Relating Description Complexity to Entropy

Reijo Jaakkola
Tampere University, Finland

Antti Kuusisto
Tampere University, Finland
University of Helsinki, Finland

Miikka Vilander
Tampere University, Finland

Abstract

We demonstrate some novel links between entropy and description complexity, a notion referring to the minimal formula length for specifying given properties. Let MLU be the logic obtained by extending propositional logic with the universal modality, and let GMLU be the corresponding extension with the ability to count. In the finite, MLU is expressively complete for specifying sets of variable assignments, while GMLU is expressively complete for multisets. We show that for MLU, the model classes with maximal Boltzmann entropy are the ones with maximal description complexity. Concerning GMLU, we show that expected Boltzmann entropy is asymptotically equivalent to expected description complexity multiplied by the number of proposition symbols considered. To contrast these results, we prove that this link breaks when we move to considering first-order logic FO over vocabularies with higher-arity relations. To establish the aforementioned result, we show that almost all finite models require relatively large FO-formulas to define them. Our results relate to links between Kolmogorov complexity and entropy, demonstrating a way to conceive such results in the logic-based scenario where relational structures are classified by formulas of different sizes.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics; Theory of computation → Finite Model Theory

Keywords and phrases finite model theory, entropy, formula size, randomness, formula size game

Digital Object Identifier 10.4230/LIPIcs.STACS.2023.38

Funding Antti Kuusisto was supported by the Academy of Finland project *Theory of computational logics*, grant numbers 352419, 352420, 353027, 324435, 328987. Furthermore, Antti Kuusisto and Miikka Vilander were supported by the Academy of Finland consortium project *Explaining AI via Logic* (XAILOG), grant number 345612.

1 Introduction

In this article we investigate links between description complexity and entropy. By description complexity of a model, we mean the minimal length of a formula that specifies the model up to a maximal possible extent. With a strong enough logic, this amounts to investigating the length of formulas specifying models up to isomorphism, but this is by no means the only interesting scenario. By the description complexity of a class of models, we mean the minimal length of a formula defining that class. In this paper we are particularly interested in the description complexity of completely specified model classes, i.e., equivalence classes of logics. The main objective of the paper is to point out links between description complexity and entropy. By entropy, we refer essentially to Shannon's entropy and the earlier notion of Boltzmann entropy from statistical mechanics.

We first consider models with unary relational vocabularies. We study two related logics, MLU and GMLU. The logic MLU is the extension of propositional logic with the universal modality $\blacklozenge$, also known as global modality. The truth definition states that $\mathfrak{M}, w \vDash \blacklozenge\varphi$ if

$\mathfrak{M}, u \vDash \varphi$ for some u in the domain of $\mathfrak{M}$. Thus, in the finite, this logic is tuned to specify precisely which variable assignments are present in the model considered. The system GMLU is the extension of MLU with the ability to count: we have $\mathfrak{M}, w \vDash \blacklozenge^{\geq d}\varphi$ if $\mathfrak{M}, u \vDash \varphi$ for at least d points u in the domain of $\mathfrak{M}$. We note that when limiting to models with a finite unary vocabulary and a fixed finite bound on domain size, GMLU is expressively complete, being able to define all classes of models closed under isomorphism. While MLU can fully specify which *set* of assignments is present in a model, GMLU can lift this specification to the level of *multisets*.

Let τ be a finite unary relational vocabulary, and let $\mathrm{Mod}_n(\tau)$ denote the class of τ-models over the fixed domain $W = \{1, \ldots, n\}$. Let $\equiv_{\mathrm{MLU}}$ and $\equiv_{\mathrm{GMLU}}$ denote the logical equivalence relations of MLU and GMLU over $\mathrm{Mod}_n(\tau)$. We first prove that among the classes of $\equiv_{\mathrm{MLU}}$, the class with the largest description complexity is the class with the largest Boltzmann entropy. This means that the models with the largest description complexity belong to the class that has the largest Boltzmann entropy. We then move on to investigating GMLU. Let $\langle H_B \rangle$ denote the expected Boltzmann entropy over the equivalence classes of $\equiv_{\mathrm{GMLU}}$, with the probability of an individual class being its size divided by the size of $\mathrm{Mod}_n(\tau)$. Let $\langle C \rangle$ denote the expected description complexity of a model chosen randomly from $\mathrm{Mod}_n(\tau)$, and let $|\tau|$ denote the size of the vocabulary τ. We will prove that

$$\langle H_B \rangle \sim |\tau| \langle C \rangle \tag{1}$$

that is, $\langle H_B \rangle$ is asymptotically equivalent to $|\tau|\langle C \rangle$. This gives an intimate relationship between $\langle C \rangle$ and Boltzmann entropy. To obtain a link to Shannon entropy, we simply note that the Shannon entropy of the distribution of models based on $\equiv_{\mathrm{GMLU}}$ is equal to $\langle H_B \rangle - \log(|\mathrm{Mod}_n(\tau)|)$.

We then move on to investigating general (finite) relational vocabularies. Our main result there is that the expected description complexity of classes of FO grows asymptotically faster with domain size than the corresponding expected Boltzmann entropy. To establish this result, we show that almost all models require relatively large FO-formulas to define them.

Concerning related work, there exist well known relationships between Kolmogorov complexity and entropy. Notably, for any computable distribution, the expected Kolmogorov complexity can be linked, within a constant, to Shannon entropy. See for example [6, 9, 10, 14] for discussions of the issue. The article [14] discusses some generalizations and shows, e.g., that the relationship fails in the general case for Rényi and Tsallis entropies. Links between description lengths and entropy are fundamentally interesting, linking syntactic issues to semantic randomness. Most notable results in the field concern variants of Kolmogorov complexity. The aim of the current article is to provide one way of demonstrating how these results extend beyond the realm of binary strings and descriptions via programs. The link given in Equation (1) elucidates nicely the relationship between the syntax of GMLU and models with unary vocabularies. The result on FO provides contrast to this and warns against overselling the analogy between description complexities and entropy. However, we conjecture that even for FO, a monotone Galois connection can be demonstrated between description complexities of a relevant collection of classes and related Boltzmann entropies, but this is left for future work for lack of space.

Links between description complexity and the properties of described model classes relate also to the relationship between classifier size and classified data. This topic is relevant, for example, from the point of view of current research on explainability in AI. For work on this topic, see, e.g., [8, 2].

Concerning other related work, we turn attention to the proof techniques used in the paper. One of the main tools we use is the framework of logic-related games. We note that standard Ehrenfeucht-Fraïssé games, and their variants such as bisimulation games, do not

suffice for the purposes of this article. Thus we utilize *formula size games* for MLU and GMLU instead. Generally, the first formula size game was defined for propositional logic by Razborov in [13]. A better known version is the game of Adler and Immerman for CTL in [1]. The game for MLU resembles the similar game developed in [7] which was there also used to demonstrate a nonelementary succinctness gap between modal logic and FO. For GMLU, we develop a suitable game by extending the game for MLU. The hard part is using the games in a suitable way. We accomplish this by using the fact that models in classes with large Boltzmann entropy realize a rich number of types. In addition to games, we also use various techniques for estimating Boltzmann entropy and description complexity, e.g., Stirling's approximation, the weak law of large numbers and counting arguments.

2 Preliminaries

Let $f, g : \mathbb{N} \to \mathbb{R}_{\geq 0}$. We use $f = \mathcal{O}(g)$ to denote that $f \leq Cg(n)$, for some constant $C > 0$ and large enough n. If we want to emphasize that the implied constant C depends on some parameter p (which is independent of n), we will write $f = \mathcal{O}_p(g)$. We use $f = \Omega(g)$ to denote that $f(n) \geq Cg(n)$, for some constant $C > 0$ and large enough n. Finally, we use $f = \Theta(g)$ to denote that $f = \mathcal{O}(g)$ and $f = \Omega(g)$. We say that f is **asymptotically** g, if $\lim_{n\to\infty} f/g = 1$ and we denote this by $f \sim g$. By log we mean logarithm to base two.

The following variants of classical results will be useful for our purposes.

▶ **Proposition 1** (Stirling's approximation [5]). $\log(n!) = n\log(n) - n\log(e) + \Theta(\log(n))$

▶ **Proposition 2** (Weak law of large numbers [11]). *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of Bernoulli random variables with success probability $p := \Pr[X_n = 1]$. Then for every $\delta > 0$ we have that*

$$\lim_{n\to\infty} \Pr\left[\left|p - \tfrac{1}{n}\sum_{i=1}^{n} X_n\right| < \delta\right] = 1.$$

We next define the logics studied in this work. Let $\tau = \{p_1, \ldots, p_k\}$ be a set of proposition symbols. The syntax of **graded universal modal logic** GMLU[τ] is generated as follows.

$$\varphi := \blacklozenge^{\geq d}\psi \mid \blacksquare^{<d}\psi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \blacklozenge^{\geq d}\varphi \mid \blacksquare^{<d}\varphi$$
$$\psi := p \mid \neg p \mid \psi \vee \psi \mid \psi \wedge \psi$$

Here $p \in \tau$ and $d \in \mathbb{N}$. Notice that by design, the formulas of GMLU[τ] only contain proposition symbols that occur in the scope of a global modal operator $\blacklozenge^{\geq d}$ or $\blacksquare^{<d}$. Additionally, all formulas are in negation normal form. (In the sequel, the notation $\neg\varphi$ will always mean the negation normal form formula, where the negation has been pushed to the level of literals.) Now, let $\mathfrak{M}$ be a Kripke model with universe W. The semantics of the global graded modalities are defined as follows: $(\mathfrak{M}, w) \models \blacklozenge^{\geq d}\varphi \Leftrightarrow$ there are at least d points $v \in W$ such that $(\mathfrak{M}, v) \models \varphi$. Additionally, $(\mathfrak{M}, w) \models \blacksquare^{<d}\varphi \Leftrightarrow (\mathfrak{M}, w) \models \neg\blacklozenge^{\geq d}\neg\varphi$. Intuitively this means that all points in $\mathfrak{M}$ satisfy φ, except for less than d exceptions. The rest of the semantics is defined as usual in propositional logic. Note that $\blacklozenge^{\geq d}$ and $\blacksquare^{<d}$ are dual to each other. (We note that in this article, modal logics will always have a strictly unary vocabulary, so Kripke models will not have an accessibility relation as part of the relational structure involved.)

Given a Kripke model $\mathfrak{M}$ over τ and $\varphi \in$ GMLU[τ], we define the point-free truth relation such that $\mathfrak{M} \models \varphi \Leftrightarrow$ for every $w \in W$, we have $\mathfrak{M}, w \models \varphi$. Since no propositional symbol occurs outside the scope of a global modality, $\mathfrak{M} \models \varphi$ iff there is some $w \in W$ for which $\mathfrak{M}, w \models \varphi$. Hence the truth of any formula of GMLU is independent of the evaluation point w. The property that truth is always independent of the evaluation point is the reason we defined

GMLU so that proposition symbols must occur in the scope of modalities. The fragment of GMLU[τ] where $d = 1$ for all modalities is called **universal modal logic** MLU[τ]. This logic has only the modalities $\blacklozenge^{\geq 1}$ and $\blacksquare^{<1}$, and we denote these with $\blacklozenge$ and $\blacksquare$ for simplicity.

A 1-**type** π over τ is a maximally consistent set of literals (propositional symbols and their negations). This means that π has exactly one of p or $\neg p$ for each $p \in \tau$. The set of all 1-types over τ is denoted by $\boldsymbol{\alpha}_\tau$. Given a Kripke model $\mathfrak{M}$ over τ and $w \in W$, we let $\mathrm{tp}_{\mathfrak{M}}[w]$ denote the unique 1-type that w **realizes**.

The **size** of a formula $\varphi \in \mathrm{GMLU}[\tau]$, denoted $size(\varphi)$, is defined as follows:

- $size(\alpha) = 1$ for a literal α,
- $size(\varphi \vee \psi) = size(\varphi \wedge \psi) = size(\varphi) + size(\psi) + 1$,
- $size(\blacklozenge^{\geq d}\varphi) = size(\blacksquare^{<d}\varphi) = size(\varphi) + d$.

We emphasize that according to our definition *all literals have the same size*. The motivation for this is to consider negative (i.e., negated) information and positive (i.e., non-negated) information as equal in relation to formula size. This also explains the convention of defining GMLU such that formulas are in negation normal form.

We will also consider standard first-order logic FO. Let $\tau = \{R_1, \ldots, R_k\}$ be a set of relation symbols. The syntax of FO[τ] is generated by the following grammar:

$$\varphi := x = y \mid \neg x = y \mid R(\overline{x}) \mid \neg R(\overline{x}) \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \exists x\varphi \mid \forall x\varphi,$$

where $\overline{x}$ is a tuple of variables. We use the standard semantics of FO[τ]. The **size** of a formula $\varphi \in \mathrm{FO}[\tau]$, denoted $size(\varphi)$, is defined as follows:

- $size(\alpha) = 1$ for a literal α,
- $size(\varphi \vee \psi) = size(\varphi \wedge \psi) = size(\varphi) + size(\psi) + 1$,
- $size(\exists x\varphi) = size(\forall x\varphi) = size(\varphi) + 1$

Again we emphasize that according to our definition, all literals have the same size.

Let $\mathcal{L} = (L, \models)$ be a logic and $\mathcal{M}$ a *finite class of models*. The class $\mathcal{M}$ is here considered fixed and known from the context. We say that a formula $\varphi \in L$ of $\mathcal{L}$ **defines** a set $M \subseteq \mathcal{M}$ if for all $\mathfrak{M} \in \mathcal{M}$, we have $\mathfrak{M} \models \varphi$ iff $\mathfrak{M} \in M$. Such a set M is called $\mathcal{L}$**-definable** (with respect to $\mathcal{M}$). Given an $\mathcal{L}$-definable set M, its $\mathcal{L}$**-description complexity** $C_{\mathcal{L}}(M)$ is the size of a minimum size formula $\varphi \in \mathcal{L}$ which defines M. Now, if $\mathcal{L}$ is closed under negation (as all the logics in this paper are), then the relation "$\mathfrak{M}$ and $\mathfrak{N}$ satisfy the same $\mathcal{L}$-formulas" induces a partition of $\mathcal{M}$ denoted by $\equiv_{\mathcal{L}}$. The $\mathcal{L}$-description complexity *of a model* $\mathfrak{M}$ with respect to $\equiv_{\mathcal{L}}$ is $C_{\equiv_{\mathcal{L}}}(\mathfrak{M}) := C_{\mathcal{L}}(M)$, where M is the equivalence class of $\mathfrak{M}$. For brevity, we formulate the results below only for description complexities of classes rather than models.

For an example of description complexity, consider MLU[τ] for the singleton alphabet $\tau = \{p\}$. Models, where p is true in every point is a class of the partition $\equiv_{\mathrm{MLU}[\tau]}$. The description complexity of this class is 2 as the minimum size formula that defines the class is $\blacksquare p$.

Let $\mathcal{M}$ be a finite class of models and let $\equiv$ be an *arbitrary* equivalence relation over $\mathcal{M}$. Given an equivalence class $M \subseteq \mathcal{M}$, we define its **Boltzmann entropy** as $H_B(M) := \log(|M|)$. This terminology is borrowed from statistical mechanics, where the Boltzmann entropy of a macrostate is the quantity $k_B \ln(\Omega)$. Here k_B is the Boltzmann constant, ln the natural logarithm and Ω the number of microstates associated with the macrostate. Note that in our definition, we use the binary logarithm. As a measure of randomness, it is natural to define the Boltzmann entropy *of a model* $\mathfrak{M}$ as $H_B^{\equiv}(\mathfrak{M}) := H_B(M)$, where M is the equivalence class of $\mathfrak{M}$. This reflects the *informal intuition* that often the randomness of an object x is in fact more related to the size (or richness) of a similarity class of objects that x belongs to rather than to x itself. Consider, for example, the equivalence classes that a sufficiently weak logic defines over the universe of binary strings of a fixed finite length. As a general intuition, it is natural to associate a formula φ (or the class it defines) with a macrostate, while the models of φ are then the corresponding microstates.

Let $\{M_i \mid i \in I\}$ enumerate the equivalence classes of $\equiv$. As they form a partition of $\mathcal{M}$, we have the following natural probability distribution over the equivalence classes: $p_{\equiv}(M_i) := |M_i|/|\mathcal{M}|$. Given a random variable $X : \{M_i \mid i \in I\} \to \mathbb{R}_{\geq 0}$, we use $\langle X \rangle$ to denote its expected value with respect to $p_{\equiv}$. Now, suppose we are in a context where we have fixed a finite universe $\mathcal{M}$ of models. Let $\equiv_{\mathrm{GMLU}} \subseteq \mathcal{M} \times \mathcal{M}$ be the corresponding equivalence relation of GMLU. Suppose $\{M_i \mid i \in I\}$ enumerates the equivalence classes of $\equiv_{\mathrm{GMLU}}$. Recall that $C_{\mathrm{GMLU}}(M_i)$ denotes the GMLU description complexity of the class M_i. Let $p_{\equiv_{\mathrm{GMLU}}}(M_i)$ be the corresponding probability $|M_i|/|\mathcal{M}|$. In this paper, we denote by $\langle C \rangle$ the expected description complexity of GMLU, that is, $\langle C \rangle = \sum_{i \in I} p_{\equiv_{\mathrm{GMLU}}}(M_i) C_{\mathrm{GMLU}}(M_i)$. The class $\mathcal{M}$ will be clear from the context. Note that trivially the same expected value is obtained for the description complexity of *models* over $\mathcal{M}$ if we give every model $\mathfrak{M} \in \mathcal{M}$ the probability $1/|\mathcal{M}|$ (the uniform distribution).

We note that it would be natural to define the Boltzmann entropy of an equivalence relation $\equiv$ as the expected value $\langle H_B \rangle$ of H_B with respect to the above natural probability distribution $p_{\equiv}$. The value $\langle H_B \rangle$ is closely related to the **Shannon entropy** $H_S(\equiv)$ of $\equiv$, which we define as the expected value of the random variable $M_i \mapsto -\log(p_{\equiv}(M_i))$. More explicitly, we define that $H_S(\equiv) := -\sum_{i \in I} p_{\equiv}(M_i) \log(p_{\equiv}(M_i))$. Note that this expression is always well-defined, since $M_i \neq \varnothing$, for every $i \in I$. The following result is established in Appendix A.1. Note that the expected value of $H_B^{\equiv}$ over the uniform distribution on $\mathcal{M}$ is equal to $\langle H_B \rangle$, so the result could also be formulated for single models.

▶ **Proposition 3.** *Let $\mathcal{M}$ be a finite class of models and $\equiv \subseteq \mathcal{M} \times \mathcal{M}$ an equivalence relation over $\mathcal{M}$. Then $H_S(\equiv) + \langle H_B \rangle = \log(|\mathcal{M}|)$.*

There exist results in the literature on entropy similar to the above, see, e.g., [3] and [15]. By the proposition, both the Shannon entropy of $\equiv$ and the expected Boltzmann entropy of $\equiv$ cannot be simultaneously large (meaning close to their maximum value $\log(|\mathcal{M}|)$). Indeed, suppose we do not alter $\mathcal{M}$, so $\log(|\mathcal{M}|)$ is constant. Now suppose we alter $\equiv$ so that $H_S(\equiv)$ is increased. This lowers $\langle H_B \rangle$. Vice versa, increasing $\langle H_B \rangle$ lowers $H_S(\equiv)$. Shannon entropy and expected Bolzmann entropy are complementary quantities, summing to a constant.

3 MLU: The largest class has maximal description complexity

Fix $\tau = \{p_1, \ldots, p_k\}$ and let $\mathrm{Mod}_n(\tau)$ be the set of Kripke models over τ and the *fixed universe* $W = \{1, \ldots, n\}$. In this section we consider the equivalence $\equiv_{\mathrm{MLU}[\tau]}$ as defined above and denote it by $\equiv$. We show that in this canonical partition, the largest class, which is the one with the largest Boltzmann entropy, has maximal MLU[τ]-description complexity.

The equivalence classes of $\equiv$ can be described easily. For Kripke models $\mathfrak{M}_1, \mathfrak{M}_2 \in \mathrm{Mod}_n(\tau)$, we have $\mathfrak{M}_1 \equiv \mathfrak{M}_2 \Leftrightarrow \{\mathrm{tp}_{\mathfrak{M}_1}[w] \mid w \in W\} = \{\mathrm{tp}_{\mathfrak{M}_2}[w] \mid w \in W\}$. That is, each equivalence class is uniquely determined by the 1-types realized in it. As the number of 1-types over τ is 2^k, the number of equivalence classes of $\equiv$ is $2^{2^k} - 1$. Given a set $\Pi \subseteq \boldsymbol{\alpha}_\tau$, we let M_Π be the equivalence class that has the models that realize exactly the 1-types in Π.

Note the equivalence class M_Π of any set $\Pi \subseteq \boldsymbol{\alpha}_\tau$ can be defined by the following formula:

$$\varphi(\Pi) := \bigwedge_{\pi \in \Pi} \blacklozenge \psi(\pi) \wedge \blacksquare \Big(\bigwedge_{\pi \in \boldsymbol{\alpha}_\tau \setminus \Pi} \neg \psi(\pi) \Big),$$

where $\psi(\pi)$ is the conjunction of the literals in the 1-type π. For $\Pi \neq \boldsymbol{\alpha}_\tau$, the size of the formula $\varphi(\Pi)$ is $k2^{k+1} + |\Pi|$. For $\Pi = \boldsymbol{\alpha}_\tau$, the size is $k2^{k+1} + 2^k - 1$. We see that the classes with at most one type missing are tied for the largest formula size.

Using, e.g., standard probabilistic arguments, one can show that for Kripke models of size n, where n is much larger than 2^k, the largest equivalence class is the one realizing all the 1-types. In fact, the largest class will contain "almost all" of the models of size n.

▶ **Proposition 4.** *If n is large with respect to k, then $|M_\Pi| < |M_{\boldsymbol{\alpha}_\tau}|$, for every $\Pi \subset \boldsymbol{\alpha}_\tau$.*

On the other hand, we can show that the equivalence class containing models which realize all the 1-types is one of the most difficult ones to define. To prove this, we will start by introducing a formula size game for MLU[τ].

The formula size game for MLU[τ], denoted $\mathrm{FS}^\tau_{r_0}(\mathcal{A}_0, \mathcal{B}_0)$ has two players: Samson and Delilah. We refer to them as S and D, or he and she, respectively. The game has three parameters: a natural number $r_0 \geq 1$ and two sets of Kripke-models $\mathcal{A}_0$ and $\mathcal{B}_0$. Positions of the game are of the form $(r, \mathcal{A}, \mathcal{B})$ and the starting position is $(r_0, \mathcal{A}_0, \mathcal{B}_0)$.

In each position, S makes a move. The moves available for S in position $(r, \mathcal{A}, \mathcal{B})$ are:

- p-move: S chooses a τ-literal α. The game ends. If $\mathcal{A} \models \alpha$ and $\mathcal{B} \models \neg\alpha$, then S wins. Otherwise D wins. S cannot make this move if he has not made a ♦-move so far.
- $\vee$-move: S chooses $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$ such that $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{A}$ and $r_1, r_2 \geq 1$ such that $r_1 + r_2 + 1 = r$. D chooses whether the next position is $(r_1, \mathcal{A}_1, \mathcal{B})$ or $(r_2, \mathcal{A}_2, \mathcal{B})$.
- $\wedge$-move: The same as a $\vee$-move with the roles of $\mathcal{A}$ and $\mathcal{B}$ switched.
- ♦-move: For every $(\mathfrak{M}, w) \in \mathcal{A}$, S chooses $v \in W$. Let $\mathcal{A}'$ be the set of models $(\mathfrak{M}, v)$ chosen this way. Let $\mathcal{B}' := \{(\mathfrak{M}, v) \mid (\mathfrak{M}, w) \in \mathcal{B} \text{ for some } w \in W, v \in W\}$. The next position of the game is $(r-1, \mathcal{A}', \mathcal{B}')$. S cannot make this move if $r = 1$.
- ■-move: The same as a ♦-move with the roles of $\mathcal{A}$ and $\mathcal{B}$ switched.

▶ **Theorem 5.** *The following statements are equivalent:*

1. *S has a winning strategy in the game $\mathrm{FS}^\tau_r(\mathcal{A}, \mathcal{B})$.*
2. *There is $\varphi \in \mathrm{MLU}[\tau]$ with size at most r such that $\mathcal{A} \models \varphi$ and $\mathcal{B} \models \neg\varphi$.*

Proof. Simple proof by induction. A version for basic modal logic can be found in [7]. ◀

Suppose that $\pi_1, \ldots, \pi_n$, where $n = 2^{|\tau|}$, enumerates all the 1-types over τ. Let $\mathfrak{M}_0$ denote a Kripke model with domain $\{1, \ldots, n\}$ and with the property that for every $1 \leq i \leq n$ the 1-type realized by i is π_i. For every $i \neq j$, we let $\mathfrak{M}_{i,j}$ denote the Kripke model obtained from $\mathfrak{M}_0$ by specifying that the 1-type of i is π_j. We further denote $\mathfrak{M}_i := \mathfrak{M}_{i,1}$ for $2 \leq i \leq n$ and $\mathfrak{M}_1 := \mathfrak{M}_{1,2}$. Each model $\mathfrak{M}_i$ is now missing the type π_i and is otherwise identical to $\mathfrak{M}_0$. We let $\mathcal{A}_0 = \{(\mathfrak{M}_0, 1)\}$ and $\mathcal{B}_0 = \{(\mathfrak{M}_i, 1) \mid 1 \leq i \leq n\}$. We will next show that separating these two sets requires a large MLU[τ] formula.

▶ **Lemma 6.** *D has a winning strategy in the game $\mathrm{FS}^\tau_{k2^{k+1}+2^k-2}(\mathcal{A}_0, \mathcal{B}_0)$.*

Proof. We use the following notation for the set of different underlying models that occur in a set X of pointed models: $\mathrm{Md}(X) = \{\mathfrak{M} \mid (\mathfrak{M}, i) \in \mathcal{A} \text{ for some } i\}$.

We define a measure for a position of the game called hardness. We use hardness as an invariant to show that a position is too hard for S to handle with the available resource r. Let π_i be a type and let $P = (r, \mathcal{A}, \mathcal{B})$ be a position of the game. We define four different kinds of types and the hardness of those types as follows:

1. If no ♦-moves have been made in the game so far, $\mathcal{A} \neq \emptyset$ and $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B})$, then π_i is of kind 1 and $h_i(P) = 2k$.
2. Otherwise, if there are propositionally equivalent $(\mathfrak{M}_0, j) \in \mathcal{A}$ and $(\mathfrak{M}_i, l) \in \mathcal{B}$, then π_i is of kind 2 and $h_i(P) = 2k$.

3. Otherwise, if $(\mathfrak{M}_0, i) \in \mathcal{A}$ and $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B})$, then π_i is of kind 3 and

$$h_i(P) = 2 \cdot |\{(\mathfrak{M}_i, j) \in \mathcal{B} \mid j \neq i, \pi_i \text{ and } \pi_j \text{ differ by exactly one proposition}\}| - 1.$$

4. Otherwise, π_i is of kind 4 and $h_i(P) = 0$.

We further denote the number of types with positive hardness by $\#h^+(P)$ and define the hardness $h(P)$ of the position P as $h(P) = \sum_{1 \leq i \leq n} h_i(P) + \#h^+(P) - 1$.

We will describe the winning strategy for D in terms of maintaining the following two conditions in each position P of the game:

(a) $r < h(P)$,

(b) there is at most one type of kind 3 in position P.

We will show that while these conditions hold, S cannot win. Since the resource r of S will run out eventually, this is a winning strategy for D.

In the starting position P_0 no ♦-moves have been made and $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B}_0)$ for each $1 \leq i \leq n$ so all types π_i are of kind 1 and have $h_i(P_0) = 2k$. Thus condition (b) holds and

$$r = k2^{k+1} + 2^k - 2 < k2^{k+1} + 2^k - 1 = 2k2^k + 2^k - 1 = h(P_0)$$

p-move: In each position P of the game, we have $r \geq 1$ so while $r < h(P)$ holds, we have $h(P) \geq 2$. Using this, we show that any p-move made by S while $r < h(P)$ leads to a win for D. If no ♦-moves have been made, then S cannot make a p-move. If there is a type π_i of kind 2, then there are propositionally equivalent $(\mathfrak{M}_0, j) \in \mathcal{A}$ and $(\mathfrak{M}_i, l) \in \mathcal{B}$ so no literal separates them. If neither of the above hold, then by condition (b), there is a type π_i of kind 3 with $(\mathfrak{M}_0, i) \in \mathcal{A}$ and $(\mathfrak{M}_i, j), (\mathfrak{M}_i, l) \in \mathcal{B}$, where π_j and π_l differ from π_i by exactly one proposition. Again no literal separates $\mathcal{A}$ and $\mathcal{B}$.

$\vee$-move: Similar to the $\wedge$-move case below. Full details in the Appendix.

$\wedge$-move: We show that one of the positions P_1, P_2 satisfies the conditions (a) and (b).

Let $\mathcal{B}_1, \mathcal{B}_2 \subseteq \mathcal{B}$ and $r_1, r_2 \geq 1$ be the choices of S. Let π_i be a type. If π_i is of kind 1, then $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B}_1)$ or $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B}_2)$ so π_i is still of kind 1 and $h_i(P_1) = 2k$ or $h_i(P_2) = 2k$. If π_i is of kind 2 with propositionally equivalent models $(\mathfrak{M}_0, j) \in \mathcal{A}$ and $(\mathfrak{M}_i, l) \in \mathcal{B}$, then $(\mathfrak{M}_i, l) \in \mathcal{B}_1$ or $(\mathfrak{M}_i, l) \in \mathcal{B}_2$ so π_i is still of kind 2 and $h_i(P_1) = 2k$ or $h_i(P_2) = 2k$.

Finally if π_i is a type of kind 3, then S can split the models $(\mathfrak{M}_i, j) \in \mathcal{B}$, where π_i and π_j differ by one proposition, between the sets $\mathcal{B}_1$ and $\mathcal{B}_2$.

Assume that S puts all these models on the same side. Then $h_i(P_1) = h_i(P)$ or $h_i(P_2) = h_i(P)$. Thus $h_i(P_1) + h_i(P_2) \geq h_i(P)$. Additionally $\#h^+(P_1) + \#h^+(P_2) \geq \#h^+(P)$ so

$$\begin{aligned} h(P_1) + h(P_2) &= \sum_{1 \leq i \leq n} h_i(P_1) + \sum_{1 \leq i \leq n} h_i(P_2) + \#h^+(P_1) + \#h^+(P_2) - 2 \\ &\geq \sum_{1 \leq i \leq n} h_i(P) + \#h^+(P) - 1 - 1 = h(P) - 1. \end{aligned}$$

Now $r_1 + r_2 = r - 1 < h(P) - 1 \leq h(P_1) + h(P_2)$ so we have $r_1 < h(P_1)$ or $r_2 < h(P_2)$. Now assume that S splits some models $(\mathfrak{M}_i, j)$ to both sides. Now $h_i(P_1) + h_i(P_2) \geq h_i(P) - 1$. In addition, the type π_i has positive hardness in both positions P_1 and P_2 so $\#h^+(P_1) + \#h^+(P_2) \geq \#h^+(P) + 1$. These two deviations from the above case, that only concern the single type π_i of kind 3, cancel each other out so again $h(P_1) + h(P_2) \geq h(P) - 1$ and therefore $r_1 < h(P_1)$ or $r_2 < h(P_2)$. Finally, all types are of the same kind as in position P so condition (b) holds.

♦-move: Let $(\mathfrak{M}_0, i)$ be a choice of S. For each $j \neq i$ with $\mathfrak{M}_j \in \mathrm{Md}(\mathcal{B})$, we have $(\mathfrak{M}_j, i) \in \mathcal{B}'$ so π_j is a type of kind 2 and $H_j(P') = 2k$. If there are multiple versions of $\mathfrak{M}_0$ in $\mathcal{A}$ and S makes another choice $(\mathfrak{M}_0, l)$, then all types π_j with $\mathfrak{M}_j \in \mathrm{Md}(\mathcal{B})$ work the same

way. If S only chooses $(\mathfrak{M}_0, i)$ and we have $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B})$, then π_i becomes a type of kind 3 with $(\mathfrak{M}_i, j) \in \mathcal{B}'$ for all $j \neq i$ so $h_i(P') = 2k - 1$. We additionally note that by the definition of hardness, $h(P) \leq 2k \cdot |\mathrm{Md}(\mathcal{B})| + |\mathrm{Md}(\mathcal{B})| - 1$. Thus

$$h(P') \geq 2k \cdot |\mathrm{Md}(\mathcal{B})| - 1 + |\mathrm{Md}(\mathcal{B})| - 1 \geq h(P) - 1 > r - 1 = r'$$

and condition (b) is maintained.

■**-move:** For each $\mathfrak{M}_i \in \mathrm{Md}(\mathcal{B})$, S chooses at least one $(\mathfrak{M}_i, l) \in \mathcal{B}'$. Let π_j be the type this model realizes. Now $(\mathfrak{M}_0, j) \in \mathcal{A}'$ realizes the same type, π_i is of kind 2 and $h_i(P') = 2k$. Thus $h(P') = 2k \cdot |\mathrm{Md}(\mathcal{B})| + |\mathrm{Md}(\mathcal{B})| - 1 \geq h(P) > r - 1 = r'$ and condition (b) holds. ◀

We have shown that the largest class $M_{\boldsymbol{\alpha}_\tau}$ requires a formula of size at least $k2^{k+1} + 2^k - 1$ to define. Since any of the classes can be defined via a formula of *precisely* this size, we see that in the case of MLU[τ] the largest class is maximally difficult to define.

▶ **Proposition 7.** *The largest equivalence class $M_{\boldsymbol{\alpha}_\tau}$ of $\equiv_{\mathrm{MLU}[\tau]}$ has maximal* MLU[τ]*-description complexity.*

4 GMLU: Relating entropy and description complexity asymptotically

Fix $\tau = \{p_1, \ldots, p_k\}$ and let $\ell = 2^k$. A Kripke model $\mathfrak{M}$ with universe $W = \{1, \ldots, n\}$ can be described in GMLU[τ] up to isomorphism. Hence the equivalence classes of $\equiv_{\mathrm{GMLU}[\tau]}$, hereafter denoted $\equiv$, over $\mathrm{Mod}_n(\tau)$ are the isomorphism classes. Since $\mathfrak{M}$ can be described up to isomorphism by listing how many times each 1-type is realized, there is a one-to-one correspondence between isomorphism classes and tuples $(n_1, \ldots, n_\ell)$, where $n_1 + \cdots + n_\ell = n$. We will use $[n_1, \ldots, n_\ell]$ to denote the isomorphism class consisting of those Kripke models of size n in which the ith type is realized precisely n_i-times. Note that $|[n_1, \ldots, n_\ell]| = \binom{n}{n_1, \ldots, n_\ell}$.

In this section we show that the expected Boltzmann entropy $\langle H_B \rangle$ is asymptotically $|\tau|$ times the expected GMLU[τ]-description complexity with respect to the distribution $p_\equiv$.

4.1 Expected Boltzmann entropy

In this subsection we will establish that $\langle H_B \rangle \sim |\tau| n$. Using Proposition 1 we get the following alternative asymptotic formula for $\langle H_B \rangle$

$$\begin{aligned}
&\sum_{n_1+\cdots+n_\ell=n} p_\equiv([n_1, \ldots, n_\ell]) \log \binom{n}{n_1, \ldots, n_\ell} \\
&= \sum_{n_1+\cdots+n_\ell=n} p_\equiv([n_1, \ldots, n_\ell]) \left(n \left(\log(n) - \sum_{i=1}^{\ell} \frac{n_i}{n} \log(n_i) \right) + \Theta\big(\log(n)\big) + \sum_{i=1}^{\ell} \Theta(\log(n_i)) \right) \\
&= \left(\sum_{n_1+\cdots+n_\ell=n} p_\equiv([n_1, \ldots, n_\ell]) \left(\sum_{i=1}^{\ell} \frac{n_i}{n} \log\left(\frac{n}{n_i}\right) \right) \right) n \\
&\quad + \Theta\big(\log(n)\big) + \sum_{n_1+\cdots+n_\ell=n} p_\equiv([n_1, \ldots, n_\ell]) \sum_{i=1}^{\ell} \Theta(\log(n_i))
\end{aligned}$$

We will show that

$$\left(\sum_{n_1+\cdots+n_\ell=n} p_\equiv([n_1, \ldots, n_\ell]) \left(\sum_{i=1}^{\ell} \frac{n_i}{n} \log\left(\frac{n}{n_i}\right) \right) \right) n \tag{2}$$

is asymptotically $|\tau|n$, which will of course entail that $\langle H_B \rangle \sim |\tau|n$. Note that

$$\sum_{i=1}^{\ell} \frac{n_i}{n} \log\left(\frac{n}{n_i}\right) \tag{3}$$

is the Shannon entropy of the distribution on $\{1, \ldots, \ell\}$ which assigns to each $1 \leq i \leq \ell$ the weight $\frac{n_i}{n}$. Thus we can use $\log(\ell) = |\tau|$ to bound the formula

$$\sum_{n_1+\cdots+n_\ell=n} p_{\equiv}([n_1, \ldots, n_\ell]) \left(\sum_{i=1}^{\ell} \frac{n_i}{n} \log\left(\frac{n}{n_i}\right) \right) \tag{4}$$

from above. Hence $|\tau|n$ is an upper bound on (2).

We will next bound (4) from below by using Proposition 2. For every $1 \leq i \leq \ell$ and $j \in \mathbb{Z}_+$ we let X_j^i denote a random Bernoulli variable with success probability $2^{-|\tau|}$. Intuitively speaking, X_j^i is an indicator function for the event "the jth element received the ith 1-type". Now, for every $1 \leq i \leq \ell$ and for all $\delta > 0$ the law of large numbers implies that

$$\lim_{n\to\infty} \Pr\left[\left|\tfrac{1}{n}\textstyle\sum_{j=1}^{n} X_j^i - 2^{-|\tau|}\right| < \delta\right] = 1.$$

Thus it follows from the union bound that the following probability

$$\Pr\left[\forall\ 1 \leq i \leq \ell : \left|\frac{1}{n}\sum_{j=1}^{n} X_j^i - 2^{-|\tau|}\right| < \delta\right] \tag{5}$$

approaches 1 as $n \to \infty$. Fix $\delta > 0$. For every n we let I_n^δ denote the following set:

$$\left\{(n_1, \ldots, n_\ell) \mid n_1 + \cdots + n_\ell = n \text{ and } \forall\ 1 \leq i \leq \ell : \left|\frac{n_i}{n} - 2^{-|\tau|}\right| < \delta\right\}.$$

The set I_n^δ includes the tuples $(n_1, \ldots, n_\ell)$, where the numbers add up to n and are very close to each other. The probability result above intuitively means that a randomly chosen tuple is almost always in I_n^δ. Thus, roughly speaking, we only need to consider models, where the points are split between all of the types very evenly.

Let n be large enough so that (5) is larger than $(1-\delta)$. For every $(n_1, \ldots, n_\ell) \in I_n^\delta$ we want to estimate the formula (3) from below. Fix a tuple $(n_1, \ldots, n_\ell) \in I_n^\delta$. Now, for every $1 \leq i \leq \ell$ we have $2^{-|\tau|} - \delta < n_i/n < 2^{-|\tau|} + \delta$, which also entails that $n/n_i > 2^{|\tau|}/(1 + \delta 2^{|\tau|})$. Thus for every $1 \leq i \leq \ell$ we have that

$$2^{|\tau|}(2^{-|\tau|} - \delta) \log\left(\frac{2^{|\tau|}}{(1+\delta 2^{|\tau|})}\right) < \sum_{i=1}^{\ell} \frac{n_i}{n} \log\left(\frac{n}{n_i}\right).$$

Now we can bound the formula (4) from below by

$$2^{|\tau|}(2^{-|\tau|} - \delta) \log\left(\frac{2^{|\tau|}}{(1+\delta 2^{|\tau|})}\right) \cdot \sum_{(n_1,\ldots,n_\ell)\in I_n^\delta} p_{\equiv}([n_1, \ldots, n_\ell]).$$

Notice that the right-hand side expresses the probability that a random τ-model $\mathfrak{A}$ of size n belongs to $[n_1, \ldots, n_\ell]$, for some $(n_1, \ldots, n_\ell) \in I_n^\delta$, which we know is at least $(1-\delta)$, since we chose n to be large enough. Thus we have, for every $\delta > 0$ and n sufficiently large, the following lower bound for the formula (4):

$$f(\delta) := 2^{|\tau|}(2^{-|\tau|} - \delta) \log\left(\frac{2^{|\tau|}}{(1+\delta 2^{|\tau|})}\right) \cdot (1-\delta).$$

Observe that $f(\delta) \to |\tau|$ as $\delta \to 0$. Hence, for every $\varepsilon > 0$ we have that $(1-\varepsilon)|\tau| < f(\delta)$, for sufficiently small δ. Combining this with our upper bound of $|\tau|n$ for (2) one can easily show that (2) is asymptotically $|\tau|n$. This concludes our proof of the following theorem.

▶ **Theorem 8.** $\langle H_B \rangle \sim |\tau|n$.

4.2 Expected description complexity

In this subsection we show that $\langle C \rangle \sim n$. Let M be an equivalence class of $\equiv$. For a 1-type π we denote $|\pi|_M := |\{w \in W \mid (\mathfrak{M}, w) \vDash \pi\}|$, where $\mathfrak{M} \in M$. The number $|\pi|_M$ is the number of points that satisfy the type π in the models of the class M. Since we will focus on a single class M we will omit the subscript in the sequel. Let π_m be the 1-type with the largest number of points in the models of the class M. Let $I := \{1 \leq i \leq 2^{|\tau|} \mid |\pi| \geq 1\}$. The set I consists of the indices of types that are realized in the class M. In this subsection we show that the formula size required to define such a class M is in the order of $\min(n, 2(n - |\pi_m|))$.

For upper bounds, we define a class M via two different formulas, one of them using the largest type π_m defined above:

$$\varphi_1 := \bigwedge_{i \in I} \blacklozenge^{\geq |\pi_i|} \psi(\pi_i)$$

$$\varphi_2 := \blacksquare^{<1}\Big(\bigvee_{i \in I} \psi(\pi_i)\Big) \wedge \bigwedge_{i \in I \setminus \{m\}} \blacklozenge^{\geq |\pi_i|} \psi(\pi_i) \wedge \bigwedge_{i \in I \setminus \{m\}} \blacksquare^{<|\pi_i|+1} \neg\psi(\pi_i)$$

It is easy to verify that $size(\varphi_1) = n + \mathcal{O}_{|\tau|}(1)$ and $size(\varphi_2) = 2(n - |\pi_m|) + \mathcal{O}_{|\tau|}(1)$.

For the lower bounds, we utilize a formula size game $\mathrm{FSC}_r^\tau(\mathcal{A}, \mathcal{B})$ for GMLU$[\tau]$. The rules of the game are the same as in the MLU$[\tau]$-game except the $\blacklozenge$-moves and $\blacksquare$-moves are replaced with the following new moves:

- $\blacklozenge^{\geq d}$-move: S chooses a number $d \in \mathbb{N}$. If $r \leq d$, the game ends and D wins. Otherwise, for every $(\mathfrak{M}, w) \in \mathcal{A}$, S chooses d different points $v \in W$. Let $\mathcal{A}'$ be the set of models $(\mathfrak{M}, v)$ chosen this way. For every $(\mathfrak{M}, w) \in \mathcal{B}$, S chooses $n - d + 1$ different points $v \in W$. Let $\mathcal{B}'$ again be the set of models chosen. The next position of the game is $(r - d, \mathcal{A}', \mathcal{B}')$.
- $\blacksquare^{<d}$-move: The same as a $\blacklozenge^{\geq d}$-move with the roles of $\mathcal{A}$ and $\mathcal{B}$ switched.

The equivalent of Theorem 5 can be proved for this new game in a very similar manner.

For an example of using these new moves, consider a model $\mathfrak{M} \in \mathcal{A}$ with three points, where p is true and another model $\mathfrak{M}' \in \mathcal{B}$, with only two points, where p is true. To separate these models S might make a $\blacklozenge^{\geq d}$-move choosing $d = 3$. S would choose the three points in $\mathfrak{M}$ with p and the $n - 2$ points in $\mathfrak{M}'$ with $\neg p$. Now all three versions of $\mathfrak{M} \in \mathcal{A}'$ have p while all $n - 2$ versions of $\mathfrak{M}' \in \mathcal{B}'$ have $\neg p$ so S wins by making a p-move.

We now define the starting model sets of our formula size game. As before, we assume the domain of the models is $W = \{1, \ldots, n\}$. Let $\mathcal{A}_0 := \{(\mathfrak{M}, 1)\}$, where $\mathfrak{M} \in M$. We additionally assume that the points 1 and 2 of the model are propositionally equivalent. We do not need to fix the model $\mathfrak{M}$ any more precisely but note that there is only one model in the set $\mathcal{A}_0$. Now let $(i, j) \in I \times I$ with $i \neq j$ and let $w \in W$ be the largest number with $(\mathfrak{M}, w) \vDash \pi_i$. The model $\mathfrak{M}_{i \to j}$ has $(\mathfrak{M}_{i \to j}, w) \vDash \pi_j$ and is otherwise identical to $\mathfrak{M}$. In other words, $\mathfrak{M}_{i \to j}$ has one less point of the type π_i and one more of the type π_j compared to $\mathfrak{M}$. We let $\mathcal{B}_0 := \{(\mathfrak{M}_{i \to j}, 1) \mid i, j \in I, i \neq j\}$. There are $|I|^2$ models in the set $\mathcal{B}_0$. Note that all models in $\mathcal{A}_0$ and $\mathcal{B}_0$ have propositionally equivalent starting points.

Let us now consider the formula size game $\mathrm{FSC}^{\tau}_{r_0}(\mathcal{A}_0, \mathcal{B}_0)$. For any position $(r, \mathcal{A}, \mathcal{B})$ of this game, we define a directed graph $\mathcal{G}(\mathcal{A}, \mathcal{B}) := (V, E)$ by setting $V := I$ and $(i, j) \in E$ iff there are propositionally equivalent $(\mathfrak{M}, w) \in \mathcal{A}$ and $(\mathfrak{M}_{i\to j}, v) \in \mathcal{B}$. We call a set $C \subseteq \{i^+, i^- \mid i \in I\}$ a **cover** of $\mathcal{G}(\mathcal{A}, \mathcal{B})$ if for every $(i, j) \in E$ we have $i^+ \in C$ or $j^- \in C$. The cost of a cover C is

$$r(C) := \sum_{i^+ \in C} |\pi_i|_M + \sum_{i^- \in C} |\pi_i|_M.$$

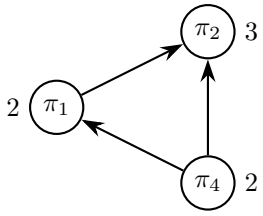

Figure 1 The graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ and the number of points for each type.

We give an example of a graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ and a cover for this graph. Consider the alphabet $\tau = \{p, q\}$ and the class M of models of size 7, where the type $\pi_1 = \{p, q\}$ is realized in two points, the type $\pi_2 = \{\neg p, q\}$ in three points and the type $\pi_4 = \{\neg p, \neg q\}$ in two points. The type $\pi_3 = \{p, \neg q\}$ is not realized in models of M. We assume that $(\mathfrak{M}, w) \in \mathcal{A}$ for some $\mathfrak{M} \in M$ and we have propositionally equivalent $(\mathfrak{M}_{i\to j}, v) \in \mathcal{B}$ for the pairs $(1, 2)$, $(4, 2)$ and $(4, 1)$. The graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is pictured below. The set $C = \{2^-, 4^+\}$ is a cover of $\mathcal{G}(\mathcal{A}, \mathcal{B})$. To see this, note that the inclusion of 4^+ covers all edges from π_4 to other nodes and 2^- covers edges from other nodes to π_2. This covers all edges so C is a cover. The cost of C is $|\pi_2| + |\pi_4| = 3 + 2 = 5$.

We are now ready for the crucial Lemma of this subsection.

▶ **Lemma 9.** *Let $P := (r, \mathcal{A}, \mathcal{B})$ be a position of the game $\mathrm{FSC}^{\tau}_{r_0}(\mathcal{A}_0, \mathcal{B}_0)$ and let $R(P) := \min\{r(C) \mid C \text{ is a cover of } \mathcal{G}(\mathcal{A}, \mathcal{B})\}$. If $r < R(P)$, then D has a winning strategy in the game from the position P.*

Proof. We show that any move S makes either leads to D winning the game immediately or maintains the conditions of the claim given the correct choice by D.

***p*-move:** Since $R(P) > 0$, there are propositionally equivalent pointed models on both sides of the game so clearly D wins if S makes any p-move.

$\vee$-move: Let $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$ and $r_1, r_2 \geq 1$ be the choices of S and let $P_1 = (r_1, \mathcal{A}_1, \mathcal{B})$ and $P_2 = (r_2, \mathcal{A}_2, \mathcal{B})$. For each edge $e = (i, j) \in E$, there are propositionally equivalent models $(\mathfrak{M}, w) \in \mathcal{A}$ and $(\mathfrak{M}_{i\to j}, v) \in \mathcal{B}$. Since $\mathcal{A}_1 \cup \mathcal{A}_2 = \mathcal{A}$, every model $(\mathfrak{M}, w) \in \mathcal{A}$ is in $\mathcal{A}_1$ or $\mathcal{A}_2$ so every edge of the graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is present in at least one of the graphs $\mathcal{G}(\mathcal{A}_1, \mathcal{B})$ and $\mathcal{G}(\mathcal{A}_2, \mathcal{B})$. We claim that $r_1 < R(P_1)$ or $r_2 < R(P_2)$. Assume for contradiction that $r_1 \geq R(P_1)$ and $r_2 \geq R(P_2)$. Then there is a cover C_1 of $\mathcal{G}(\mathcal{A}_1, \mathcal{B})$ with $r(C_1) \leq r_1$ and the same for P_2. Now $C_1 \cup C_2$ is a cover of $\mathcal{G}(\mathcal{A}, \mathcal{B})$. Additionally $r(C_1 \cup C_2) \leq r(C_1) + r(C_2) \leq r_1 + r_2 \leq r$. This means that $R(P) \leq r$, which is a contradiction with the condition $r < R(P)$. Thus D can choose a position that maintains the condition of the claim.

$\wedge$-move: Very similar to the above case with the models in $\mathcal{B}$ split between $\mathcal{B}_1$ and $\mathcal{B}_2$.

$\Diamond^{\geq d}$-move: Let $d \in \mathbb{N}$ be the number chosen by S. For each $(\mathfrak{M}, w) \in \mathcal{A}$, S chooses d different points from the model $\mathfrak{M}$. Let A be the set of all points chosen this way. For each $(\mathfrak{M}_{i\to j}, w) \in \mathcal{B}$, let $B_{i\to j}$ be the set of $n - d + 1$ points chosen by S. Let $\mathrm{tp}_{\mathfrak{M}}(X)$ be the set of types realized by a set X of points in the model $\mathfrak{M}$. We consider the following two cases:

1. The model $\mathfrak{M}$ has at least $d+1$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Let $e=(i,j)\in E$. The model $\mathfrak{M}_{i\to j}$ only differs from $\mathfrak{M}$ by the type of one point so $\mathfrak{M}_{i\to j}$ has at least d points that realize types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Since $|B_{i\to j}|=n-d+1$, there is at least one point in $B_{i\to j}$ with a type from $\mathrm{tp}_{\mathfrak{M}}(A)$. Thus there are propositionally equivalent $(\mathfrak{M},w')\in\mathcal{A}'$ and $(\mathfrak{M}_{i\to j},v')\in\mathcal{B}'$. Thus the edge $e=(i,j)$ is still present in the graph $\mathcal{G}(\mathcal{A}',\mathcal{B}')$ of the following position. This applies for every $e\in E$ so $R(P')=R(P)$.
2. The model $\mathfrak{M}$ has exactly d points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Now A is the set of those d points. We first consider edges $e=(i,j)\in E$ with $\pi_i\notin\mathrm{tp}_{\mathfrak{M}}(A)$ or $\pi_j\in\mathrm{tp}_{\mathfrak{M}}(A)$. For any edge of this kind, the model $\mathfrak{M}_{i\to j}$ has at least d points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$ so at least one of the $n-d+1$ points in $B_{i\to j}$ has a type from $\mathrm{tp}_{\mathfrak{M}}(A)$. As in case 1, this means that all these edges are still present in the graph $\mathcal{G}(\mathcal{A}',\mathcal{B}')$.
   Let us then consider the rest of the edges $e=(i,j)\in E$ with $\pi_i\in\mathrm{tp}_{\mathfrak{M}}(A)$ and $\pi_j\notin\mathrm{tp}_{\mathfrak{M}}(A)$. For an edge of this kind, the model $\mathfrak{M}_{i\to j}$ has only $d-1$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Thus if S chooses the $n-d+1$ points of $B_{i\to j}$ to be exactly the points with types not in $\mathrm{tp}_{\mathfrak{M}}(A)$, then $\mathfrak{M}_{i\to j}$ has no propositionally equivalent counterpart on the other side and the edge e is not present in the graph $\mathcal{G}(\mathcal{A}',\mathcal{B}')$.
   We then consider the condition of the claim in the position $P'=(r-d,\mathcal{A}',\mathcal{B}')$. By the above arguments, the only way S could remove edges when moving from $\mathcal{G}(\mathcal{A},\mathcal{B})$ to $\mathcal{G}(\mathcal{A}',\mathcal{B}')$, was to choose in each version of the model $\mathfrak{M}$ exactly all of the points that satisfy some set $\mathrm{tp}_{\mathfrak{M}}(A)$ of types. Any edge eliminated this way originates from an index i of a type in $\mathrm{tp}_{\mathfrak{M}}(A)$. All of these edges can be covered via the cover $C_A=\{i^+\mid\pi_i\in\mathrm{tp}_{\mathfrak{M}}(A)\}$. The cost of this cover is the total number of points of the model $\mathfrak{M}$ with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Since A contains exactly all points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$, we have $r(C_A)=|A|=d$. Let C' be a cover of $\mathcal{G}(\mathcal{A}',\mathcal{B}')$ with minimal cost so $R(P')=r(C')$. Now $C'\cup C_A$ is a cover of $\mathcal{G}(\mathcal{A},\mathcal{B})$ with cost $R(P')+d$. Thus $r<R(P)\leq R(P')+d$ so $r-d<R(P')$ and the condition of the claim is maintained.

■$^{<d}$-move: Similar to the **♦$^{\geq d}$**-move with $n-d+1$ points chosen from models in $\mathcal{A}$ and d points chosen from models in $\mathcal{B}$. Full details in the Appendix. ◀

By the above Lemma, the formula size required to define a class M of the equivalence $\equiv$ comes down to calculating the minimum cost of a cover.

▶ **Theorem 10.** *Let M be an equivalence class of the relation $\equiv$ and let π be the propositional type with most satisfying points in models in M. If the formula $\varphi\in\mathrm{GMLU}[\tau]$ defines the class M, then φ has size at least* $\min(n,2(n-|\pi|))$.

Proof. Let $s=\min(n,2(n-|\pi|))$. We use the above Lemma to show that D has a winning strategy in the game $\mathrm{FSC}^{\tau}_{s}(\mathcal{A}_0,\mathcal{B}_0)$, thus proving the claim.

It suffices to show that the minimum cost of a cover of $\mathcal{G}(\mathcal{A}_0,\mathcal{B}_0)=(V,E)$ is equal to s. First we see that $\mathcal{G}(\mathcal{A}_0,\mathcal{B}_0)$ is a complete irreflexive directed graph. We begin by noting that $C^+:=\{i^+\mid i\in I\}$ is a cover with cost n and adding any i^- or replacing i^+ with i^- does not reduce the cost. Thus if all indices are used, C^+ is a minimum cost cover. Next, we consider covers C_i, where there is an index $i\in I$ with $\{i^+,i^-\}\cap C_i=\emptyset$. Note that i is the only such index. Indeed, if there were a second such index j, then the edge (i,j) would not be covered. Now, for any $j\in I$, $j\neq i$ we have $j^+\in C_i$ since it is the only way to cover the edge (j,i). In the same way $j^-\in C_i$ since the edge (i,j) must be covered. Thus $C_i=\{j^+,j^-\mid j\in I,j\neq i\}$. The cost of C_i is

$$r(C_i)=\sum_{j^+\in C_i}|\pi_j|+\sum_{j^-\in C_i}|\pi_j|=n-|\pi_i|+n-|\pi_i|=2(n-|\pi_i|).$$

The cost minimal cover of this type is clearly the one where i is the index of the type with the most satisfying points. Thus the minimal cover size is $\min(n,2(n-|\pi|))$. ◀

▶ **Theorem 11.** $\langle C \rangle \sim n$.

Proof. Since $C(M) \leq n + \mathcal{O}_{|\tau|}(1)$ for any equivalence class M, we have $\langle C \rangle \leq n + \mathcal{O}_{|\tau|}(1)$. For the lower bound, recall from the previous section that for any $\delta > 0$ and n sufficiently large we have that

$$\sum_{(n_1,\ldots,n_\ell) \in I_n^\delta} p_{\equiv}([n_1, \ldots, n_\ell]) > (1 - \delta).$$

Observe that if $(n_1, \ldots, n_\ell) \in I_n^\delta$, for δ sufficiently small, then Theorem 10 entails that $C([n_1, \ldots, n_\ell]) \geq n$ as every 1-type is realized less than $n/2$-times. Thus, for any $\delta > 0$ and n sufficiently large, we have $\langle C \rangle \geq (1-\delta)n$. Using these bounds it is easy to show $\langle C \rangle \sim n$. ◀

The desired relation between Boltzmann entropy and description complexity now follows directly from Theorems 8 and 11.

▶ **Corollary 12.** $\langle H_B \rangle \sim |\tau| \langle C \rangle$

5 FO: Expected description complexity for polyadic vocabularies

We saw in the previous section that the ratio of expected Boltzmann entropy of an isomorphism class and its GMLU-description complexity is asymptotically the size of the underlying fixed vocabulary. Given that the main characteristic of GMLU is that it can characterize finite monadic structures up to isomorphism, one might guess that a similar behaviour would extend to FO, which can characterize arbitrary finite structures up to isomorphism. The purpose of this section is to show that surprisingly this is not the case: the expected description complexity grows faster than the expected Boltzmann entropy.

Given a relation symbol R we will use $ar(R)$ to denote its arity. Fix a finite relational vocabulary τ and let $m := \max\{ar(R) \mid R \in \tau\}$. For the rest of this section we will assume that $m \geq 2$. The following result, which fails for unary vocabularies, is established in [4].

▶ **Proposition 13.** *The number of non-isomorphic τ-models of size n is asymptotically $2^{p(n)}/n!$, where $p(n) = \sum_{R \in \tau} n^{ar(R)}$.*

In [12] the authors mention (without a proof) that with high probability, defining a single graph of size n up to isomorphism in FO requires a sentence of size $\Omega\big(\frac{n^2}{\log(n)}\big)$. Here we prove a version of this statement for an arbitrary (but finite) relational vocabulary. For the proof, recall that $\equiv_{\mathrm{FO}[\tau]}$ is over $\mathrm{Mod}_n(\tau)$.

▶ **Theorem 14.** *With high probability we have that $C_{\mathrm{FO}[\tau]}(M) = \Omega\big(\frac{n^m}{\log(n)}\big)$, when the isomorphism class M is selected uniformly at random.*

Proof. The proof is a counting argument: we will show that the ratio between "short" formulas and isomorphism classes of models of size n approaches 0 as n increases. Fix n and consider some $s \geq 2$. We will start by bounding the number of FO[τ]-sentences of size s in which w.l.o.g. only variables from the set $\{x_1, \ldots, x_n\}$, which consists of pairwise distinct variables, occur.

Each FO[τ]-sentence of size s can be viewed as a labeled tree with s nodes, the labels being literals and symbols from the set $\{\wedge, \vee, \exists, \forall\}$. Since a tree with s nodes has $s-1$ edges, the syntax tree of each FO[τ]-sentence of size s can be encoded using

$$f(s) := s + 1 + (s-1)2\log(s) + s\log(N_\tau + 4)$$

bits, where $N_\tau := \sum_{R\in\tau} n^{ar(R)}$ is the number of atomic τ-formulas over $\{x_1,\dots,x_n\}$. Thus there are at most $2^{f(s)}$ sentences of size s. Using this bound we can also bound the number of FO[τ]-sentences of size *at most* s. Indeed, the number of such sentences is at most $\sum_{i=2}^{s} 2^{f(i)} \leq s2^{f(s)} = 2^{f(s)+\log(s)}$.

Observe that $\log(N_\tau+4) \leq d\log(n)$, for some $d>0$ (and n sufficiently large). We now set

$$s = \frac{n^m}{10(m+d)\log(n)},$$

which implies (using very crude bounds such as $s \leq s\log(s)$) that

$$\begin{aligned} f(s)+\log(s) &\leq 5s\log(s) + s\log(N_\tau+4) \\ &\leq 5s(\log(s)+d\log(n)) \\ &\leq 5(m+d)s\log(n) = \frac{n^m}{2} \end{aligned}$$

Thus there are at most $2^{n^m/2}$ sentences of size at most s.

Now the number of non-isomorphic τ-models of size n is asymptotically

$$2^{p(n)}/n! \geq 2^{n^m}/n! \geq 2^{n^m - n\log(n)} = 2^{\left(1-\frac{\log(n)}{n^{m-1}}\right)n^m}$$

for sufficiently large n. Combining these two estimates we have that

$$\frac{2^{n^m/2}}{2^{p(n)}/n!} \leq \frac{2^{n^m/2}}{2^{\left(1-\frac{\log(n)}{n^{m-1}}\right)n^m}} = \left(2^{\frac{\log(n)}{n^{m-1}}-\frac{1}{2}}\right)^{n^m}$$

Since $\log(n)/n^{m-1} \to 0$, by taking n to be sufficiently large we have that $2^{\frac{\log(n)}{n^{m-1}}-\frac{1}{2}} < 1$. Thus with high probability we have that $C_{\mathrm{FO}[\tau]}(M) = \Omega\left(\frac{n^m}{\log(n)}\right)$. ◀

▶ Remark 15. Since for every isomorphism class M we have that $C_{\mathrm{FO}[\tau]}(M) = \mathcal{O}(n^m)$, there is a small gap between this upper bound and the lower bound established in Theorem 14. Even in the case of graphs it seems an open problem to determine the average case FO-description complexity of an isomorphism class, see [12] for more discussion.

Consider now the partition $\equiv_{\mathrm{FO}[\tau]}$ of $\mathrm{Mod}_n(\tau)$. In Appendix A.5 we use Theorem 14 to establish the following result.

▶ **Proposition 16.** *The expected description complexity of $\equiv_{\mathrm{FO}[\tau]}$ grows asymptotically faster than its expected Boltzmann entropy.*

Note that Proposition 16 does not follow immediately from Theorem 14, since there we consider the uniform distribution over the isomorphism classes, while here we need to consider $p_{\equiv_{\mathrm{FO}[\tau]}}$ which a priori could place negligible probabilities on isomorphism classes with high description complexity. However, it follows from the results of [4] that for large n the distribution $p_{\equiv_{\mathrm{FO}[\tau]}}$ is quite close to the uniform distribution.

6 Conclusion

The current paper has demonstrated links between description complexity and entropy. Concretely, we have shown that in MLU, the largest class has maximal Boltzmann entropy, while for GMLU, the expected description complexity is asymptotically equivalent to expected Boltzmann entropy. Corresponding links to Shannon entropy follow from Proposition 3. We also contrast our findings by proving that for first-order logic, description complexity grows asymptotically faster than expected Boltzmann entropy.

In general, our results relate to links between Kolmogorov complexity and entropy, developed this time in the logic-based scenario where relational structures are classified by formulas of different sizes. The results demonstrate, for example, how the size of data classifiers relates to the randomness of the classified data.

In the future we shall expand this study to further logics and more general settings. The overall aim is to investigate the interplay of description complexity and the described classes, involving, e.g., suitable Galois connections between sizes of formulas and classes. Further connections to, e.g., statistical physics should also be investigated.

References

1. Micah Adler and Neil Immerman. An *n!* lower bound on formula size. *ACM Trans. Comput. Log.*, 4(3):296–314, 2003. doi:10.1145/772062.772064.
2. Pablo Barceló, Mikaël Monet, Jorge Pérez, and Bernardo Subercaseaux. Model interpretability through the lens of computational complexity. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL: https://proceedings.neurips.cc/paper/2020/hash/b1adda14824f50ef24ff1c05bb66faf3-Abstract.html.
3. Stephen J. Blundell and Katherine M. Blundell. *Concepts in Thermal Physics*. Oxford University Press, October 2009. doi:10.1093/acprof:oso/9780199562091.001.0001.
4. Ronald Fagin. The number of finite relational structures. *Discret. Math.*, 19(1):17–21, 1977. doi:10.1016/0012-365X(77)90116-9.
5. William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., 1968.
6. Peter Grünwald and Paul M. B. Vitányi. Shannon information and Kolmogorov complexity. *CoRR*, cs.IT/0410002, 2004. doi:10.48550/arXiv.cs/0410002.
7. Lauri Hella and Miikka Vilander. Formula size games for modal logic and μ-calculus. *J. Log. Comput.*, 29(8):1311–1344, 2019. doi:10.1093/logcom/exz025.
8. Reijo Jaakkola, Tomi Janhunen, Antti Kuusisto, Masood Feyzbakhsh Rankooh, and Miikka Vilander. Explainability via short formulas: the case of propositional logic with implementation. In *Joint Proceedings of (HYDRA 2022) and the RCRA Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion*, volume 3281 of *CEUR Workshop Proceedings*, pages 64–77, 2022.
9. Sik K. Leung-Yan-Cheong and Thomas M. Cover. Some equivalences between Shannon entropy and Kolmogorov complexity. *IEEE Trans. Inf. Theory*, 24(3):331–338, 1978. doi:10.1109/TIT.1978.1055891.
10. Ming Li and Paul M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications, 4th Edition*. Texts in Computer Science. Springer, 2019.
11. Michel Loève. *Probability Theory*. Graduate texts in mathematics. Springer, 1963.
12. Oleg Pikhurko and Oleg Verbitsky. Logical complexity of graphs: A survey. In Martin Grohe and Johann A. Makowsky, editors, *Model Theoretic Methods in Finite Combinatorics - AMS-ASL Joint Special Session, Washington, DC, USA, January 5-8, 2009*, volume 558 of *Contemporary Mathematics*, pages 129–180. American Mathematical Society, 2009.
13. Alexander A. Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Comb.*, 10(1):81–93, 1990. doi:10.1007/BF02122698.
14. Andreia Teixeira, Armando Matos, Andre Souto, and Luis Filipe Coelho Antunes. Entropy measures vs. Kolmogorov complexity. *Entropy*, 13(3):595–611, 2011. doi:10.3390/e13030595.
15. Pasko Zupanovic and Domagoj Kuic. Relation between Boltzmann and Gibbs entropy and example with multinomial distribution. *Journal of Physics Communications*, 2:045002, 2018. doi:10.1088/2399-6528/aab7e1.

A Appendix

A.1 Proof of Proposition 3

Letting $\{M_i \mid i \in I\}$ enumerate the equivalence classes of $\equiv$, we have the following chain of identities.

$$\begin{aligned}
&H_S(\equiv) + \langle H_B \rangle \\
&= -\sum_{i\in I} p_\equiv(M_i) \log p_\equiv(M_i) + \sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|) \\
&= -\sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|/|\mathcal{M}|) + \sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|) \\
&= -\sum_{i\in I} p_\equiv(M_i)(\log(|[M_i]_\equiv|) - \log(|\mathcal{M}|)) + \sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|) \\
&= -\sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|) + \sum_{i\in I} p_\equiv(M_i) \log(|\mathcal{M}|) + \sum_{i\in I} p_\equiv(M_i) \log(|[M_i]_\equiv|) \\
&= \log(|\mathcal{M}|) \sum_{i\in I} p_\equiv(M_i) \\
&= \log(|\mathcal{M}|)
\end{aligned}$$

A.2 Proof of Proposition 4

The following standard calculation shows that if n is large enough, then the probability that a random τ-model of size n does not realize all the 1-types is less than $1/2$.

$$\begin{aligned}
&\Pr[\exists \pi : \mathfrak{M} \text{ does not realize } \pi] \le \sum_{\pi} \Pr[\mathfrak{M} \text{ does not realize } \pi] \\
&= \sum_{\pi} \prod_{a\in W} \Pr[a \text{ does not realize } \pi] = \sum_{\pi} \prod_{a\in W} (1 - \Pr[a \text{ does realize } \pi]) \\
&= \sum_{\pi} \prod_{a\in W} (1 - 2^{-k}) = n(1-2^{-k})^n \to 0, \text{ as } n \to \infty
\end{aligned}$$

In the inequality we used union bound while in the first equality we used the fact that the events "a does not realize π", for $a \in W$, are independent.

A.3 Proof of Lemma 6 continued

$\vee$-move: We show that for any $\vee$-move S makes, D can choose one of the following positions P_1, P_2 that satisfies both conditions (a) and (b).

Let $\mathcal{A}_1, \mathcal{A}_2$ and r_1, r_2 be the choices of S. We assume $\mathcal{A}_1, \mathcal{A}_2 \neq \emptyset$. Let π_i be a type. If π_i is of kind 1, then π_i is still of kind 1 in both following positions and $h_i(P) = 2k = h_i(P_1) = h_i(P_2)$, since $\mathcal{B}$ remains unchanged in both positions. If π_i is of kind 2, then there are propositionally equivalent $(\mathfrak{M}_0, j) \in \mathcal{A}$ and $(\mathfrak{M}_i, l) \in \mathcal{B}$. We have $(\mathfrak{M}_0, j) \in \mathcal{A}_1$ or $(\mathfrak{M}_0, j) \in \mathcal{A}_2$ so π_i is still a type of kind 2 in one of the following positions and $h_i(P_1) = 2k$ or $h_i(P_2) = 2k$. Similarly if π_i is of kind 3, then $(\mathfrak{M}_0, i) \in \mathcal{A}_1$ or $(\mathfrak{M}_0, i) \in \mathcal{A}_2$ so π_i remains a type of kind 3 in one of the following positions and $h_i(P_1) = h_i(P)$ or $h_i(P_2) = h_i(P)$. Furthermore, each type with positive hardness in P still has positive hardness in at least one of P_1 or P_2 so $\#h^+(P_1) + \#h^+(P_2) \ge \#h^+(P)$. Thus

$$h(P_1) + h(P_2) = \sum_{1 \le i \le n} h_i(P_1) + \sum_{1 \le i \le n} h_i(P_2) + \#h^+(P_1) + \#h^+(P_2) - 2$$
$$\geq \sum_{1 \le i \le n} h_i(P) + \#h^+(P) - 1 - 1 = h(P) - 1.$$

Now $r_1 + r_2 = r - 1 < h(P) - 1 \leq h(P_1) + h(P_2)$ so we have $r_1 < h(P_1)$ or $r_2 < h(P_2)$. In addition, since all types are of the same kind as in position P, condition (b) still holds.

A.4 Proof of Lemma 9 continued

■$^{<d}$**-move:** Let $d \in \mathbb{N}$ be the number chosen by S. For each $(\mathfrak{M}, w) \in \mathcal{A}$, S chooses $n - d + 1$ different points from the model $\mathfrak{M}$. Let A be the set of all points chosen this way. For each $(\mathfrak{M}_{i\to j}, w) \in \mathcal{B}$, let $B_{i\to j}$ be the set of d points chosen by S. Let $\mathrm{tp}_{\mathfrak{M}}(X)$ be the set of types realized by the set X of points in the model $\mathfrak{M}$. We consider the following two cases:

1. The model $\mathfrak{M}$ has at least $n - d + 2$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Let $e = (i, j) \in E$. The model $\mathfrak{M}_{i\to j}$ only differs from $\mathfrak{M}$ by the type of one point so $\mathfrak{M}_{i\to j}$ has at least $n - d + 1$ points that satisfy types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Since $|B_{i\to j}| = d$, there is at least one point in $B_{i\to j}$ with a type from $\mathrm{tp}_{\mathfrak{M}}(A)$. Thus there are propositionally equivalent $(\mathfrak{M}, w') \in \mathcal{A}'$ and $(\mathfrak{M}_{i\to j}, v') \in \mathcal{B}'$. Thus the edge $e = (i, j)$ is still present in the graph $\mathcal{G}(\mathcal{A}', \mathcal{B}')$ of the following position. This applies for every $e \in E$ so $R(P') = R(P)$.
2. The model $\mathfrak{M}$ has exactly $n - d + 1$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Now A is the set of those $n - d + 1$ points. We first consider edges $e = (i, j) \in E$ with $\pi_i \notin \mathrm{tp}_{\mathfrak{M}}(A)$ or $\pi_j \in \mathrm{tp}_{\mathfrak{M}}(A)$. For any edge of this kind, the model $\mathfrak{M}_{i\to j}$ has at least $n - d + 1$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$ so at least one of the d points in $B_{i\to j}$ has a type from $\mathrm{tp}_{\mathfrak{M}}(A)$. As in case 1, this means that all these edges are still present in the graph $\mathcal{G}(\mathcal{A}', \mathcal{B}')$.
   Let us then consider the rest of the edges $e = (i, j) \in E$ with $\pi_i \in \mathrm{tp}_{\mathfrak{M}}(A)$ and $\pi_j \notin \mathrm{tp}_{\mathfrak{M}}(A)$. For an edge of this kind, the model $\mathfrak{M}_{i\to j}$ has only $n - d$ points with types from $\mathrm{tp}_{\mathfrak{M}}(A)$. Thus if S chooses the d points of $B_{i\to j}$ to be exactly the points with types not in $\mathrm{tp}_{\mathfrak{M}}(A)$, then $\mathfrak{M}_{i\to j}$ has no propositionally equivalent counterpart on the other side and the edge e is not present in the graph $\mathcal{G}(\mathcal{A}', \mathcal{B}')$.
   We then consider the condition of the claim in the position $P' = (r - d, \mathcal{B}', \mathcal{A}')$. We saw above that S can only eliminate an edge $e = (i, j)$ if $\pi_i \in \mathrm{tp}_{\mathfrak{M}}(A)$, $\pi_j \notin \mathrm{tp}_{\mathfrak{M}}(A)$ and $\mathrm{tp}_{\mathfrak{M}_{i\to j}}(B_{i\to j}) \subseteq \mathrm{tp}_{\mathfrak{M}}(W) \setminus \mathrm{tp}_{\mathfrak{M}}(A)$. Thus we denote $\mathrm{tp}(B) := \mathrm{tp}_{\mathfrak{M}}(W) \setminus \mathrm{tp}_{\mathfrak{M}}(A)$. All edges of this kind can be covered via the cover $C_B = \{j^- \mid \pi_j \in \mathrm{tp}(B)\}$. The cost of this cover is the total number of points with types from $\mathrm{tp}(B)$ in the model $\mathfrak{M}$. By the definition of $\mathrm{tp}(B)$ the cost is $r(C_B) = n - |A| = n - (n - d + 1) = d - 1$. Let C' be a cover of $\mathcal{G}(\mathcal{A}', \mathcal{B}')$ with minimal cost so $R(P') = r(C')$. Now $C' \cup C_B$ is a cover of $\mathcal{G}(\mathcal{A}, \mathcal{B})$ with cost $R(P') + d - 1$. Thus $r < R(P) \leq R(P') + d - 1$ so $r - d < R(P')$ and the condition of the claim is maintained.

A.5 Proof of Proposition 16

In this section we use $\equiv$ to denote $\equiv_{\mathrm{FO}[\tau]}$. Our goal is to show that the expected Boltzmann entropy of $\equiv$ grows asymptotically slower than its expected description complexity.

We start by bounding the expected Boltzmann entropy from above. For every equivalence class M of $\equiv$ we have by Proposition 1 that

$$\log(|M|) \leq \log(n!) = n\log(n) - n\log(e) + \Theta(\log(n)),$$

which in turn implies that $\langle H_B \rangle \leq n\log(n) - n\log(e) + \Theta(\log(n))$.

Next we will derive a lower bound on the expected description complexity of $\equiv$. Let c be a constant such that with high probability $C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)$. (Theorem 14 guarantees that such a constant exists.) In [4] it was proved that with high probability a random τ-model is rigid, i.e., it has no non-trivial automorphism. Since the isomorphism class of a rigid τ-model is of size $n!$, we have that with high probability a random member of $\equiv$ has size $n!$. Using a union bound argument we have that

$$\lim_{n\to\infty} \Pr\left[C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right) \text{ and } |M| = n!\right] = 1 \tag{6}$$

In particular, the above probability is at least, say, $1/2$ when n is large enough. In other words, for n large enough, at least half of the isomorphism classes (of models of size n) have size $n!$ and their description complexity is at least $c\left(\frac{n^m}{\log(n)}\right)$.

Now we can bound the expected description complexity from below. First, we have that

$$\sum_M p_{\equiv}(M) C_{\mathrm{FO}[\tau]}(M) \geq \sum_{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)} p_{\equiv}(M) C_{\mathrm{FO}[\tau]}(M)$$

$$\geq c\left(\frac{n^m}{\log(n)}\right) \sum_{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)} p_{\equiv}(M) = c\left(\frac{n^m}{\log(n)}\right) \frac{1}{2^{p(n)}} \sum_{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)} |M|.$$

We want a constant lower bound on the expression

$$\frac{1}{2^{p(n)}} \sum_{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)} |M|$$

which expresses the probability that the isomorphism class of random model of size n has description complexity at least $c\left(\frac{n^m}{\log(n)}\right)$. Using Equation (6), we have for n large enough the following estimates:

$$\frac{1}{2^{p(n)}} \sum_{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right)} |M| \geq \frac{1}{2^{p(n)}} \sum_{\substack{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right) \\ |M| = n!}} |M| = \frac{n!}{2^{p(n)}} \sum_{\substack{C_{\mathrm{FO}[\tau]}(M) \geq c\left(\frac{n^m}{\log(n)}\right) \\ |M| = n!}} 1$$

$$\geq \frac{n!}{2^{p(n)}} (1/2) \frac{2^{p(n)}}{n!} = 1/2.$$

Thus for n large enough we have that $\langle C_{\mathrm{FO}[\tau]} \rangle \geq (1/2) c\left(\frac{n^m}{\log(n)}\right)$, which certainly grows faster than $n\log(n) - n\log(e) + \Theta(\log(n)) \geq \langle H_B \rangle$.