**, and also in **IBQT** : just add to T_5 a least upper-bound m for b and c such that $m < \mathbb{N}$, one obtains a join-tree. It is not a quasi-tree because A7 does not hold for the triple $(0, b, c)$. Hence, we have **QT** \subset **IBQT** \cap **BO**.

Note that S_2 is not in **IBQT** but its induced substructure S_5 is. \square

Figure 7 shows how these examples are located in the different classes of betweenness relations. The structures S_1 and S_4 are finite, S_2 and S_5 are infinite, which is necessary because the finite structures in **BO** and **QT** are the same.

Remark 2.16 : An alternative notion of betweenness for an O-forest $F = (N, \leq)$ could be defined as $B'_F := B_{\widehat{F}}[N]$. As it is an induced betweenness in a join-tree, this definition does not bring anything new. If F is an O-tree, we have : $(x, y, z) \in B'_F$ if and only if $\neq (x, y, z)$ and, either $x < y \leq m \geq z$ or $z < y \leq m \geq x$ for some m that need not be the join of x and z .

3 Axiomatizations

3.1 First-order axiomatizations

Our first main result is Theorem 3.1 that provides a first-order axiomatization of the class **IBQT**, among countable (finite or countably infinite) structures. All our constructions are relative to countable structures.

3.1.1 Induced betweenness in quasi-trees

The letter B designates always ternary relations.

Theorem 3.1 : The class **IBQT** is axiomatized by the first-order properties A1-A6 and A8.

Lemma 3.2 : Let $S = (N, B)$ satisfy Axioms A1-A6 and $r \in N$. Let \leq_r be the binary relation on N such that $x \leq_r y \iff x = y \vee y = r \vee B(x, y, r)$.

- (1) $T(S, r) := (N, \leq_r)$ is an O-tree.
- (2) If $(x, y, z) \in B$, $x <_r y$ and $z <_r y$, then $y = x \sqcup_r z$.
- (3) If $(x, y, z) \in B$ and $x <_r w <_r y$, we do not have $z <_r w$.

Proof : (1) It is easy to check that \leq_r is a partial order and that for any $x \in N$, the set $\{y \in N \mid y \geq_r x\}$ is linearly ordered.

(2) Assume $(x, y, z) \in B$, $x <_r y$, $z <_r y$. We cannot have $x <_r z$ or $z <_r x$, because otherwise, we have $(x, z, y) \in B$ or $(z, x, y) \in B$, contradicting $(x, y, z) \in B$. Assume for a contradiction, that $x <_r w <_r y$ and $z <_r w <_r y$. Then, we have $(x, w, y) \in B$ and $(z, w, y) \in B$, whence $B^+(x, w, y, z)$, and $B^+(z, w, y, x)$, which gives $(w, y, z) \in B$ and $(z, w, y) \in B$, contradicting A2 \wedge A3.

(3) This assertion follows immediately. \square

Lemma 3.3 : Let $S := (N, B)$ satisfy A1-A6 and A8, and $r \in N$.

(1) Let x and y are incomparable with respect to \leq_r . If $z <_r y$, then $(x, y, z) \in B$.

- (2) If $(x, y, z) \in B$, then $x <_r y$ or $z <_r y$.
- (3) We have $B \subseteq B_{T(S, r)}$ if N is finite.

Proof : In this proof, $<$, \leq and \sqcup will denote $<_r, \leq_r$ and \sqcup_r .

(1) Let x and y are incomparable and $z <_r y$. The root r is not any of x, y, z . If $(x, r, y) \in B$, then, since $(r, y, z) \in B$, we have $B^+(x, r, y, z)$ hence $(x, y, z) \in B$. Otherwise, $A(x, y, r)$ does not hold, and as we have $(z, y, r) \in B$, we get $(z, y, x) \in B$ by A8[?], hence $(x, y, z) \in B$.

(2) Let $(x, y, z) \in B$. We have three cases.

Case 1 : The nodes x, y, z are pairwise incomparable w.r.t. $<$. Then, if $(x, r, y) \in B$, we have $B^+(x, r, y, z)$, hence $(r, y, z) \in B$ and $z <_r y$. Otherwise, $A(x, y, r)$ does not hold, hence by A8[?], we have $(z, y, r) \in B$, hence $z <_r y$.

Case 2 : x and y are comparable. If $x <_r y$, we are done. If $y <_r x$, we have $(y, x, r) \in B$, hence $B^+(z, y, x, r)$ which gives $z <_r y$. The case where z and y are comparable is similar because $(z, y, x) \in B$.

Case 3 : x and z are comparable. If $x <_r z$, we have $B(x, z, r)$, hence $B^+(x, y, z, r)$ which gives $x <_r y$. If $z <_r x$, the proof is similar.

(3) Let $(x, y, z) \in B$. As N is finite, x and z have a join $x \sqcup_r z$. We have $x <_r y$ or $z <_r y$ by (2). If $x <_r y$, there are two cases : if $y \leq_r x \sqcup_r z$, we have $(x, y, z) \in B_{T(S, r)}$; if $x \sqcup_r z <_r y$, we cannot have $x \sqcup_r z = y$ because then $(x, z, y) \in B$, and we cannot have $x \perp_r z$ because then $x <_r x \sqcup_r z <_r y$ and $z <_r x \sqcup_r z <_r y$ and $(x, y, z) \notin B$ by Lemma 3.2(2). The case $z <_r y$ is similar. \square

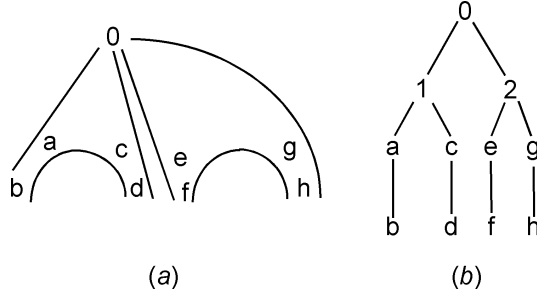


Figure 10: (a) shows S_8 and (b) shows T_8 , Example 3.4(c) and Remark 3.16.

Examples 3.4 : (a) In statement (3) above, we may have a proper inclusion. Consider S_6 defined as (N_6, B_6) with $N_6 := \{0, 1, 2, a, c\}$, $B_6^+(0, 1, 2, a)$, $B_6^+(0, 1, 2, c)$ and $r := 0$. Then $T(S_6, 0) = T[N_6]$ where T is the join-tree of Figure 1. We have $(a, 2, c)$ in $B_{T(S_6, 0)}$ but not in B_6 .

(b) The inclusion $B \subseteq B_{T(S, r)}$ may be false if S is infinite. Consider $S_7 = (\mathbb{N} \cup \{a, b, c\}, B_7)$ defined as S_2 in the proof of Proposition 2.15 (see Figure 8), augmented with the triples (a, b, c) and (c, b, a) . Then $T(S_7, 0) = T_2$ of this proof, but $(a, b, c) \notin B_{T(S_7, 0)}$.

(c) We give an example showing how we will prove Theorem 3.1. Let $S_8 := (N_8, B_8)$ such that $N_8 := \{0, a, b, c, d, e, f, g, h\}$ and B_8 is defined by the following properties :

$$\begin{aligned} & B_8^+(0, a, b), B_8^+(0, c, d), B_8^+(0, e, f), B_8^+(0, g, h), \\ & B_8^+(b, a, c, d), B_8^+(f, e, g, h), \\ & B_8^+(b, a, 0, e, f), B_8^+(d, c, 0, e, f), B_8^+(b, a, 0, g, h), B_8^+(d, c, 0, g, h). \end{aligned}$$

Figure 10(a) shows this structure drawn as Figures 3 and 5 (right part). We do not show the last four conditions for the purpose of clarity.

By adding new nodes 1 and 2 to $T(S_8, 0)$ such that $a < 1 < 0, c < 1 < 0, e < 2 < 0$ and $g < 2 < 0$, we get the rooted tree T_8 of Figure 10(b). Then $B_7 = B_{T_8}[N_7]$, hence, is in **IBQT**.

The proof of Theorem 3.1 will consist in adding new elements to trees $T(S, r)$ for such cases.

(d) If $S = (N, B)$ satisfies A1-A7 (and thus A8 by Lemma 2.7), then, for each $r \in N$, the O-tree $T(S, r)$ is a join-tree and $B = B_{T(S, r)}$ by Proposition 5.6 of [5]. \square

Definitions 3.5 : *Directions in O-trees.*

(a) Let $T = (N, \leq)$ be an O-tree¹³. Let $L \subseteq N$ be linearly ordered and upwards closed¹⁴. We denote by $N_{<}(L)$ the set $\{x \in N \mid x < L\}$. Two nodes x

¹³Or an O-forest, but we will use the notion of direction only for O-trees.

¹⁴In particular, if $X \neq \emptyset$, the set $L_{>}(X) := \{y \in N \mid y > X\}$ is linearly ordered and upwards closed.

and y in $N_{<}(L)$ are *in the same direction w.r.t. L* if $x \leq u$ and $y \leq u$ for some $u \in N_{<}(L)$. This is an equivalence relation that we denote by \sim_L . Clearly, $x \leq y$ implies $x \sim_L y$. Each equivalence class is called a *direction relative to L* . We denote by $Dir_L(x)$ the direction relative to L that contains x such that $x < L$. The O-tree is *binary* if each such L has at most two directions.

(b) Let $S = (N, B)$ satisfy A1-A6 (and not necessarily A8) and r be any node taken as root. Let $T = (N, \leq_r)$ be the O-tree $T(S, r)$. We will denote \leq_r by \leq . Related notations are $<$, \sqcup and \perp . If x and y in N are incomparable, denoted by $x \perp y$, we let $L_{>}(x, y) := L_{>}(\{x, y\})$. This set is an upwards closed line that contains r , but not x and y . We denote by \mathfrak{L} the countable set of such lines.

(c) For $L \in \mathfrak{L}$, we denote by $\mathfrak{D}(L)$ the set of directions relative to L . We have $L = L_{>}(N_{<}(L))$. \square

Examples 3.6 : (1) In the O-tree T_3 of Figure 9(a), $L_{>}(b, c) = \mathbb{N}$ and the corresponding three directions are $\{a, b\}, \{c, d\}$ and the set of negative integers.

(2) Let consider again the join-tree T (defined by Fraïssé, Example 1.2(4)). We recall that $T := (Seq_+(\mathbb{Q}), \preceq)$ partially ordered by \preceq as follows :

$$(x_n, \dots, x_0) \preceq (y_m, \dots, y_0) \text{ if and only if} \\ n \geq m, (x_{m-1}, \dots, x_0) = (y_{m-1}, \dots, y_0) \text{ and } x_m \leq y_m.$$

The join of two incomparable nodes $x := (x_n, \dots, x_0)$ and $y := (y_m, \dots, y_0)$ is $z := (x'_p, \dots, x_0)$ such that $p \leq n$, $p < m$, $(y_p, \dots, y_0) = (x'_p, \dots, x_0)$ and $x_p < x'_p$. The directions relative to $L = L_{>}(x, y) = L_{>}(z)$ are :

$$Dir_L(x) = \{(u_q, \dots, u_{p+1}, u_p, x_{p-1}, \dots, x_0) \mid u_q, \dots, u_p \in \mathbb{Q}, u_p \leq x_p\} \\ \text{and} \\ Dir_L(y) = \{(u_q, \dots, u_{p+1}, x'_p, \dots, x_0) \mid u_q, \dots, u_{p+1} \in \mathbb{Q}\}.$$

We have a structuring of T consisting of the axis $\{(x) \mid x \in \mathbb{Q}\}$ and the lines $\{(u, x_p, \dots, x_0) \mid u \in \mathbb{Q}\}$ for all $x_p, \dots, x_0 \in \mathbb{Q}$ and $p \geq 0$. \square

We will examine the directions relative to the sets $L = L_{>}(x, y)$. Clearly, $N_{<}(L)$ is the disjoint union of the directions relative to L , and there are at least two different ones, those of x and y

Lemma 3.7 : Let $S = (N, B), r$ and $<$ be as in Definition 3.5(b) and $L \in \mathfrak{L}$. Let $u, v \in D$ for some direction D relative to L , and $w < m$ with $m \in L$. Then $(u, m, w) \in B$ if and only if $(v, m, w) \in B$.

Proof : We have $\{u, v\} < a < m$ for some $a \in D$. Then $(u, a, m) \in B$ and $(v, a, m) \in B$. If $(u, m, w) \in B$, then we have $B^+(u, a, m, w)$, hence $(a, m, w) \in B$. With $(v, a, m) \in B$, we get $B^+(v, a, m, w)$, hence $(v, m, w) \in B$. \square

It follows that we can define, for $D, D' \in \mathfrak{D}(L)$ and $m \in L$:

$$B(D, m, D') : \iff B(u, m, w) \text{ for some } u \in D \text{ and } w \in D'.$$

This is actually equivalent to : $B(u, m, w)$ for all $u \in D$ and $w \in D'$.

Lemma 3.8 : Let $S = (N, B)$ satisfy A1-A6 and A8. Let $r \in N$, $T = T(S, r) := (N, \leq_r)$ and $m \in L \in \mathfrak{L}$. The binary relation $\neg B(D, m, D')$ for $D, D' \in \mathfrak{D}(L)$ is an equivalence relation.

Proof : Reflexivity and symmetry are clear. Assume that we have $\neg B(D, m, D')$ and $\neg B(D', m, D'')$ for distinct directions D, D', D'' . Hence, $\neg B(u, m, v)$ and $\neg B(v, m, w)$ for some u, v, w respectively in D, D', D'' , and for a contradiction, assume that $B(u, m, w)$ holds.

Hence, we have $\neg B(u, m, v)$ and also $\neg B(m, u, v)$ and $\neg B(m, v, u)$ because $u \perp v$. Hence we have $\neg A(m, u, v) \wedge B(v, m, w)$, and so, A8' gives $B(v, m, w)$, contradicting the hypothesis that $B(D', m, D'')$ does not hold. Hence, $\neg B(u, m, w)$ holds for all u, w respectively in D, D'' , so that $\neg B(D, m, D'')$. \square

Definition 3.9 : *Incompatible directions.*

(a) If $D, D' \in \mathfrak{D}(L)$, we define $D \approx_L D'$ if $B(D, m, D')$ holds for no $m \in L$. By Lemma 3.2(2), $B(D, m, D')$ can hold only if m is the smallest element of L . Hence, $D \approx_L D'$ holds if and only if, either L has no smallest element or $B(D, \min(L), D')$ does not hold. Hence, by Lemma 3.3, \approx_L is an equivalence relation¹⁵. We say that D and D' are incompatible.

(b) For each $D \in \mathfrak{D}(L)$, we denote by \overline{D} the union of the directions that are \approx_L -equivalent to D . The sets \overline{D} form a partition of $N_{<}(L)$. We define $\mathcal{C} := \mathcal{C}_1 \uplus \mathcal{C}_2$ as the set of downward closed subsets of N such that :

$\mathcal{C}_1 := \{N_{\leq}(x) \mid x \in N\}$ (in particular $N = N_{\leq}(r)$) and

$\mathcal{C}_2 := \{\overline{D} \mid D \in \mathfrak{D}(L), L \in \mathfrak{L} \text{ and } \overline{D} \text{ is the union of at least two directions}\}.$

Lemma 3.10 : Let S be as in Lemma 3.8.

(1) The family \mathcal{C} is not overlapping.

(2) It is first-order definable in S .

Proof : (1) Consider E and E' in \mathcal{C} such that $w \in E \cap E'$.

There are three possible cases to consider.

Case 1 : $E = N_{\leq}(x), E' = N_{\leq}(y)$. Then $x \leq y$ or $y \leq x$ because $w \leq x$ and $w \leq y$, which gives $E \subseteq E'$ or $E' \subseteq E$.

Case 2 : $E = N_{\leq}(x), w \leq x, E' = \overline{D}, D = \text{Dir}_L(w)$ where $L \in \mathfrak{L}$. Then $x < L$ (in particular if $x = w$) or $x \in L$, which gives $E \subseteq D \subseteq E'$ or $E' \subseteq E$.

Case 3 : $E = \overline{D}, D \in \mathfrak{D}(L)$, and $E' = \overline{D'}, D' \in \mathfrak{D}(L')$. Then $L \cup L' \subseteq L_{>}(w)$, hence $L' \subset L$ or $L \subset L'$ or $L = L'$. In the first case, we have $\text{Dir}_L(w) \subseteq E \subseteq N_{\leq}(x)$ for any $x \in L - L'$. We have $x < L'$. Then, $N_{\leq}(x) \subseteq \text{Dir}_{L'}(w) \subseteq E'$. The second case is similar and the last one gives $\text{Dir}_L(w) = \text{Dir}_{L'}(w)$, hence, $E = E'$.

¹⁵Not to be confused with \sim_L of Definition 3.5(a), whose classes are the directions relative to L .

(2) We recall that \mathcal{C} is relative to a rooted O-tree $T(S, r)$ that depends on a chosen $r \in N$. There exists an FO formula $\varphi(X, r)$ (not depending on S) such that for every r and $X \subseteq N$,

$$S = (N, B) \models \varphi(X, r) \text{ if and only if } X \in \mathcal{C}.$$

Since \mathcal{C} is defined from $T(S, r)$, this formula has free variable r . The partial order \leq_r (denoted by \leq) is FO definable in S in terms of r .

An FO formula $\varphi_1(X, r)$ can express that $X = N_{\leq}(x)$ for some $x \in N$.

Next we consider the sets \overline{D} . Let x and y be incomparable in $T = T(S, r) = (N, \leq)$. Let $L = L_{>}(x, y)$ and $u, v < L$. The nodes u and v are in a same set \overline{D} for some $D \in \mathfrak{D}(L)$ (actually $D = \text{Dir}_L(u)$) if and only if :

$$(N, B) \models \forall z.(z \in L \implies \neg B(u, z, v)),$$

which can be expressed by an FO formula $\sigma(r, x, y, u, v)$ because $z \in L$ is FO expressible¹⁶ in terms of r, x and y .

If $u < L$, then $\overline{\text{Dir}_L(u)}$ is the union of at least two directions in $\mathfrak{D}(L)$ if and only if :

$$(N, B) \models u < L \wedge \exists v.(v < L \wedge \neg \sigma(r, x, y, u, v))$$

which is expressed by an FO formula $\delta(r, x, y, u)$ (for convenience, this formula includes the condition $u < L$).

Let $\varphi_2(X, r)$ be the FO formula expressing that :

$$\begin{aligned} \exists x, y.[x \perp y \wedge \exists u.(u \in X \wedge \delta(r, x, y, u)) \wedge \\ \forall u.(u \in X \implies \forall v.(v \in X \iff \sigma(r, x, y, u, v)))] \end{aligned}$$

(The condition $x \perp y$ is FO expressible in terms of r). It expresses that $X = \overline{\text{Dir}_{L_{>}(x, y)}(u)}$ for some incomparable elements x, y and some $u < L_{>}(x, y)$, and that X is the union of at least two directions in $\mathfrak{D}(L_{>}(x, y))$.

Hence, the formula $\varphi_1(X, r) \vee \varphi_2(X, r)$ expresses that $X \in \mathcal{C}$. \square

We will use $F(\mathcal{C})$ (cf. Definition 1.3(a)), rather denoted by $T(\mathcal{C})$ as it is an O-tree, with root $N_{\leq}(r) = N$. We have $T \subseteq T(\mathcal{C})$, where we identify a node x of T with its image under the embedding $T \rightarrow T(\mathcal{C})$ that map x to $N_{\leq}(x)$. With this notation, we have the following obvious facts :

Lemma 3.11 : For all $x, y \in N$, $D \in \mathfrak{D}(L)$, $D' \in \mathfrak{D}(L')$ and $L, L' \in \mathfrak{L}$ we have :

- (1) $N_{\leq}(x) \subset N_{\leq}(y)$ if and only if $x < y$.
- (2) $N_{\leq}(x) \subset \overline{D}$ if and only if $x < L$ and $\overline{\text{Dir}_L(x)} = \overline{D}$,
- (3) $\overline{D} \subset N_{\leq}(x)$ if and only if $x \in L$,
- (4) $\overline{D} \subset \overline{D'}$ if and only if $L' \subset L$; if $\overline{D} \subset \overline{D'}$, we have $\overline{D} \subseteq N_{\leq}(x) \subseteq \overline{D'}$ for some x in $L - L'$.

In the next three lemmas, S and the related objects are as in the previous Lemmas of this section.

¹⁶This is a key point of the proof. In the proof of Theorem 3.25, we will use an alternative description of sets L in \mathfrak{L} in which membership is still FO expressible.

Lemma 3.12: $T(\mathcal{C})$ is a join-tree.

Proof: Let E and E' be incomparable elements in $T(\mathcal{C})$. We will prove they have a join $E \sqcup_{T(\mathcal{C})} E'$ in $T(\mathcal{C})$. These sets are disjoint. There are three cases and several subcases.

Case 1 : $E = N_{\leq}(x), E' = N_{\leq}(y)$ where $x \perp y$.

Subcase 1.1 : $(x, m, y) \notin B$ for all m in $L := L(x, y)$. Then $Dir_L(x) \approx_L Dir_L(y)$ and $E'' := \overline{Dir_L(x)} \supseteq E \uplus E'$. We have $\overline{Dir_L(x)} \in \mathcal{C}$ because $Dir_L(x) \neq Dir_L(y)$.

We prove by contradiction that $E'' = E \sqcup_{T(\mathcal{C})} E'$. If this is not the case, we may have $E'' \supset N_{\leq}(z) \supseteq E \uplus E'$. But then $x, y < z$, hence $z \in L$ and $N_{\leq}(z) \supseteq N_{\leq}(L)$. So we cannot have $N_{\leq}(z) \subset E'' \subseteq N_{\leq}(L)$.

If $E'' \supset \overline{D'} \supseteq E \uplus E'$ then $\overline{D'} = \overline{Dir_{L'}(x)} = \overline{Dir_{L'}(y)}$ where $L \subset L'$. Let $z \in L' - L$. Then $x, y < z$, hence $z \in L$, contradicting the choice of z .

Note that E'' is not of the form $N_{\leq}(z)$ for any z because it is the disjoint union of at least two directions. If $E'' = N_{\leq}(z)$, then z would belong to one direction, say D'' , and all these directions, in particular $Dir_L(x)$ and $Dir_L(y)$, would be included in D'' hence equal to D'' because directions do not overlap.

Subcase 1.2 : $(x, m, y) \in B$ where $m = x \sqcup_T y = \min(L)$. Let $E'' := N_{\leq}(m) \supset E \uplus E'$.

We prove by contradiction that $E'' = E \sqcup_{T(\mathcal{C})} E'$. If this is not the case, we might have $E'' = N_{\leq}(m) \supset N_{\leq}(z) \supseteq E \uplus E'$. But then $\{x, y\} < z < m$, hence m is not the join of x and y .

If $E'' = N_{\leq}(m) \supset \overline{D'} \supseteq E \uplus E'$ then $\overline{D'} = \overline{Dir_{L'}(x)} = \overline{Dir_{L'}(y)}$ where $L \subset L'$. Let $z \in L' - L$. Then $\{x, y\} < z < m$, hence m is not the join of x and y .

Case 2 : $E = N_{\leq}(x), E' = \overline{Dir_L(y)}$. Since $N_{\leq}(x) \cap \overline{Dir_L(y)} = \emptyset$, we do not have $Dir_L(y) \approx_L Dir_L(y)$, hence we have $(x, m, y) \in B$ for some m that must be $x \sqcup_T y = \min(L)$. We claim that $N_{\leq}(m) = E \sqcup_{T(\mathcal{C})} E'$. The proof by contradiction is as in Subcase 1.2.

Case 3 : $E = \overline{D}, D \in \mathfrak{D}(L)$, and $E' = \overline{D'}, D' \in \mathfrak{D}(L')$. If $L = L'$ then, as $\overline{D} \neq \overline{D'}$, we have $B(D, m, D')$ with $m = \min(L) = x \sqcup_T y \in L$, and then $E \sqcup_{T(\mathcal{C})} E' = N_{\leq}(m)$, as in Case 2. Otherwise, if $L \subset L'$, let $y \in L' - L$, $x \in \overline{D}$, and $(x, m, y) \in B$ for some $m \in L$. Hence, $m = \min(L)$ and $N_{\leq}(m) = E \sqcup_{T(\mathcal{C})} E'$. \square

Lemma 3.13 : $B \subseteq B_{T(\mathcal{C})}[N]$.

Proof : We recall that $<$ denotes $<_r = <_{T(S, r)}$ that is, by Fact (1) of Lemma 3.11, the restriction of $<_{T(\mathcal{C})}$ to N . The joins in $T(S, r)$ and $T(\mathcal{C})$ are not always the same.

Consider $(x, y, z) \in B$. By Lemma 3.3(2), we have $x < y$ or $z < y$. Assume the first.

If $y < z$ then $x <_{T(\mathcal{C})} y <_{T(\mathcal{C})} z$, hence $(x, y, z) \in B_{T(\mathcal{C})}[N]$.

If $z < y$, then $y = x \sqcup_{T(S, r)} z$, by Lemma 3.2(2). We are in Subcase 1.2 of Lemma 3.12, hence, $y = x \sqcup_{T(\mathcal{C})} z$ and $(x, y, z) \in B_{T(\mathcal{C})}$.

If $y \perp z$, then, let $E := y \sqcup_{T(\mathcal{C})} z$. We have $x < y <_{T(\mathcal{C})} E$, hence $(x, y, E) \in B_{T(\mathcal{C})}$, and also $(y, E, z) \in B_{T(\mathcal{C})}$, hence $(x, y, z) \in B_{T(\mathcal{C})}$.

The case $z < y$ is similar. \square

Lemma 3.14 : $B_{T(\mathcal{C})}[N] \subseteq B$.

Proof : Let $x, y, z \in N$ such that $(x, y, z) \in B_{T(\mathcal{C})}$.

If we have $x < y < z$ or $z < y < x$, then $(x, y, z) \in B$ by the definition of $<$ as $<_{T(S,r)}$.

Otherwise $x \perp z$ and, $x < y \leq_{T(\mathcal{C})} u >_{T(\mathcal{C})} z$ where $u = y \sqcup_{T(\mathcal{C})} z = x \sqcup_{T(\mathcal{C})} z$ or, similarly, $x <_{T(\mathcal{C})} u \geq_{T(\mathcal{C})} y > z$. We assume the first.

Case 1 : $y \perp z$. Then we have $(x, y, z) \in B$ by Lemma 3.3(1).

Case 2 : If y and z comparable, we must have $y > z$. As $x < y \leq_{T(\mathcal{C})} u = x \sqcup_{T(\mathcal{C})} z$, we must have $y = u$. This means that we cannot be in Subcase 1.1 of Lemma 3.12 (for the definition of $x \sqcup_{T(\mathcal{C})} z$); hence we are in Subcase 1.2 with $y = x \sqcup_{T(S,r)} z$ and $(x, y, z) \in B$.

This completes the proof. \square

Proof of Theorem 3.1 : From (N, B) satisfying A1-A6 and A8, we have built a join-tree $T(\mathcal{C})$ whose nodes \mathcal{C} contains N (with x identified with $N_{\leq}(x)$) such that, by Lemmas 3.13 and 3.14, the restriction of its betweenness relation to N is B . Hence, together with Proposition 2.9, a structure (N, B) is in **IBQT** if and only if it satisfies A1-A6 and A8. \square

We know from Definition 10 and Proposition 17 of [4] that a quasi-tree (N, B) is the betweenness relation of a tree if and only if B is *discrete*, i.e., that each set $[x, y]_B := \{x, y\} \cup \{z \in N \mid B(x, z, y)\}$ is finite (cf. Definition 2.8(a)).

Corollary 3.15: A nontrivial structure (N, B) is an induced betweenness relation in a tree if and only if it satisfies axioms A1-A6, A8 and is discrete. These conditions are monadic second-order expressible.

Proof: An induced relation (N, B) of a discrete one is discrete, which gives the only if directions by Proposition 2.9.

If $S = (N, B)$ satisfies axioms A1-A6, A8 and is discrete, then for all $x, y \in N$ such that $x <_{T(S,r)} y$, the set $\{z \in N \mid x <_{T(S,r)} z <_{T(S,r)} y\}$ is finite. Hence, $T(S, r)$ is a rooted tree.

From Lemma 3.11(4), we get that, for all $x, y \in N_{T(\mathcal{C})}$ such that $x <_{T(\mathcal{C})} y$, the set $\{z \in N_{T(\mathcal{C})} \mid x <_{T(\mathcal{C})} z <_{T(\mathcal{C})} y\}$ is finite. Hence, $T(\mathcal{C})$ is a rooted tree.

The property that an interval of a linear order is finite is monadic second-order expressible as recalled in Section 1.4.

Examples and remarks 3.16 : *About the proof of Theorem 3.1.*

(1) Consider the structure S_8 of Figure 10(a). The O-tree $T(S_8, 0)$ is T_8 (in Figure 10(b)) minus the nodes 1 and 2. There are four directions relative to $L := \{0\} = L_{>}(a, c) : D(a)$, the direction of \underline{a} , and similarly, $D(c), D(e)$ and $D(g)$. The two equivalence classes of $\mathfrak{D}(L)$ are $\underline{D(a)} = D(a) \uplus D(c) = \{a, b, c, d\}$ and $\underline{D(e)} = D(e) \uplus D(g) = \{e, f, g, h\}$. The nodes 1 and 2 of Figure 10(b) represent the two nodes $\underline{D(a)}$ and $\underline{D(e)}$ added to $T(S_8, 0)$ to form the tree T_8 such that $S_8 = B_{T_8}$.

(2) Consider the O-tree of Figure 8 and its betweenness relation to which we add the fact $B^+(a, b, c)$. Let $L := \mathbb{N}$. The two directions relative to L are $\{a, b\}$ and $\{c\}$. They are \approx_L -equivalent. Only one node is added : $\{a, b, c\} = D(a) \uplus D(c)$.

(3) Let $T = (N, \leq)$ be a join-tree with root r . Let $S := (N, B_T)$. Then, $T = T(S, r)$. We now apply the construction of Theorem 3.1. Each $L \in \mathcal{L}$ has a minimal element, because T is a join-tree. It follows that no two different directions relative to L are equivalent with respect to \approx_L . Hence, The family \mathcal{C} consists only of the sets $N_{\leq}(x)$ and so, $T(\mathcal{C}) = T(S, r) = T$.

(4) If $S = (N, B)$ is an induced betweenness in a quasi-tree, then any node r can be taken as root for defining an O-tree $T(S, r)$ and from it, a join-tree $T(\mathcal{C})$. This fact generalizes the observation that the betweenness in a tree T does not depend on any root. Informally, quasi-trees and induced betweenness in quasi-trees are "undirected notions". This will not be true for betweenness in O-trees. See the remark about U in the proof of Proposition 2.15, Part (2). \square

3.1.2 Betweenness in rooted O-trees

We let $\mathbf{BO}_{\text{root}}$ be the class of betweenness relations of rooted O-trees. These relations satisfy A1-A6.

Proposition 3.17 : The class $\mathbf{BO}_{\text{root}}$ is axiomatized by a first-order sentence.

Proof: Consider $S = (N, B)$. If B is the betweenness relation of an O-tree (N, \leq) with root r , then, \leq is nothing but \leq_r defined in Lemma 3.2 from B and r . Let φ be the FO sentence that expresses properties A1-A6 (relative to B) and the following one :

A9 : there exists $r \in N$ such that the O-tree $T(S, r) = (N, \leq_r)$ whose partial order is defined by $x \leq_r y :\iff x = y \vee y = r \vee B(x, y, r)$ has a betweenness relation $B_{T(S, r)}$ equal to B .

That S satisfies A1-A6 insures that (N, \leq_r) is an O-tree with root r . The sentence φ holds if and only if S is in $\mathbf{BO}_{\text{root}}$. When it holds, the found node r defines via \leq_r the relevant O-tree. \square

The following counter-example shows that we do not obtain an FO axiomatization of the class \mathbf{BO} .

Example 3.18 : $\mathbf{BO}_{\text{root}}$ is properly included in \mathbf{BO} .

Let T be the O-tree with set of nodes \mathbb{Q} and defining partial order \preceq such that $x \preceq y :\iff x \leq y \wedge y \in \mathbb{Q} - \mathbb{Z}$ (see Figure 11). Any two elements of \mathbb{Z} are incomparable and no two incomparable elements have a join. We claim that B_T is not in $\mathbf{BO}_{\text{root}}$.

Assume that $B_T = B_U$ for some O-tree U with root $r \in \mathbb{Q}$. We will derive a contradiction.

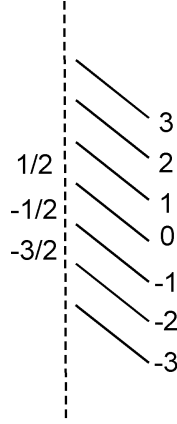


Figure 11: The O-tree of Example 3.18.

If $r \in \mathbb{Z}$ we take, without loss of generality, $r = 0$. Let $a = -1/2$ and $b = -3/2$. These nodes are incomparable in U otherwise, we would have $(0, a, b)$ or $(0, b, a)$ in $B_U = B_T$ which is false. Hence $(a, 0, b) \in B_U$, but $(a, 0, b) \notin B_T$.

If $r \in \mathbb{Q} - \mathbb{Z}$ we take, without loss of generality, $r = 1/2$. Let $a = 1$ and $b = 2$. These nodes are incomparable in U otherwise, we would have $(1/2, a, b)$ or $(1/2, b, a)$ in $B_U = B_T$ which is false. Hence $(a, 1/2, b) \in B_U$, but $(a, 1/2, b) \notin B_T$.

3.2 Monadic second-order axiomatizations

3.2.1 Betweenness in O-trees.

We will prove that the class **BO** is axiomatized by a monadic second-order sentence. In the proof of Proposition 3.17, we have defined from $S = (N, B)$ satisfying A1-A6 and $r \in N$ a candidate partial order \leq_r for (N, \leq_r) to be an O-tree with root r whose betweenness relation would be B . The order \leq_r being expressible by a first-order sentence, we finally obtained a first-order characterization of **BO**_{root}. For **BO**, a candidate order will be defined from a *line*, not from a single node. It follows that we will need for our construction a set quantification.

The next lemma is Proposition 5.3 of [5].

Lemma 3.19 : Let (L, B) satisfy properties A1-A7'. Let a, b be distinct elements of L . There exist a unique linear order \leq on L such that $a < b$ and $B_{(L, \leq)} = B$. This order is quantifier-free definable in the logical structure (L, B) in terms of a and b .

We will denote this order by $\leq_{L,B,a,b}$. There is a quantifier-free formula λ , written with the ternary relation symbol B , such that, for all a, b, u, v in L , $(L, B) \models \lambda(a, b, u, v)$ if and only if $u \leq_{L,B,a,b} v$. We recall from Definition 1.1 that a *line* L in an O-tree T is a linearly ordered set that is convex : $x \leq_T y \leq_T z$ and $x, z \in L$ imply $y \in L$.

Lemma 3.20 : Let $T = (N, \leq_T)$ be an O-tree, L a maximal line in T that has no largest node. Let $a, b \in L$, such that $a <_L b$, where $<_L$ is the restriction of $<_T$ to L .

- (1) The partial order \leq_T is first-order definable in a unique way in the structure (N, B_T) in terms of L, \leq_L, a and b .
- (2) It is first-order definable in (N, B_T) in terms of L, a and b . \square

Maximality of L is for set inclusion. This condition implies that L is upwards closed, and furthermore, infinite.

Proof: Let $x, y \in N$. We first prove the following facts.

Fact 1 : If $x, y \in L$, then $x <_T y$ if and only if $x <_L y$.

Fact 2 : If $x \notin L, y \in L$, then $x <_T y$ if and only if $(x, y, z) \in B_T$ for some $z \in L$ such that $z >_L y$.

Fact 3 : If $x, y \notin L$, then $x <_T y$ if and only if $B_T^+(x, y, z, u)$ holds for some z, u in L , such that $u >_L z$.

Fact 1 is clear from the definitions.

For Fact 2, we have some $z >_L y$ because L has no largest element. If $x <_T y <_L z$, then $(x, y, z) \in B_T$.

Assume now that $(x, y, z) \in B_T$ for some $z >_L y$. By the definition of B_T , we have $x <_T y \leq_T x \sqcup_T z$ or $z <_T y \leq_T x \sqcup_T z$. Since $z >_L y$, we cannot have $z <_T y$. Hence, $x <_T y$. (We have actually $(x, y, z) \in B_T$ for every $z >_L y$).

For Fact 3, we note that for every $y \notin L$, we have some $z \in L, z >_T y$: take for z any upper-bound of y and some element of L , then $z \in L$ because T is an O-tree. Hence, we have $z, u \in L$ such that $y <_T z <_L u$ because L has no largest element, hence $(y, z, u) \in B_T$ by Fact 2.

If $x <_T y$, we have $x <_T y <_T z$ hence, $(x, y, z) \in B_T$ and $B_T^+(x, y, z, u)$ holds.

Assume now for the converse that $B_T^+(x, y, z, u)$ holds for $z, u \in L$ such that $z <_L u$. We have $(x, y, z) \in B_T$ and $z >_T y$ by Fact 2. By the definition of B_T , we have $x <_T y \leq x \sqcup_T z$ or $z <_T y \leq x \sqcup_T z$. Since $z >_T y$, we cannot have $z <_T y$, hence, $x <_T y$.

We now prove the two assertions of the statement.

(1) The above four facts show that \leq_T is first-order definable in (N, B_T) in terms of L, \leq_L, a and b . More precisely, Facts 1,2 and 3 can be expressed as a first-order formula θ written with the relation symbols L, B and R of respective arities 1,3 and 2, such that, if L is a maximal line in T that has no largest node, $a, b \in L$ and $a <_L b$, then, for all $u, v \in N$, $(N, L, B_T, \leq_L) \models \theta(a, b, u, v)$ if and

only if $u \leq_T v$. For the validity of $\theta(a, b, u, v)$, B_T is the value of B , and \leq_L is that of R .

(2) However, \leq_L is FO definable in $(L, B_T[L])$ by Lemma 3.20. By replacing the atomic formulas $R(x, y)$ by $\lambda(a, b, x, y)$, we ensure that R is \leq_L , hence, we obtain a first-order formula $\psi(a, b, u, v)$, written with L and B such that, for $u, v \in N$ we have $(N, B_T) \models \psi(a, b, u, v)$ if and only if $u <_T v$ where B_T is the value of B . \square

A *line* in a structure $S = (N, B)$ that satisfies A1-A6 is a set $L \subseteq N$ of at least 3 elements in which any 3 different elements are aligned (cf. Definition 2.1(c)) and that is convex, *i.e.*, $[x, y]_B \subseteq L$ for all x, y in L .

Theorem 3.21 : The class **BO** is axiomatized by a monadic second-order sentence.

Proof : Let $\varphi(L)$ be the monadic second-order formula expressing the following properties of a structure $S = (N, B)$ and a set $L \subseteq N$:

- (i) S satisfies A1-A6,
- (ii) L is a maximal line in S ,
- (iii) there are $a, b \in L$ such that the formula $\psi(a, b, u, v)$ of Lemma 3.20 defines a partial order \leq on N such that $a < b$,
- (iv) (N, \leq) is an O-tree U , in which L is a maximal line without largest element, and
- (v) $B_U = B$.

We need a set quantification to express the maximality of L . All other conditions are first-order expressible.

If $S = (N, B_T)$ is the betweenness relation of an O-tree $T = (N, \leq)$ without root, and L is a maximal line in T , then L is also a maximal line in S . As T has no root, L has no largest element. Then $\varphi(L)$ holds where $a, b \in L$ are such that $a <_L b$. Hence, $S \models \exists L. \varphi(L)$.

Conversely, if $S = (N, B)$ satisfies $\exists L. \varphi(L)$, then, conditions (iv) and (v) show that S is in the class **BO**.

Together with Proposition 3.17, we can express by an MSO sentence that (S, N) is the betweenness relation of an O-tree, with or without root.

A structure $S = (N, B)$ is the betweenness relation of an O-forest if and only if its connected components (cf. Remark 2.14) that are the betweenness relations of O-trees. Hence, we get a monadic second-order sentence expressing that a structure S is the betweenness relation of an O-forest. \square

3.2.2 Induced betweenness in O-trees.

Next we examine in a similar way the class **IBO**. It is easy to see that **IBO** = **IBO_{root}**.

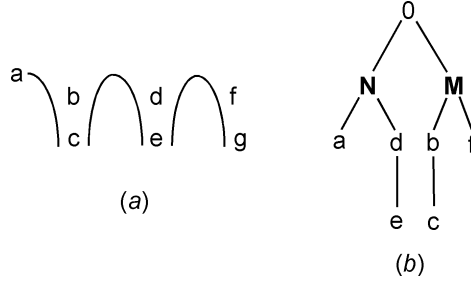


Figure 12: The structure U of Proposition 3.22 (the counter-example) and the O-tree T of Remark 3.23.

Proposition 3.22 : Every structure in the class **IBO** satisfies Properties A1-A6 but these properties do not characterize this class.

Proof: Every structure S in the class **IBO** is an induced substructure of some S' in **BO**, that thus satisfies Properties A1-A6. Hence, S satisfies also these properties as they are expressed by universal sentences.

Now, we give an example of a structure $U = (N, B)$ that satisfies Properties A1-A6 but is not in **IBO**_{root}.

We let $N := \{a, b, c, d, e, f, g\}$ and B such that $B^+(a, b, c)$, $B^+(c, b, d, e)$, $B^+(e, d, f, g)$ hold, and nothing else. See Figure 12(a), with the conventions of Figures 3 and 5. Assume that $B = B_T[N]$ where T is an O-tree (M, \leq) such that $N \subseteq M$. We will consider several cases leading each to $B \subset B_T[N]$, hence to a contradiction.

(1) We first assume that a, c, e, g are pairwise incomparable.

The joins $a \sqcup c$, $c \sqcup e$ and $e \sqcup g$ must be defined (because (a, b, c) , (c, b, e) and (e, f, g) are in B_T) and furthermore $b \leq a \sqcup c$, $b \leq c \sqcup e$, $d \leq c \sqcup e$, $d \leq e \sqcup g$ and $f \leq e \sqcup g$. The joins $a \sqcup c$ and $c \sqcup e$ must be comparable and so must be $c \sqcup e$ and $e \sqcup g$.

(1.1) These joins are pairwise distinct, otherwise $B_T[N]$ contains triples not in B , as we now prove.

(1.1.1) Assume $a \sqcup c = c \sqcup e = e \sqcup g = \alpha$. At least one of $a \sqcup e$, $c \sqcup g$ and $a \sqcup g$ is defined and equal to α .

If $a \sqcup e = \alpha = a \sqcup c = c \sqcup e$, then either $c < d \leq \alpha$ or $e < d \leq \alpha$ because $(c, d, e) \in B_T$. Hence, we have (a, d, c) or (a, d, e) in $B_T[N]$ but these triples do not belong to B . All other proofs will be of this type.

If $c \sqcup g = \alpha = c \sqcup e = e \sqcup g$, then (c, f, e) or (c, f, g) is in $B_T[N] - B$ if, respectively, $e < f \leq \alpha$ or $g < f \leq \alpha$ (because $(e, f, g) \in B_T$).

If $a \sqcup g = \alpha = c \sqcup e = e \sqcup g$, then (a, f, g) or (c, f, e) is in $B_T[N] - B$, if, respectively, $g < f \leq \alpha$ or $e < f \leq \alpha$ (because $(e, f, g) \in B_T$).

(1.1.2) We now consider the cases where only two of $a \sqcup c$, $c \sqcup e$ and $e \sqcup g$ are equal.

If $a \sqcup c = c \sqcup e = \alpha$, then if $\alpha < e \sqcup g$, then (a, b, g) or (c, b, g) is in $B_T[N] - B$ (because $(a, b, c) \in B_T$); if $e \sqcup g < \alpha$, then (c, f, e) or (c, f, g) is in $B_T[N] - B$ because $\alpha = c \sqcup e = c \sqcup g$.

If $c \sqcup e = e \sqcup g = \alpha$ and $a \sqcup c < \alpha$, then $e < d \leq \alpha$ or $c < d \leq \alpha$ which gives (a, d, e) or (a, d, c) in $B_T[N] - B$; if $\alpha < a \sqcup c$, then (a, f, g) or (a, f, e) is in $B_T[N] - B$.

If $a \sqcup c = e \sqcup g = \alpha$, then we have $c \sqcup e < \alpha$ and $a \sqcup e = \alpha$. Hence, (a, d, c) or (a, d, e) is in $B_T[N] - B$. We cannot have $\alpha < c \sqcup e$ because then $c, e < \alpha < c \sqcup e$.

(1.2) If $a \sqcup c$ and $e \sqcup g$ are incomparable, then $a \sqcup c < c \sqcup e$ and $e \sqcup g < c \sqcup e$. We have then $c \sqcup e = c \sqcup g = a \sqcup g$. Hence, we get that (a, b, g) or (c, b, g) is in $B_T[N] - B$.

(1.3) Hence, $a \sqcup c$, $c \sqcup e$ and $e \sqcup g$ are pairwise different but comparable. We have six cases to consider : $a \sqcup c < c \sqcup e < e \sqcup g$ and five other ones, corresponding to the six sequences of three objects.

If $a \sqcup c < c \sqcup e < e \sqcup g$ then, $a < b < a \sqcup c$ or $c < b < a \sqcup c$ and (a, b, g) or $(c, b, g) \in B_T[N] - B$.

The verifications are similar in the five other cases.

(2) We consider cases where a, c, e, g are not pairwise incomparable.

Observation : If $u < x, (x, y, z) \in B_T$ and we do not have $x > z$, then $B_T^+(u, x, y, z)$ holds. (If $x > z$, then x may not be the join of u and z).

If $a > c$, then we have $a > b > c$ and $c \sqcup e > c$. Hence $c \sqcup e \geq b$, or $b > c \sqcup e$. We get triples (e, b, c) or (a, b, e) in $B_T[N] - B$.

If $a < c$, then we have $a < b < c \leq c \sqcup e$. Hence $(a, c, e) \in B_T[N] - B$.

Hence $a \perp c$. By the observation, we cannot have $e < c, g < c, e < a$ or $g < a$.

If $c < e$, then, if $a \sqcup c \leq e$ we have (e, b, c) or (e, b, a) in $B_T[N] - B$; if $e < a \sqcup c$, then $(a, e, c) \in B_T[N] - B$.

Hence, $c \perp e$. By the observation, we cannot have $a < c, a < e$, or $g < e$.

If $e < g$, then, either $c \sqcup e \leq g$ or $g < c \sqcup e$ which gives $(g, b, c), (g, b, e)$ or (c, g, e) in $B_T[N] - B$.

Hence, $e \perp g$. By the observation, we cannot have $a < g$ or $c < g$.

All cases yield $B \subset B_T[N]$. Hence, S is not in **IBO**. \square

Remarks 3.23 : (1) If we modify U of the previous proof by replacing $B^+(c, b, d, e)$ by $B^+(c, d, e)$ (but we keep b in the set of nodes), we get a modified structure U' for which the same result holds, by a similar proof.

(2) If we delete g from U , we get a structure W that is in **IBO_{root}**. A witnessing O-tree T is shown in Figure 12(b) where **N** and **M** represent two copies of \mathbb{N} ordered top-down as in the O-tree T_2 of Figure 8 and the proof of Proposition 2.15.

(3) For every finite structure $H = (N_H, B_H)$, let φ_H be a first-order sentence expressing that a given structure (N, B) has an induced substructure isomorphic to H . Hence, every structure in **IBO** satisfies properties A1-A6 and $\neg\varphi_U \wedge \neg\varphi_{U'}$.

We do not know whether this first-order sentence axiomatizes the class **IBO**, and more generally, whether there exists a finite set of "excluded" finite induced

structures like U and U' , that would characterize the class **IBO**. The existence of such a set would give a first-order axiomatization of **IBO**.

The construction of Theorem 3.21 does not extend to **IBO** because, as we noted in the proof of Proposition 2.15 (point (3)), a *finite* structure in **IBO** may not be an induced betweenness relation of any finite O-tree. No construction like that of $T(\mathcal{C})$ in the proof of Theorem 3.1 can produce an infinite structure from a finite one. Nevertheless :

Conjecture 3.24 : The class **IBO** is characterized by a monadic second-order sentence.

3.3 Logically defined transformations of structures

Each betweenness relation is a structure $S = (N, B)$ defined from a *marked O-tree*, a structure $T = (P, \leq, N)$ where (P, \leq) is an O-tree and $N \subseteq P$, the set of *marked* nodes, is handled as a unary relation. The different cases are shown in Table 1. In each case a first-order formula can check whether the structure (P, \leq, N) is of the appropriate type, and another one can define the relation B in (P, \leq, N) . Hence, the transformation of (P, \leq, N) into (N, B) is a first-order transduction (Definition 1.6).

<i>Structure</i> (N, B)	<i>Axiomatization</i>	<i>Source structure</i>	<i>From (N, B) to a source structure</i>
QT	FO : A1-A7, Prop. 2.9	join-tree (N, \leq, N)	FOT
IBQT	FO : A1-A6, A8, Thm 3.1	join-tree (P, \leq, N)	MSOT
BO	MSO : Theorem 3.21	O-tree (N, \leq, N)	MSOT
IBO	MSO ? : Conjecture 3.24	O-tree (P, \leq, N)	not MSOT

Table 1

The last column indicates which type of transduction, FO transduction (*FOT*) or MSO transduction (*MSOT*) can produce, from a structure (N, B) , a relevant marked O-tree (P, \leq, N) . For **QT**, this follows from the proof of Theorem 2.9(1) : if $S = (N, B)$ satisfies A1-A7 and $r \in N$, then, the O-tree $T(S, r) = (N, \leq_r)$ is a join-tree and $B = B_{T(S, r)}$. For **BO**, the MSO sentence that axiomatizes the class constructs a relevant O-tree (it guesses one and checks that the guess is correct). For **IBO**, we observed that the source tree may need to be infinite for defining a finite betweenness structure, which excludes the existence of an MSO transduction, because these transformations produce structures whose domain size is linear in that of the input structure. (cf. Definition 1.6, and Chapter 7 of [8]).

It remains to prove that the transformation of $S \in \mathbf{IBQT}$ into a witnessing marked O-tree (P, \leq, N) is a monadic second-order transduction. This is the content of the following statement.

Theorem 3.25 : A marked join-tree witnessing that a given structure S is in **IBQT** can be defined from S by MSO formulas. \square

We first describe the proof strategy. We want to prove that, for a given structure $S = (N, B)$ that satisfies Axioms A1-A6 and A8, the tree $T(\mathcal{C})$ of the proof of Theorem 3.1 can be constructed by MSO formulas (of course independent of S).

The first step is the construction of $T(S, r) = (N, \leq_r)$: one chooses a node r from which the partial order \leq_r is FO definable in S by using r as value of a variable.

The nodes of $T(\mathcal{C})$ (constructed from $T(S, r)$) are the sets in \mathcal{C} (cf. the proof of Theorem 3.1) and they are of two types :

either $N_{\leq}(z)$, they are in \mathcal{C}_1 ,

or $\overline{Dir_L(u)}$ for $u < L$ and $L \in \mathcal{L}$ such that $\overline{Dir_L(u)}$ is the union of at least two directions (cf. Lemma 3.10); they are in \mathcal{C}_2 .

A set $N_{\leq}(z)$ is represented by its maximal element z in a natural way, and T embeds in $T(\mathcal{C})$ (cf. Definition 1.3(a)).

A set $\overline{Dir_L(u)}$ is a new node added to T . In order to make the transformation of $S \mapsto T(\mathcal{C})$ into a transduction as in Definition 1.6(b), we define $N_{T(\mathcal{C})}$ as $(N \times \{1\}) \uplus (M \times \{2\})$ where $(x, 1)$ encodes $N_{\leq}(x)$ and each $w \in M \subseteq N$ encodes (bijectively) some set $\overline{Dir_L(u)} \in \mathcal{C}_2$. An MSO formula will express that a node z encodes $U = \overline{Dir_L(u)}$, for some L and u .

Lemma 3.10(2) has shown that each set $\overline{Dir_L(u)}$ in \mathcal{C}_2 can be defined from three nodes x, y and u . We need a definition by a single node, in order to obtain a monadic second-order transduction. The sets U in \mathcal{C}_2 are FO definable but not pairwise disjoint. Hence, one cannot select arbitrarily an element of U to represent it. We will use a notion of structuring of O-trees, similar to the one defined in [5] for join-trees, that we will also use in Section 4. We will also have to prove that the partial order $\leq_{T(\mathcal{C})}$ is defined by MSO formulas, but this will be straightforward by means of the formula expressing that a node z encodes a set in \mathcal{C}_2 .

Definition 3.26: *Strict upper-bounds.*

Let (N, \leq) be a partial order and $X \subseteq N$. A *strict upper-bound* of X is an element y such that $y > X$. We denote by $lsub(X)$ the *least strict upper-bound* of X if it exists. If X has no maximum element but has a least upper-bound m , then $lsub(X) = m$. If X has a maximum element m , its least strict upper-bound if it does exist *covers* m , that is, $lsub(X) > m$ and there is no x such that $lsub(X) > x > m$.

Definition 3.27 : *Structurings of O-trees.*

In the following definitions, $T = (N, \leq)$ is an O-tree.

(a) If U and W are two lines (convex linearly ordered subsets of N), we say that W *covers* U , denoted by $U \prec W$, if $U < w$ for some w in W and, for such w and any $x \in N$, if $U < x < w$, then $x \in W$. (See Example 3.28(1) below). Note that $lsub(U)$ may not exist, but if it does, it is in W .

(b) A *structuring* of T is a set \mathcal{U} of nonempty lines that forms a partition of N and satisfies the following conditions:

1) One distinguished line called the *axis* is upwards closed.

- 2) There are no two lines $U, U' \in \mathcal{U}$ such that $U < U'$.
3) For each x in N , $L_{\geq}(x) = I_k \uplus I_{k-1} \uplus \dots \uplus I_0$ for nonempty intervals I_0, \dots, I_k of $(L_{\geq}(x), \leq)$ such that:

- 3.1) $x = \min(I_k)$ and $I_k < I_{k-1} < \dots < I_0$,
3.2) for each j , there is a line $U \in \mathcal{U}$ such that $I_j \subseteq U$, and it is denoted by U_j ; U_0 is the axis,
3.3) each I_j is upwards closed in U_j , that is, if $x \in U_j$ and $x > y \in I_j$ then $x \in I_j$.

Hence, $U_j \neq U_{j'}$ if $j \neq j'$, and $U_j < U_{j-1}$ for $j = 1, \dots, k$. The sequence I_0, I_1, \dots, I_k is unique for each x , and k is called the *depth* of x and also of U_k . We denote by $U(x)$ the line that contains $x \in N$.

We say that $T = (N, \leq, \mathcal{U})$ is a *structured O-tree*. \square

Examples 3.28 : (1) Let $T = (N, \leq)$ be the O-tree such that $N = \{a, b\} \times \mathbb{Q}$ ordered such that $(a, x) \leq (a, y)$ and $(b, x) \leq (b, y)$ if and only if $x \leq_{\mathbb{Q}} y$, and $(b, x) \leq (a, y)$ if and only if $y >_{\mathbb{Q}} \alpha$ where α is an irrational real number. Then, the lines $W := \{a\} \times \mathbb{Q}$ and $U := \{b\} \times \mathbb{Q}$ form a structuring of T with axis W , and $U < W$.

(2) A structuring of the tree T_3 of Figure 9(a) consists of the axis $\{a, b\} \cup \mathbb{N}$ and two lines that are $\{c, d\}$ and the set of negative integers.

(3) Let T be the join-tree of Example 1.2(4). We recall that $T := (Seq_+(\mathbb{Q}), \leq)$, where $Seq_+(\mathbb{Q})$ is the set of finite nonempty sequences of rational numbers partially ordered as follows: $(x_n, \dots, x_0) \leq (y_m, \dots, y_0)$ if and only if $n \geq m$, $(x_{m-1}, \dots, x_0) = (y_{m-1}, \dots, y_0)$ and $x_m \leq_{\mathbb{Q}} y_m$. It has a structuring consisting of $\{(x_0) \mid x_0 \in \mathbb{Q}\}$ as an axis and the lines $\{(x_n, \dots, x_0) \mid x_n \in \mathbb{Q}\}$, for each $n \geq 1$ and $x_0, \dots, x_{n-1} \in \mathbb{Q}$. A node (x_n, \dots, x_0) is at depth n .

(4) Figure 13 shows a structuring of a join-tree with axis U_0 and lines U_0, \dots, U_6 such that $U_1 < U_0, U_3 < U_2 < U_0, U_6 < U_2$ and $U_5 < U_4 < U_0$. We have $L_{\geq}(i) = I_2 \uplus I_1 \uplus I_0$ where $I_2 = U_3 \cap L_{\geq}(i), I_1 = U_2 \cap L_{\geq}(g), I_0 = U_0 \cap L_{\geq}(e)$. \square

Proposition 3.29 : Let \mathcal{U} be a structuring of an O-tree $T = (N, \leq)$. Then, T is a join-tree if and only if each $U \in \mathcal{U}$ that is not the axis has a least strict upper-bound, and $lsub(U) \in W$ where W is the line in \mathcal{U} that covers U .

Proof : Clear from Definition 3.27. \square

Proposition 3.30 : Every O-tree has a structuring.

Proof : The proof is similar to that of [5] establishing that every join-tree has a structuring. We give it for completeness. Let $T = (N, \leq)$ be an O-tree. We choose an enumeration $x_0, x_1, \dots, x_n, \dots$ of N and a maximal line B_0 ; it is thus upwards closed.

For each $i > 0$, we choose a maximal line B_i containing the first node not in $B_{i-1} \cup \dots \cup B_0$. We define $U_0 := B_0$ and, for $i > 0$, $U_i := B_i - (U_{i-1} \uplus \dots \uplus U_0) = B_i - (B_{i-1} \cup \dots \cup B_0)$. We define \mathcal{U} as the set of lines U_i . It is a structuring of J . The axis is U_0 . Condition 2) is guaranteed because we choose a maximal line B_i at each step. \square

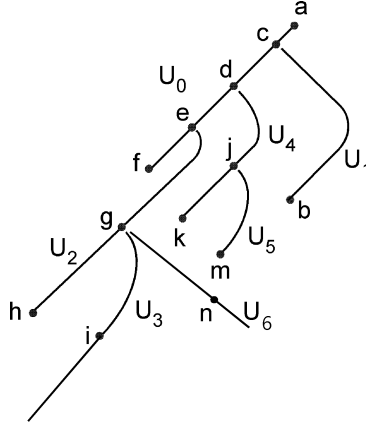


Figure 13: A structuring, Example 3.28(4).

Lemma 3.31 : If (N, \leq, \mathcal{U}) is a structured O-tree, we define $S(N, \leq, \mathcal{U})$ as the relational structure (N, \leq, N_0, N_1) such that N_0 is the set of nodes at even depth and $N_1 := N - N_0$.

(1) The class of structures (N, \leq, N_0, N_1) that represent a structured O-tree is MSO definable.

(2) There is a first-order formula $\nu(X, N_0, N_1)$ expressing in every structure $S(N, \leq, \mathcal{U})$ representing a structured O-tree that a set X belongs to \mathcal{U} .

Proof : (1) The proof is, up to minor details, that Proposition 3.7(1) in [5]. We let $\sigma(N_0, N_1)$ be the corresponding MSO formula.

(2) We let $\nu(X, N_0, N_1)$ express that :

- (i) X is nonempty, linearly ordered and convex,
- (ii) $X \subseteq N_0$ or $X \subseteq N_1$,
- (iii) if $x \in N_0 \cap X$ and $[x, y] \subseteq N_0$ or $[y, x] \subseteq N_0$ then $y \in X$,
- (iv) the same holds for N_1 instead of N_0 .

Let $X \in \mathcal{U}$. Condition 3) of Definition 3.27 yields that, if $x < y$, then $[x, y] \subseteq N_0$ or $[x, y] \subseteq N_1$ if and only if x and y belong to the same line in \mathcal{U} (in particular because if $[x, y] \subseteq N_0$ or $[x, y] \subseteq N_1$, then $[x, y] \subseteq I_k \subseteq U_k$). Conditions (i)-(iv) hold.

Conversely, assume that $\nu(X, N_0, N_1)$ holds. Let $x \in X$. We have $X \subseteq U(x)$: let $y \in X$; if $x < y$, then $[x, y] \subseteq N_0 \cap X$ or $[x, y] \subseteq N_1 \cap X$. Hence, $y \in U(x)$ by the above remark ; if $y < x$, then $x \in U(y)$ and so $y \in U(x)$.

If there is $z \in U(x) - X$, then, as X is an interval, we have $z < X$ or $X < z$. The intervals $[z, x]$ (or $[x, z]$) is contained in N_0 or in N_1 , hence, $z \in X$ by (iii). Contradiction. Hence, $X = U(x)$. The formula $x \in X \wedge \nu(X)$ expresses that $X = U(x)$. \square

Some more notation : Let $T = (N, \leq, \mathcal{U})$ be a structured O-tree with axis A . Let $x \in N - A$ and $L_{\geq}(x) = I_k \uplus I_{k-1} \uplus \dots \uplus I_0$ as in Definition 3.27(b). We define $L^+(x) := I_{k-1} \uplus \dots \uplus I_0$. We have $U_{k-1} = W_{k-1} \uplus I_{k-1}$ for some interval W_{k-1} of U_{k-1} such that $W_{k-1} < I_{k-1}$. With these hypotheses and notation :

- Lemma 3.32** : (1) The interval W_{k-1} is not empty.
(2) For every $y \in \downarrow(W_{k-1})$, we have $L_{>}(x, y) = L^+(x)$.
(3) Every set $L \in \mathfrak{L}$ is of the form $L^+(z)$ for some z .

Proof : (1) If W_{k-1} is empty, then $U_k < I_{k-1} = U_{k-1}$, contradiction with Condition 2) of Definition 3.27(b).

(2) Clear from Condition 2) of Definition 3.27(b).

(3) Let $L = L_{>}(x, y)$. We have $L_{\geq}(x) = I_k \uplus I_{k-1} \uplus \dots \uplus I_0$ and $L_{\geq}(y) = J_{\ell} \uplus J_{\ell-1} \uplus \dots \uplus J_0$ (cf. Condition 3) of Definition 3.27(b)). We have three cases:

Case 1: $I_{m-1} \uplus \dots \uplus I_0 = J_{m-1} \uplus \dots \uplus J_0$ for some $m \leq \min(k, \ell)$ such that $I_m \cap J_m = \emptyset$.

Then $L_{>}(x, y) = L^+(z)$ for any z in $I_m \cup J_m$ (or even in $U_m \cup U'_m$, where $J_m \subseteq U'_m \in \mathcal{U}$). We have also :

$$\begin{aligned} L_{>}(x, y) &= L_{>}(x', y') = L_{>}(x', u) = L_{>}(y', u) \\ &\text{for every } x' \in \downarrow(I_m), y' \in \downarrow(J_m) \\ &\text{and } u \in \downarrow(U_{m-1} - I_{m-1}) = \downarrow(W_{m-1}), \text{ (cf. (1) and (2)).} \end{aligned}$$

Case 2 : $I_{m-1} \subset J_{m-1}$ and $I_p = J_p$ for every $p < m - 1$.

Then $L_{>}(x, y) = L^+(z)$ for any z in I_m (or even in U_m). We have also

$$\begin{aligned} L_{>}(x, y) &= L_{>}(x', u) \text{ for every } x' \in \downarrow(I_m), \text{ and} \\ u &\in \downarrow(U_{m-1} - I_{m-1}) = \downarrow(W_{m-1}). \end{aligned}$$

Case 3 : Similar to Case 2 by exchanging x and y . \square

Example and remarks 3.33 : (1) In Case 1, the sets $\downarrow(I_m)$, $\downarrow(J_m)$ and $\downarrow(U_{m-1} - I_{m-1})$ are three different directions relative to L . In Case 2, $\downarrow(I_m)$ and $\downarrow(U_{m-1} - I_{m-1})$ are similarly different directions.

(2) In the example of Figure 13, we have :

$$\begin{aligned} L_{>}(i, n) &= L_{>}(h, n) = L^+(i) = L^+(n) = L_{\geq}(g) \text{ illustrating Cases 1 and 2,} \\ L_{>}(g, m) &= L_{>}(h, m) = L^+(j) = L^+(k) = L_{\geq}(d) \text{ and} \\ L_{>}(k, m) &= L^+(m) = L_{\geq}(j) \text{ illustrating Case 2.} \square \end{aligned}$$

Lemma 3.34 : There exist FO formulas $\alpha(N_0, N_1, r, x, z)$ and $\beta(N_0, N_1, r, x, z)$ that express the following properties in a structure (N, B, N_0, N_1, r) that satisfies A1-A6 and A8 and defines a structuring of the O-tree $T((N, B), r)$; the corresponding set \mathcal{C}_2 is as in Definition 3.9.

(1) The formula $\alpha(N_0, N_1, r, x, z)$ expresses that $x \in L^+(z)$.

(2) The formula $\beta(N_0, N_1, r, X, z)$ expresses that $X = \overline{Dir_{L^+(z)}(z)}$ and $X \in \mathcal{C}_2$.

Proof : (1) The property $x \in L^+(z)$ is expressed by the following FO formula $\alpha(N_0, N_1, r, x, z)$ defined as :

$$[z \in N_0 \wedge \exists y.(z < y \leq x \wedge y \in N_1)] \vee \\ [z \in N_1 \wedge \exists y.(z < y \leq x \wedge y \in N_0)].$$

(2) Lemma 2.19(2) shows that $X = \overline{Dir_L(z)} \wedge X \in \mathcal{C}_2$ is FO expressible provided $x \in L$ is. Assertion (1) shows precisely that $x \in L^+(z)$ is FO expressible. \square

Proof of Theorem 3.25 : By using the previous lemmas, we now prove the existence of MSO formulas that define in a structure $S = (N, B)$ that satisfies A1-A6 and A8, a marked join-tree T such that $N_T \supseteq N$ and $B = B_T[N]$. In the technical terms of [8] there is a monadic second-order transduction that transforms a structure $S = (N, B)$ into such a marked join-tree (N_T, \leq_T, N) .

The formulas implement the following steps, assuming that S that satisfies A1-A6 and A8.

First step: One chooses $r \in N$, there is no constraint on this choice. One obtains an O-tree $T(S, r)$.

Second step: One guesses a partition (N_0, N_1) of N that defines a structuring of $T(S, r)$, according to Lemma 3.31. As the order on $T(S, r)$ depends on r , the formula $\sigma(N_0, N_1)$ of Lemma 3.31 is transformed into $\sigma'(N_0, N_1, r)$, written with r to define \leq_r .

Third step : All this yields the set $\mathcal{C} = \mathcal{C}_1 \uplus \mathcal{C}_2$ and the associates notions of Definition 3.9 and Lemma 3.32. We will *encode* each set in \mathcal{C}_2 by a unique node z that defines a unique set $\overline{Dir_{L^+(z)}(z)} \in \mathcal{C}_2$. We may have $\overline{Dir_{L^+(z)}(z)} = \overline{Dir_{L^+(w)}(w)}$ where $z \neq w$, but we wish to have each set in \mathcal{C}_2 encoded by a unique node. For insuring this, we choose a set M of nodes such that each set in \mathcal{C}_2 is $\overline{Dir_{L^+(z)}(z)}$ for a unique node $z \in M$. That a set M is correctly chosen can be checked by using the formula β of Lemma 3.34.

We now have the set of nodes of $T(\mathcal{C})$ defined as $N_{T(\mathcal{C})} := (N \times \{1\}) \uplus (M \times \{2\})$ where $(x, 1)$ encodes $N_{\leq}(x)$ and each $w \in M$ in a pair $(w, 2)$ encodes a unique set in \mathcal{C}_2 . Then $T(\mathcal{C}) = (N_{T(\mathcal{C})}, \leq)$ where \leq is the inclusion of the sets encoded by the pairs in $N_{T(\mathcal{C})}$. This partial order is easy to define by means of the formula β .

To sum up, the formulas will use the parameters r, N_0 and M and check they are correctly chosen by existential quantifications :

- r to be the root of the O-tree $T(S, r) = (N, \leq_r)$,
- $N_0 \subseteq N$ such that the structure $(N, \leq_r, N_0, N - N_0)$ represents a structured O-tree,
- M intended to be in bijection with \mathcal{C}_2 .

First-order formulas can check that these parameters are correctly chosen. However, the choices of N_0 and M need set quantifications.

We obtain a join-tree T' with set of nodes $N_{T'} = (N \times \{1\}) \uplus (M \times \{2\})$. Then $S = (N, B)$ is isomorphic to $(N \times \{1\}, B_{T'}[N \times \{1\}])$ where $(x, 1)$ corresponds to $x \in N$. Hence, S is defined by $(N_{T'}, \leq_{T'}, N \times \{1\})$ constructed by MSO formulas. \square

Remark 3.35 : *About join-completion.*

The join-completion builds an O-tree T from the sets $U(x, y)$, cf. Definition 1.3(b). If x and y have no join, then $U(x, y)$ defined as $N_{\leq}(L_{\geq}(x, y))$ is equal to $N_{\leq}(L_{>}(x, y))$. By means of a structuring of T , such a set is of the form $N_{\leq}(L^+(z))$, hence can be encoded by a single node z . The technique of Theorem 3.25 is applicable to prove that join-completion is an MSO transduction.

4 Embeddings in the plane

In order to give a geometric characterization of join-trees and of induced betweenness in quasi-trees (equivalently, in join-trees), we show how a structured join-tree can be embedded in portions of straight lines in the plane that form a *topological tree*.

Definition 4.1 : *Trees of lines in the plane.*

(a) In the Euclidian plane, let $\mathcal{L} = (L_i)_{i \in \mathbb{N}}$ be a family of straight half-lines¹⁷ (simply called *lines* below) with respective origins $o(L_i)$, that satisfies the following conditions :

(i) if $i > 0$, then $o(L_i) \in L_j$ for some $j < i$,

(ii) for all $i, j \in \mathbb{N}$, $i \neq j$, the set $L_i \cap L_j$ is $\{o(L_i)\}$ or $\{o(L_j)\}$ or is empty.

(We may have $o(L_i) = o(L_j)$).

We call \mathcal{L} a *tree of lines* : the union of the lines L_i is a connected set $\mathcal{L}^\#$ in the plane. A *path* (resp. a *cycle*) in $\mathcal{L}^\#$ is a homeomorphism h of the interval $[0, 1]$ of real numbers (respectively of the circle S^1) into $\mathcal{L}^\#$ such that $h(0) = x$ and $h(1) = y$ in the case of a path. For any two distinct $x, y \in \mathcal{L}^\#$, there is a unique path from x to y (it "follows the lines"), and consequently, there is no cycle. This path goes through lines L_k such that $k \leq \max\{i, j\}$ where $x \in L_i$ and $y \in L_j$, hence, through finitely many of them. This path uses a single interval of each line it goes through, otherwise, there is a cycle.

(b) We obtain a ternary *betweenness* relation :

$$B_{\mathcal{L}}(x, y, z) : \iff \neq (x, y, z) \text{ and } y \text{ is on the path between } x \text{ and } z.$$

(c) On each line L_i , we define a linear order as follows :

$$x \preceq_i y \text{ if and only if } y = x \text{ or } y = o(L_i) \text{ or } y \text{ is between } x \text{ and } o(L_i).$$

On $\mathcal{L}^\#$, we define a partial order by :

$$x \preceq y \text{ if and only if } x = y \text{ or}$$

$$x \prec_{i_k} o(L_{i_k}) \prec_{i_{k-1}} o(L_{i_{k-1}}) \prec_{i_{k-2}} \dots \prec_{i_1} o(L_{i_1}) \prec_{i_0} y$$

$$\text{for some } i_0 < i_1 < \dots < i_k. \text{ If } k = 0, \text{ then } x \prec_{i_0} y.$$

¹⁷One could equivalently use bounded segments of straight lines because on each such segment, one can designate countably many points.

It is clear that $(\mathcal{L}^\#, \preceq)$ is an uncountable rooted O-tree : for each x in $\mathcal{L}^\#$, the set $\{y \in \mathcal{L}^\# \mid x \preceq y\}$ is linearly ordered with greatest element $o(L_0)$.

Definition 4.2 : *Embeddings of join-trees in trees of lines.*

Let $T = (N, \leq, \mathcal{U})$ be a structured join-tree (cf. Definition 3.27). An *embedding* of T into a tree of lines \mathcal{L} is an injective mapping $m : N \rightarrow \mathcal{L}^\#$ such that:

for each $U \in \mathcal{U}$, m is order preserving : $(U, \leq) \rightarrow (L_i, \preceq_i)$ for some $i \in \mathbb{N}$, and if U is not the axis, then $m(\text{lsub}(U)) = o(L_i)$.

Lemma 4.3 : If T is a structured join-tree embedded by m into a tree of lines \mathcal{L} , then, its betweenness satisfies :

$$B_T(x, y, z) \iff [\neq(x, y, z) \wedge B_{\mathcal{L}}(m(x), m(y), m(z))].$$

Proof sketch : Let $(x, y, z) \in B_T$. Assume that $x < y < x \sqcup z$ and let us compare $L_{\geq}(x) = I_k \uplus I_{k-1} \uplus \dots \uplus I_0$ and $L_{\geq}(z) = J_\ell \uplus J_{\ell-1} \uplus \dots \uplus J_0$ (as in the proof of Lemma 3.32(3)). There are three cases. In each of them, we have a path in T between x and z , that goes through y and is a concatenation of intervals of lines of the structuring of T . By concatenating the corresponding segments of the lines in \mathcal{L} , we get a (topological) path between $m(x)$ and $m(z)$ that contains $m(y)$. Hence, we have $(m(x), m(y), m(z))$ in $B_{\mathcal{L}}$. The proof is similar in the other direction. \square

Theorem 4.4 : If \mathcal{L} is a tree of lines and N is a countable subset of $\mathcal{L}^\#$, then $S := (N, B_{\mathcal{L}}[N])$ is in **IBQT**, *i.e.* is an induced betweenness structure in a quasi-tree. Conversely, every structure in **IBQT** is isomorphic to some $S = (N, B_{\mathcal{L}}[N])$ of the above form.

Proof : If \mathcal{L} is a tree of lines and $N \subset \mathcal{L}^\#$ is countable, then $S := (N, B_{\mathcal{L}}[N])$ is in **IBQT**. A witnessing join-tree T is built as follows. Its set of nodes is $N \cup O$ where O is the set of origins of all lines in \mathcal{L} . Its order is the restriction to $N \cup O$ of the order \preceq on $\mathcal{L}^\#$. Then $(N, B_{\mathcal{L}}[N]) = (N, B_T[N])$ hence belongs to **IBQT**.

Conversely, let $S = (N, B_T[N])$ such that T is a structured join-tree. It is isomorphic to $(N, B_{\mathcal{L}}[N])$ for some tree of lines by the following proposition. \square

Proposition 4.5 : Every structured join-tree embeds into a tree of lines \mathcal{L} .

The proof will use some notions of geometry relative to positions of lines in the plane.

Definitions 4.6 : *Angles and line drawings.*

An orientation of the plane, say the trigonometric one is fixed.

(a) Let L, K be two lines with same origin. Their *angle* $L \triangle K$ is the real number α , $0 \leq \alpha < 2\pi$, such that L becomes K by a rotation of angle α .

If $o(K)$ is in $L - \{o(L)\}$, we define $L \triangle K := L' \triangle K$ where L' is the unbounded half-line included in L with origin $o(K)$.

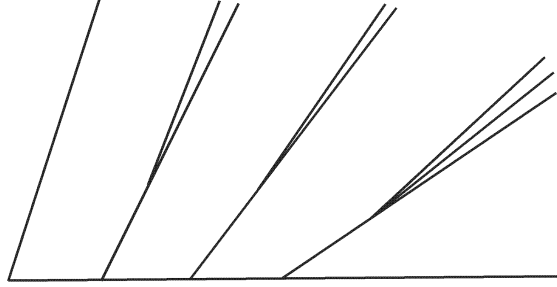


Figure 14: For the proof of Lemma 4.8.

(b) For a line L , an angle α such that $0 < \alpha < \pi$ and $O \in L$, we define $S(L, O, \alpha)$ as the union of the lines K with origin O such that $0 \leq L \triangle K < \alpha$. We call *sector* such a set.

Lemma 4.7 : For given L and α as above, one can draw countably many lines with origin $o(L)$ inside the sector $S(L, o(L), \alpha)$.

Proof : We draw $K_1, K_2, \dots, K_i, \dots$ such that $L \triangle K_1 = \alpha/2$ and $K_i \triangle K_{i+1} = \alpha/2^{i+1}$ for each i . \square

Lemma 4.8 : Let L, α be as above and X be a countable set enumerated as $\{x_1, x_2, \dots, x_i, \dots\} \subseteq L - \{o(L)\}$. One can draw lines $K_1, K_2, \dots, K_i, \dots$ in the sector $S(L, o(L), \alpha)$ in such a way that $o(K_i) = x_i$ for each i , no two lines are parallel or meet except at their origins, and none is included in L .

Proof : We must have $0 < L \triangle K_i < \alpha$ for each i . For each i , we let $\gamma_i := \alpha/2^{i+1}$ and $\beta_i := \Sigma\{\gamma_j \mid x_j \prec x_i\} < \alpha$ where $x_j \prec x_i$ means that x_i is between $o(L)$ and x_j . Then, we draw $K_1, K_2, \dots, K_i, \dots$ with respective origins $x_1, x_2, \dots, x_i, \dots$ such that $L \triangle K_i = \beta_i$. \square

For each i , the sector $S(K_i, x_i, \gamma_i)$ contains nothing else than K_i . By Lemma 3.8, one can draw inside $S(K_i, x_i, \gamma_i)$ countably many lines with origin x_i .

Proof of Proposition 4.5 : Let \mathcal{U} be a structuring of a join-tree T . Let A be the axis. Hence, $lsub(A)$ is undefined.

The depth $\partial(U)$ of $U \in \mathcal{U}$ is defined in Definition 3.27 for O-trees. It satisfies the following induction :

$$\begin{aligned} \partial(A) &= 0, \\ \partial(U) &= \partial(U') + 1 \text{ if } U' \text{ has the minimal depth such that } lsub(U) \in U'. \\ & \text{(Hence, } lsub(U) \neq lsub(U')). \end{aligned}$$

We draw lines L_0, L_1, \dots and define an embedding m such that the conditions of Definition 4.2 hold. We first draw L_0 and define m on A , as required. We choose α such that $0 < \alpha < \pi$. All further constructions will be inside the sector

$S(L_0, o(L_0), \alpha)$. By Lemmas 4.7 and 4.8, we can draw the lines of depth 1. There is space for drawing the lines of depth 2. We continue in this way¹⁸. \square

5 Conclusion

We have defined betweenness relations in different types of generalized trees, and obtained first-order or monadic second-order axiomatizations. In Section 4, we have given a geometric characterization of join-trees and the associated betweenness relations.

We have proved that the class **IBQT** of induced substructures of the first-order class **QT** of quasi-trees is first-order axiomatizable. This is not an immediate consequence of the FO axiomatization of **QT** as shown in the appendix.

We conjecture that betweenness in O-trees is *not* first-order definable (although the class of O-trees is). We also conjecture that the class **IBO** of induced betweenness relations in O-trees has a monadic second-order axiomatization.

In [5], we have defined quasi-trees and join-trees of different kinds from regular infinite terms, and proved they are equivalently the unique models of monadic second-order sentences. Both types of characterizations yield finitary descriptions and decidability results, in particular for deciding isomorphism. In a future work, we will extend these results to O-trees and to their betweenness relations.

Other works on betweenness.

Betweenness in partial orders (of any cardinality) is axiomatized by J. Lihova in [12] by an infinite set of universal first-order sentences. We prove in [7] that this set cannot be replaced by a single first-order sentence but that it can be by a single monadic second-order one.

Several betweenness notion in graphs are surveyed in [1]. Motivated by the study of convex geometries, V. Chvatal studies in [2] the betweenness in finite triangulated graphs, relative to induced paths: y is between x and z if it is an intermediate vertex on a chordless path between x and z .

6 Appendix : Induced relational structures

The following example shows that the FO characterization of **IBQT** does not follow from the FO characterization of the class **QT**.

Counter-example 6.1 : *Taking induced substructures does not preserve first-order axiomatizability.*

We prove a little more. We define an FO class \mathcal{C} of relational structures such that $Ind(\mathcal{C})$, the class of induced substructures of those in \mathcal{C} , is not MSO axiomatizable.

¹⁸The angles γ_i are of course very small as depth increases.

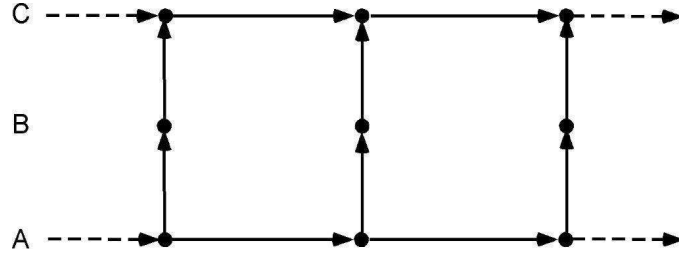


Figure 15: The ladder of Example 6.1.

Let R be a binary relation symbol and A, B, C be unary ones. We let \mathcal{C} be the class of structures $S = (V, R, A, B, C)$ that satisfy the following conditions (i) to (iv) :

- (i) The sets defined by A, B, C form a partition of V ,
- (ii) R is irreflexive.

Hence S can be considered as a loop-free directed graph whose vertices form the set V and are "colored" by A, B or C . Further conditions are as follows :

- (iii) each infinite connected component of S is a "horizontal ladder" that is infinite in both directions, and a portion of which is shown in Figure 15; the sets of A - and C -colored vertices form two biinfinite horizontal directed paths.
- (iv) Each finite connected component is a closed "ring", with two directed cycles of A - and C -colored vertices ; Figure 15 shows a portion of such a ring.

By a *successor* (or *predecessor*) of x , we mean a vertex y such that $(x, y) \in R$ (or $(y, x) \in R$ respectively).

Conditions (iii) and (iv) can be expressed by an FO sentence saying in particular :

(a) Every vertex x_A in A has a unique successor y_A in A and a unique successor x_B in B ; x_B has a unique successor x_C in C ; y_A has a unique successor in y_B in B ; y_B has a unique successor y_C in C that is also the unique successor of x_C in C .

(b) Similar condition with predecessor instead of successor for A - and C -colored vertices.

(c) There are no other edges than those specified by (a) and (b).

Let us assume that $Ind(\mathcal{C})$ is characterized by an MSO sentence ψ . We will derive a contradiction.

Let θ be an MSO sentence expressing that a structure $S = (V, R, A, B, C)$ consists of six vertices $x_A, z_A, x_B, z_B, x_C, z_C$, of directed edges x_Ax_B, x_Bx_C, z_Az_B and z_Bz_C , of a directed path p_A of A -colored vertices from x_A to z_A and of a directed path p_C of C -colored vertices from x_C to z_C . These conditions imply that V is finite. The construction of θ is routine. In particular, the existence of paths p_A and p_C can be expressed in MSO logic with set quantifications. First-order logic cannot express transitive closures. cf. [8].

Then, the structures that satisfy $\theta \wedge \psi$ are exactly those that satisfy θ and have paths p_A and p_C of equal lengths. But such an equality is not MSO expressible (cf. [8]). Hence, no MSO sentence ψ can characterize $Ind(C)$. \square

This example shows that the first-order axiomatization of the class **IBQT** (Theorem 3.1) is not an immediate consequence of the first-order axiomatization of quasi-trees. To the opposite, the proof of Proposition 2.9 has used an argument based on the structure of logical formulas.

References

- [1] M. Changat, P. Narasimha-Shenoi, and G. Seethakuttyamma, Betweenness in graphs: A short survey on shortest and induced path betweenness, *AKCE International Journal of Graphs and Combinatorics*, **16** (2019) 96-109. (Available on ScienceDirect.com, <https://doi.org/10.1016/j.akcej.2018.06.007>)
- [2] V. Chvatal, Antimatroids, betweenness, convexity, in *Research Trends in Combinatorial Optimization*, Springer, 2008, pp. 57-64.
- [3] B. Courcelle, Regularity equals monadic second-order definability for quasi-trees, in *Fields of Logic and Computation II, Lec. Notes Comput. Sci.* **9300** (2015) 129-141.
- [4] B. Courcelle, Several notions of rank-width for countable graphs, *J. Comb. Theory, Ser. B.* **123** (2017) 186-214.
- [5] B. Courcelle, Algebraic and logical descriptions of generalized trees, *Logical Methods in Computer Science* **13** (2017) Issue 3.
- [6] B. Courcelle, Betweenness in order-theoretic trees, to appear in *Fields of Logic and Computation III, Lecture Notes in Computer Science*, Springer 2020.
- [7] B. Courcelle, Betweenness of partial orders, April 2020, submitted for publication, see <https://hal.archives-ouvertes.fr/hal-02547130>
- [8] B. Courcelle and J. Engelfriet, *Graph structure and monadic second-order logic, a language theoretic approach*, Cambridge University Press, 2012.

- [9] R. Fraïssé, *Theory of relations*, *Studies in Logic*, Volume **145**, North-Holland, 2000.
- [10] W. Hodges, *Model theory*, Cambridge University Press, 1993.
- [11] L. Libkin, *Elements of finite model theory*, Springer, 2004.
- [12] J. Lihova, Strict-order betweenness, *Acta Univ. M. Belii Ser. Math.* **8** (2000) 27-33. Available from <https://actamath.savbb.sk/acta0804.shtml>.
- [13] S. Oum, Rank-width and vertex-minors. *J. Comb. Theory, Ser. B* **95** (2005) 79-100.