Computational Complexity of Determining Which Statements about Causality Hold in Different Space-Time Models

Vladik Kreinovich
*The University of Texas at El Paso*, vladik@utep.edu

Olga Kosheleva
*The University of Texas at El Paso*, olgak@utep.edu

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep

Part of the Computer Engineering Commons

Recommended Citation

https://scholarworks.utep.edu/cs_techrep/159

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.
Computational Complexity of Determining Which Statements about Causality Hold in Different Space-Time Models

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso, El Paso, TX 79968, USA
vladik@utep.edu, olgak@utep.edu

Abstract

Causality is one of the most fundamental notions of physics. It is therefore important to be able to decide which statements about causality are correct in different models of space-time. In this paper, we analyze the computational complexity of the corresponding decision problems. In particular, we show that:

- for Minkowski space-time, the decision problem is as difficult as Tarski’s decision problem for elementary geometry, while
- for a natural model of primordial space-time, the corresponding decision problem is of the lowest possible complexity among all possible spacetime models.

1 Introduction

Formulation of the problem. Causality is one of the most fundamental notions of physics. It is usually described by a reflexive binary relation $a \leq b$ meaning that an event $a$ can influence an event $b$. It is important to study different properties of the causality relation.

From the logical viewpoint, a natural way to describe such properties is to start with elementary statements of the type $a \leq b$, and to combine these statements by using propositional connectives $\lor$ (“or”), $\land$ (“and”), $\neg$ (“not”), and quantifiers $\forall a$ and $\exists a$ over events. For each space-time model, i.e., for each set of events $E$ with a given causality relation $\leq$, the resulting statements form a first order theory of this causality relation $\leq$.

We face two natural questions. The first is a fundamental question: is the corresponding first order theory decidable? In other words, is there an algorithm that, given a formula, decides whether this formula holds in the given space-time model? For several reasonable space-time models, the corresponding theory is, in principle, decidable. For such models, we face a second more practical
question: how difficult can this decision be? Or, in precise terms, what is the computational complexity of the corresponding decision problem?

Our interest in this topic dates back to the 1970s paper co-authored by one of us with D. Grigoriev [15]. In that paper, we were mainly interested in the first fundamental question: is the corresponding theory decidable? Since the 1970s, a lot of work has been done about the computational complexity of different algorithms and problems, so we decided to revisit the topic of that paper – this time from the computational complexity viewpoint.

What we plan to do. In this paper, we start with the simplest physical model of space-time – the Minkowski space-time. From the fundamental viewpoint, the decision problem is solvable: we can use, e.g., Tarski’s decision procedure for formulas from elementary geometry (see descriptions below) – or, better yet, modern faster decision procedures for these formulas. Our new (somewhat unexpected) result is that the decision problem for the first order theory of causality for the Minkowski space-time is as difficult as the general decision problem of elementary geometry.

We then analyze the general case of abstract space-time models, and we show that for every non-trivial space-time, the decision problem is (at least) PSPACE-hard.

We then turn to another extreme – possible models of “primordial” space-time, i.e., space-time near the beginning of the Universe. We provide some arguments that this space-time should be maximally symmetric, and we analyze the computational complexity of the decision problem for causality theories in this maximally symmetric model. For this model, we describe an exponential-time decision procedure – arguably the fastest possible among different causality-related decision procedures (unless it turns out that P=NP).

Finally, we provide arguments showing, crudely speaking, that there should not be any intermediate drastically different space-time models: a natural transition from the primordial space-time should lead us to causality models similar to the Minkowski one.

2 Case Study: First Order Causality Theory for Minkowski Space-Time

What is Minkowski space-time? Reminder. Let us start with the simplest model of space-time – Minkowski space-time, a 4-dimensional space $E = \mathbb{R}^4$ with the causality relation

\[
\begin{align*}
a = (a_0, a_1, a_2, a_3) &\leq b = (b_0, b_1, b_2, b_3) \iff \\
a = b \land b_0 - a_0 &\geq \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}.
\end{align*}
\]
Comment. In this formula, we assume that the units for measuring time and space are selected in such a way that the speed of light $c$ is equal to 1. If we use other units, then instead of $b_0 - a_0 \geq \ldots$ we should write $c \cdot (b_0 - a_0) \geq \ldots$

Reduction to Tarski’s elementary geometry. The above causality relation can easily be described as a propositional combination of polynomial equalities and inequalities:

$$(a_0 = b_0 \& a_1 = b_1 \& a_2 = b_2 \& a_3 = b_3) \vee$$

$$(b_0 > a_0 \& (b_0 - a_0)^2 \geq (b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2).$$

Thus, all the formulas of the Minkowski causality first order theory can be described in terms of the first order theory of real numbers, in which:

- objects are real numbers,
- elementary formulas are of the type $x = y$, $x \geq 0$, $x = y + z$, and $x = y \cdot z$, and
- general formulas can be obtained by using propositional connectives and quantifiers.

In principle, we can also add more general elementary formulas of the type $P \geq Q$, where $P$ and $Q$ are polynomials with integer coefficients. However, this addition does not seriously change the expressive ability of this language. Indeed, if we parse each polynomial, i.e., represent it as a sequence of elementary arithmetic operations, we can describe such more complex formulas in terms of simpler ones. For example, computing $x_1 + x_2^2$ means that we first compute $z_1 = x_2 \cdot x_2$, and then $y$ as $y = x_1 + z_1$. Thus, the formula $y = x_1 + x_2^2$ can be reformulated as $\exists z_1 (z_1 = x_2 \cdot x_2 \& y = x_1 + z_1)$.

In the late 1940s, A. Tarski has proven that there exists an algorithm which decides whether each formula from this first order theory of real numbers is decidable; see, e.g., [31]. He called this theory elementary geometry because if we represent each point by its coordinates, then practically all the problems of elementary geometry can be described in this language.

Due to the reduction, Tarski’s procedure leads to a decision procedure for the first order causality theory for Minkowski space-time. Since first order causality formulas can be described in terms of elementary geometry,
the existence of Tarski’s algorithm solves the first (fundamental) problem for the Minkowski space-time: namely, by using Tarski’s algorithm, we can decide whether a given first order causality formula is true or not.

What is the computational complexity of this decision problem? Tarski’s original algorithm is very inefficient: its worst-case time complexity grows faster than an exponential function $2^n$ or than any iterations of this function, such as $2(2^n)$, $2(2(2^n))$, etc.

However, later, much more efficient algorithms have been produced. This research started with the fundamental results of M. Ben-Or, D. Kozen, and J. Reif [5] and J. Canny [9]. In particular, J. Canny showed that many problems are in PSPACE – which is more efficient than EXPTIME in practice.

G. E. Collins and H. Hong [10] describe a general algorithm which requires only doubly exponential time $\sim 2(2^n)$. In J. H. Davenport and J. Heintz [11], it was proven that the doubly exponential time is also a lower bound for the general quantifier elimination problem, so exponential time is the worst-case complexity for elementary geometry.

Results from [5] and [9] were generalized by J. Renegar [27]. In particular, J. Renegar produced exponential-time decision procedures for the case when we only have universal quantifiers.

Overall, the algorithms have been improved not only in terms of “theoretical” (asymptotic) complexity, but also in the sense of practical implementations. These efficiency improvements have led to numerous practical applications ranging from control to transportation to epidemiology; see, e.g., S. Ratchan [26] and references therein.

Since, as we have mentioned, for some subclasses of the class of all first order formulas, there are faster algorithms, a natural question appears: If we only restrict ourselves to causality-related formulas, can we have a more efficient decision algorithm than in the general case?

Our first result. Our first result is that the complexity of the decision problem for causality formulas is exactly the same as the complexity of decision problem for formulas of elementary geometry.

Specifically, we have already shown that there is a reduction from the first order theory of Minkowski causality to elementary geometry. This reduction requires time which is linear in the size of the original formula and increases the size of the formula by a constant factor. It turns out that a similar reduction is possible in the other direction:

Proposition 1. There exists a linear-time reduction from the first-order theory of elementary geometry to the first-order theory of Minkowski causality relation.
Conclusion. Thus, the computational complexity of the decision problem for first order formulas of the Minkowski causality relation is the same as for the elementary geometry. As a corollary, we can conclude that for the Minkowski causality, the complexity of the decision problem is doubly exponential.

Related open question. As we have mentioned earlier, several fragments of the first order theory of real numbers have much faster decision procedures. In particular, Canny has shown that entire existential fragment is in PSPACE – hence much more efficient than EXPTIME in practice. It is desirable to find physically interesting classes of causality formulas that correspond to such fragments – e.g., to the existential fragment.

Proof.

1°. A theoretical possibility for this proof comes from the known fact that the Minkowski causality uniquely determines a linear structure on the set of events. Moreover, every mapping $\mathbb{R}^4 \rightarrow \mathbb{R}^4$ which preserves causality is a superposition of Lorentz transformations, rotations, shifts, and dilations (homotheties). This was first proven by A. D. Alexandrov [1, 4] (see also A. D. Alexandrov [2, 3] and E. C. Zeeman [33]).

Historical comment. The transformations described here are the conformal transformations. It has been well known for a long time that these transformations leave the light cones invariant. It has also been known that the only smooth (= differentiable) transformations which preserve causality are the above superpositions. However, the fact that same results holds for arbitrary mappings (without assumptions of smoothness or continuity) was only proven in 1950.

In [32], we have actually used A. D. Alexandrov’s results to provide a physical explanation for standard arithmetic operations (addition, multiplication, etc.) on the numbers field.

We cannot directly use the above proofs because these proofs use sets, i.e., second-order objects which go beyond the first-order descriptions. However, it turns out that the main ideas of these proofs can be modified into a first-order construction. This will enable us to describe formulas of elementary geometry in causality terms.

Comment. In our proof, in addition to the original ideas from Alexandrov’s papers, we also use ideas proposed by W. Benz [6, 7], J. A. Lester [19, 20, 21, 22], and K. Svozil [29, 30].

Specifically, in our proof, in effect, we use the vector space structure on $\mathbb{R}^4$ and the fact that scalar multiplication is defined on it to encode addition and multiplication of real numbers.

2°. A physically important part of the future cone $a^+ \overset{df}{=} \{ b : a \leq b \}$ of a given event $a$ is the boundary of this cone. From the physical viewpoint, the future
cone consists of all the events \( b \) which can be reached from \( a \) by transmissions with speed \( c \) or less (where \( c \) is the speed of light). The boundary of this cone corresponds to events which can be reached only by transmissions traveling exactly with the speed of light. Let us denote the fact that \( b \) is at the boundary of this cone by \( a \overset{\cdot}{\leq} b \).

This relation is a particular case of the causality relation. It can be distinguished by the general causality relation by the following fact:

- when \( a \overset{\cdot}{\leq} b \), then the interval \( \{ d : a \leq d \leq b \} \) is linearly ordered:

- when \( a \leq b \) and \( \neg (a \overset{\cdot}{\leq} b) \), then the corresponding interval is not linearly ordered:

Thus, we can describe the relation \( a \overset{\cdot}{\leq} b \) by an equivalent first-order causality formula:

\[
a \leq b \land \forall d \forall e \left( (a \leq d \leq b \land a \leq e \leq b) \rightarrow (d \leq e \lor e \leq d) \right).
\]

3°. Let \( a \overset{\cdot}{\leq} b \) and \( a \neq b \). A straight line which goes through the points \( a \) and \( b \) is called a null line because along this line, the proper time

\[
\sqrt{(b_0 - a_0)^2 - (b_1 - a_1)^2 - (b_2 - a_2)^2 - (b_3 - a_3)^2}
\]

is equal to 0.

We can describe the fact that an event \( d \in R^4 \) is on the same null line \( ab \) as the events \( a \) and \( b \) as follows: all three events must be in \( \overset{\cdot}{\leq} \) relation to each other, i.e.,

\[
(a \overset{\cdot}{\leq} d \lor d \overset{\cdot}{\leq} a) \land (b \overset{\cdot}{\leq} d \lor d \overset{\cdot}{\leq} b).
\]
Let us denote this formula by \( d \in ab \).

4°. Now, let \( a \leq b \) and \( b \neq a \). We can now describe what it means for a point \( d \) to be in a null hyper-plane – a tangent hyper-plane to the cone \( a^+ \) which passes through the line \( ab \). Geometrically, this means that \( d \) is either on the line \( ab \), or on no null cone starting at any point \( e \) from the null line \( ab \).

Indeed, any point on top of this tangent hyper-plane is covered by the boundary of \( e^+ \) for some \( e \), while points below this hyper-plane are covered by the boundary of \( e^- \) for some \( e \).

In first order terms, the condition that \( d \) is on this hyper-plane can be described as follows:

\[
(\exists e (e \in ab \& (e \leq d \lor d \leq e)) \rightarrow d \in ab.
\]

We will denote this condition by \( d \in H(ab) \).

5°. In the 4-D space, a null hyper-plane \( H(ab) \) is 3-dimensional (i.e., of codimension 1). Thus, the intersection of two null hyper-planes \( H(ab) \) and \( H(ab') \), with \( b' \notin ab \), is of dimension 2. Similarly, the intersection of three null hyper-planes \( H(ab), H(ab'), \) and \( H(ab'') \), with \( b' \notin ab, b'' \notin ab, \) and \( b'' \notin ab' \), is a straight line. Once can check that:

- this straight line is space-like (i.e., on this line, no two points are related by the causality relation), and

- an arbitrary space-like line can be thus obtained.

So, we can define space-like lines in first order terms.

Comment. If, instead of the 4-D Minkowski space, we consider a space of a different dimension \( d \geq 3 \), then we need an intersection of \( d - 1 \) such hyper-planes.

6°. Now, let us fix three events \( a, b, \) and \( b' \) such that \( a \preceq b, a \preceq b', \) and \( b' \notin ab \).

Since \( b' \notin ab \), these three points define a (2-dimensional) plane. A point \( e \) belongs to this plane if and only if there is a space-like line which contains \( e \) and intersects with both null lines \( ab \) and \( ab' \).

\[\begin{array}{c}
  a \\
  b \\
  \hline
  c \\
  b' \\
  \hline
  a
\end{array}\]

This can be easily described in first order terms, so we get a description of this plane. In the following reduction, we will only use events from this plane. (So, all the formulas obtained after reduction should start with \( \forall a \forall b \forall b' \).)

7°. To get a reduction, we need to describe:
• real numbers,
• the relation \(\leq\) on real numbers, and
• two arithmetic operations with these real numbers:
  – addition and
  – multiplication.

Once we have that, any formula from the first order theory of real numbers will be interpreted in causality terms.

8°. Real numbers will be described by events from the line \(ab\).

The event \(a\) represents 0, the event \(b\) represents 1, and an arbitrary number \(\alpha \in \mathbb{R}\) is represented by the corresponding point \(b(\alpha) = a + \alpha \cdot (b - a) \in ab\). In this representation, the standard order \(\leq\) between the real numbers corresponds to the causality relation between the corresponding events:

\[
\alpha \leq \beta \iff b(\alpha) \leq b(\beta).
\]

Because of this definition, every point \(d \in ab\) represents a number – the number \(\alpha\) for which \(b(\alpha) = d\).

This number can given a geometric interpretation if we formally introduce on \(\mathbb{R}^4\) an Euclidean metric \(\rho\) in which the distance between \(a\) and \(b\) is exactly 1: \(\rho(a, b) = 1\). In this metric, the number \(\alpha\) corresponding to the event \(d\) is simply equal to the (signed) distance \(\rho(a, d)\) between the event \(d\) and the fixed point \(a\).

9°. To describe addition, we must be able to shift “intervals” along the null line \(ab\). This can be done as follows. By using two parallel null lines and two parallel space-like lines, we can form parallelepipeds and thus, make sure that on a parallel null line, we have an interval \(a'd'\) which is of the same size as the original interval \(ad\):
Then, by using another parallelepiped, we can move the interval $a'd'$ back into the original null line $ab$, this time into a different place, as $a''d'':$

\[ a''d'' \]

By using this construction, we can arbitrarily shift any interval along the null line $ab$.

10°. Now, addition $z = x + x'$ can be described as follows. In our representation, both real numbers $x$ and $x'$ are represented by points $d \in ab$ and $d' \in ab$ for which $\rho(a, d) = x$ and $\rho(a, d') = x'$.

\[ a \quad x' \quad d' \]

By using a construction described in Part 9 of this proof, we then shift an interval $ad'$ onto the same null line $ab$ in such a way that the shifted interval $de$ starts with the point $d$.

\[ a \quad x' \quad d' \quad e \]

Since the length $\rho(a, d)$ of the interval $ad$ is equal to $x$ and the length $\rho(d, e)$ of the interval $de$ on the same line $ab$ is equal to $x'$, the length $\rho(a, e)$ of the combined interval $ae$ is thus equal to $x + x'$.

\[ a \quad x + x' \quad e \]

So, the event $e$ is the desired representation for the sum $x + x'$. 
Comment. Strictly speaking, the above procedure only works for adding positive real numbers, so we need additional constructions corresponding to different signs. The signs are easy to describe in first order terms, since:

• non-negative numbers are represented by events \( d \geq a \), while
• non-positive numbers are represented by events \( d \leq a \).

11°. To describe multiplication, we need:

• to have an alternative representation of real numbers, on the null line \( ab' \), and
• to be able to move numbers between the original and the new (alternative) representations.

In the alternative representation, we represent real numbers by points from the null line \( ab' \), with 0 represented by \( a \), 1 represented by \( b' \), and an arbitrary number \( \alpha \in \mathbb{R} \) by a point \( b'(\alpha) \triangleq a + \alpha \cdot (b' - a) \in ab' \).

If we have a point \( d \in ab \) (\( d \neq a \)) representing a real number, then we can describe the corresponding point \( d' \in ab' \) representing the same real number as follows: it is the only point \( d' \in ab' \) for which the space-like line \( dd' \) is parallel to the space-like line \( bb' \) (since we restricted ourselves to the 2-dimensional plane, parallel simply means no common points).

For an event \( b \) representing number 1, the corresponding point is a point \( b' \), at a distance \( \rho_0 \triangleq \rho(a, b') \). In general:

• If we have an event \( d \in ab \) which represent a number \( x = \rho(a, d) \), then in the alternative representation, this same number is represented by an event \( d' \in ab' \) for which \( \rho(a, d') = x \cdot \rho_0 \).
Vice versa, an arbitrary event \( d' \in ab' \) represents a real number \( x = \frac{\rho(a, d')}{\rho_0} \). In the original representation, this same number is represented by an event \( d \) for which \( \rho(a, d) = x = \frac{\rho(a, d')}{\rho_0} \).

12°. We are now ready to show how multiplication can be described.

Similarly to addition, it is sufficient to describe multiplication of positive real numbers; other cases can be handled in a similar manner. In the case of positive real numbers, we need to multiply two given real numbers \( x > 0 \) and \( x' > 0 \). According to our representation of real numbers,

- the value \( x \) is represented by an event \( d \) for which the distance \( \rho(a, d) \) is equal to \( x \), and
- the value \( x' \) is represented by an event \( d' \) for which the distance \( \rho(a, d') \) is equal to \( x' \).

By using the construction from Part 11 of this proof, we can find an event \( d'' \in ab' \) which represent the same number \( x' \) as the event \( d' \in ab \).

There exists a space-line line going through the points \( b \) and \( d'' \). We then find a line which is parallel to \( bd'' \) and which goes through \( d \). This line intersects with a null-line \( ab' \) at some point \( e' \).

Since the lines \( bd'' \) and \( de' \) are parallel, the triangles \( abd' \) and \( ade' \) are similar. Thus, we have the following relation between the lengths of the corresponding sides:

\[
\frac{\rho(a, e')}{\rho(a, d')} = \frac{\rho(a, d)}{\rho(a, b)}
\]
By our choice of metric, \( \rho(a, b) = 1 \). By our choice of \( d \) as a representation of the value \( x \), we have \( \rho(a, d) = x \); so, the ratio in the right-hand side is equal to \( x/1 = x \). From \( \frac{\rho(a, e')}{\rho(a, d'')} = x \), we conclude that \( \rho(a, e') = x \cdot \rho(a, d') \).

The event \( d'\) represents the value \( x' \) in the alternative representation, so \( \rho(a, d'') = x' \cdot \rho_0 \). Thus, we conclude that \( \rho(a, e') = x \cdot (x' \cdot \rho_0) = (x \cdot x') \cdot \rho_0 \). By definition of the alternative representation, this event represents a number \( \frac{\rho(a, e')}{\rho_0} = x \cdot x' \).

By applying to the event \( e' \in ab' \) the construction from Part 11, we get a point \( e \in ab \) which represents the same number \( x \cdot x' \) in the original representation.

13°. Summarizing: we can represent real numbers, inequality between real numbers, addition, and multiplication. We thus have the desired representation of the elementary geometry in the first order theory of Minkowski causality. The proposition is proven.

Comment. In our definition and in the proof, we mainly considered 4-dimensional space-time. However, one can easily check that our proof applies to the Minkowski space-time model of an arbitrary dimension \( \geq 3 \).

3 General Case

Discussion. In the previous section, we have shown that for the Minkowski space-time, the decision problem has double exponential complexity. A natural question is: can this decision problem be much easier for other space-time models?

In this section, we will show that for every space-time model which is non-degenerate in some reasonable sense, the decision problem of the corresponding causality theory is PSPACE-hard.

This means that (unless PSPACE=\( P \)), the decision problem always requires at least exponential time. In the following section, we will provide an example of a physically reasonable space-time model for which the exponential time is sufficient.
PSPACE: reminder for physicist readers. The main result of this section is formulated in terms of the complexity class PSPACE. This is a class which is very familiar to every computer scientist. However, our motivations come from physics, and we therefore hope that some physicists will read this paper as well; for these readers, we provide a brief description of this complexity class. Computer science readers can skip directly to the subsection titled “Our second result”.

It is well known that some algorithms are practically useful, while some other algorithms are computationally useless: even for reasonable size inputs, they require time which exceeds the number of particles in the Universe. This distinction is very different to formalize. Usually:

- algorithms for which the computation time $t_A(x)$ is bounded by some polynomial $P(n)$ of the length $n = \text{len}(x)$ of the input (e.g., linear-time, quadratic-time, etc.) are practically useful, while
- for practically useless algorithms, the computation time grows with the size of the input much faster than a polynomial.

In view of this empirical fact, in theoretical computer science, algorithms are usually considered feasible if their running time is bounded by a polynomial of $n$. The class of problems which can be solved in polynomial time is usually denoted by $P$; see, e.g., [24].

Not all practically useful problems can be solved in polynomial time. To describe such problems, researchers have defined several more general classes of problems. One of the most well known classes is the class NP. By definition, this class consists of all the problems which can be solved in non-deterministic polynomial time – meaning that if we have a guess, we can check, in polynomial time, whether this guess is a solution to our problem.

Most computer scientists believe that $\text{NP} \neq \text{P}$, i.e., that some problems from the class NP cannot be solved in polynomial time. However, this inequality has not been proven, it is still an open problem. What is known is that some problems are $\text{NP}$-hard, i.e., any problem from the class NP can be reduced to each of these problems in polynomial time. One of such NP-hard problems is the problem SAT of propositional satisfiability: given a propositional formula $F$, i.e., a formula obtained from Boolean (yes-no) variables $x_1, \ldots, x_n$ by using $\&$, $\lor$, and $\neg$, check whether there exist values $x_1, \ldots, x_n$ which make this formula true.

NP-hardness of SAT means that if $\text{NP} \neq \text{P}$ (i.e., if at least one problem from the class NP cannot be solved in polynomial time), then SAT also cannot be solved in polynomial time. In other words, SAT is the hardest of the problems from this class.

It is known that all the problems from the class NP can be solved in exponential time. Indeed, for a problem of size $n$, there are $\leq a^n$ possible guesses, where $a$ is the size of the corresponding alphabet, so we can simply try all these guesses one by one.
This sequential testing requires only a polynomial amount of space. There is a larger class of problems PSPACE, the class of all the problems which can be solved by using polynomially many computer cells. Problems from the class PSPACE can also be solved in exponential time.

In the class PSPACE, there are also problems (called PSPACE-hard) to which all other problems from this class can be reduced in polynomial time. So, if a problem is proven to be PSPACE-hard, this means that – unless PSPACE is equal to P (or to some sub-exponential class) – we cannot solve this problem is less than exponential time.

One example of such PSPACE-hard problems is quantified satisfiability (QBF). In QBF, we consider formulas which are obtained from the Boolean variables by using propositional connectives and quantifiers $\forall x$ and $\exists x$ over possible values of the corresponding Boolean variables [24].

Our second result. In view of the above reminder, proving that a problem is PSPACE-hard means, in effect, that it cannot be solved faster than in exponential time. The following simple result shows that for every non-trivial space-time model, the causality decision problem is PSPACE-hard – and thus, in effect, cannot be solved faster than in exponential time.

As we have mentioned in the Introduction, we are only interested in the causality relation. Therefore, we arrive at the following definition of a space-time model.

Definition 1. By a space-time model, we mean a pair $(E, \leq)$, where $E$ is a set and $\leq$ is a reflexive binary relation on $E$.

Definition 2. We call a space-time model non-trivial if there exist two events $a$ and $b$ for which $a \nleq b$.

Comment. The only space-time models which we exclude are the ones for which $a \leq b$ for all $a$ and $b$. For such models, the decision problem is trivial.

Proposition 2. For every non-trivial space-time model $(E, \leq)$, the decision problem for the first order causality theory of the corresponding relation $\leq$ is PSPACE-hard.

Proof. To prove that this causality-related problem is PSPACE-hard, we will reduce QBF (a known PSPACE-hard problem) to this problem. By definition, the fact that QBF is PSPACE-hard means that every problem from the class PSPACE can be reduced to QBF; since, as well will show, QBF can be reduced to the causality problem, this means that every problem from the class PSPACE can be reduced to the causality problem – i.e., that the causality problem is indeed PSPACE-hard.
In the desired reduction of QBF to the causality problem, we assign, to each Boolean variable $x_i$, a pair of event-valued variables $a_i$ and $b_i$. In the formula $F$ from QBF,

- each occurrence of the Boolean variable $x_i$ is then replaced with the formula $a_i \leq b_i$,
- propositional connectives remain intact, and
- each quantifier $\forall x_i$ is replaced with $\forall a_i \forall b_i$, and each quantifier $\exists x_i$ is replaced with $\exists a_i \exists b_i$.

For every space-time model, we can select $b_i = a_i$ and get $x_i \overset{\text{def}}{=} (a_i \leq b_i)$ to be true. Since we assumed that the space-time is non-trivial, there also exist values $a_i$ and $b_i$ for which $a_i \not\leq b_i$. Thus, each of the corresponding variables $x_i$ can take exactly two possible values: “true” and “false”.

So, the resulting formula of the causality theory is true if and only if the original formula from QBF was true. The proposition is proven.

4 Second Case Study: A Reasonable Model of Primordial Space-Time

Main physical idea. In this section, we go to the other extreme: from the Minkowski space-time which reasonably describes current space-time (at least locally), we go to the attempts to describe space-time near the beginning of the Universe.

Due to the high degree of symmetry observed today in the cosmic microwave background, cosmologists believe that the early universe was homogenous and isotropic; see, e.g., [23]. Several physicists have presented arguments that the primordial space-time should have even more symmetries: that if we have a finite set of events $a_1, \ldots, a_m$, and in principle, from the viewpoint of general causality theory, it is possible to have a new event $a$ which a certain relation to $a_i$ (e.g., $a \leq a_1$, $a \not\leq a_2$, etc.), then such a “theoretically possible” event $a$ actually exists in the primordial space-time; see, e.g., M. Droste [12].

Here, “theoretically possible” means that there exists another space-time model in which there are points $a'_i$ which are related to each other in the same way as the points $a_i$, and in which there is another point $a'$ with the desired relation to $a_i$.

Similarly, if it is possible to have two events $a$ and $b$ related to $a_i$ and to each other in a certain way, then such events must happen in the primordial space-time.

Let us describe this idea in precise terms. In this description, we will assume that the causal relation $\leq$ is a partial order.
Definition 3.

- A 1-1 onto mapping \( f : S \to S' \) between partially ordered sets is called an isomorphism if for every \( a, b \in S \), we have \( a \leq b \leftrightarrow f(a) \leq f(b) \).

- We say that an partially ordered set \((E, \leq)\) is universal if for every finite subset \( S \subseteq E \) and for every finite partially ordered set \( F' \), if \( S \) is isomorphic to a set \( S' \subseteq F' \), then this isomorphism can be extended to an isomorphism between the whole set \( F' \) and some superset \( F \supseteq S \) of the set \( S \).

Due to fundamental reasons dictated by considerations of quantum gravity, we assume that the space-time is countable; this assumption, first made in R. Sorkin’s theory of causal sets [28], is used by M. Droste in [12].

Description of the resulting primordial space-time model. It is known (see, e.g., R. Fraissé [13, 14] and W. Hodges [16]) that there exists a countable universal partially ordered set. It is also known that all such sets are isomorphic, and that these sets have a large symmetry groups: every isomorphism \( f : S \to S' \) between two subsets of \( E \) can be extended to an isomorphism of the entire space \( E \).

This means, in particular, that for every two events \( a, a' \in E \), there exists an isomorphism which maps \( a \) into \( a' \), i.e., that this space-time model is homogeneous. This also means that if \( a_1 < a_2 \) and \( a'_1 < a'_2 \), then there exists an isomorphism of \( E \) which maps \( a_1 \) into \( a'_1 \) and \( a_2 \) into \( a'_2 \), etc.

Comment. A similar universal construction for discrete space-time models has been described by M. Droste [12].

Computational complexity of the decision problem for the primordial space-time. In the previous section, we have proven that for every non-trivial space-time model, the corresponding decision problem is PSPACE-hard, and we have mentioned that, in effect, this means that we cannot solve this problem faster than in exponential time.

We have also shown that for a seemingly simple Minkowski space-time, the decision problem is much more difficult: it requires doubly exponential time. Let us now that for the (seemingly less trivial) primordial space-time, the decision problem can be actually solved in exponential time – i.e., as fast as possible.

A possible explanation lies in the high symmetry of the primordial space-time: usually, symmetry helps in computations, and here we have as much symmetry as potentially possible.

Proposition 3. There exists a polynomial \( P(n) \) and an \( O \left( P(n) \cdot 2^{n^2} \right) \) time algorithm which, given a first order causality formula \( F_0 \), decides whether this formula holds in the universal (primordial) space-time model.
Proof.

1°. By definition, every formula from the first order causality theory is obtained from the basic formulas of the type $a \leq b$ by using propositional connectives and quantifiers. For each given formula $F_0$, we can thus trace its construction to the basic formulas and thus, get a sequence of (sub)formulas such that

- the first subformulas are elementary formulas of the type $a \leq b$;
- each of the following subformulas is obtained from the previous one(s) by using a propositional connective or a quantifier; and
- at the end, we get the desired formula.

At the end, we get a closed formula (without free variables), i.e., a formula which can be true or false. On the intermediate steps of this construction, however, we can have formulas like $a \leq b$ which contain free variables. Such formulas are true or false depending on the values of these variables.

2°. In our algorithm, to determine whether the final formula holds or not in the universal partially ordered set $E$, we will follow the corresponding sequence of subformulas. We will show that for each of these subformulas $F(a_1, \ldots, a_k)$, its truth in $E$ depends only on the $\leq$-relation between the values $a_i$. Thus, all the information about the truth of each such subformula for different elements $a_1, \ldots, a_k \in E$, can be completely determined by listing all the (partial) orders on the $k$-element set $\{a_1, \ldots, a_k\}$ for which this formula is true in $E$.

This listing can be equivalently described by a propositional formula of the type

$$((a_1 \leq a_2) \& (a_2 \not\leq a_1) \& \ldots) \lor (\ldots) \lor \ldots$$

in which each conjunction $(a_1 \leq a_2) \& (a_2 \not\leq a_1) \& \ldots$ corresponds to one such partial order.

In our algorithm, we follow the construction of the formula $F_0$ step by step, and along the way, construct the corresponding lists for all the subformulas $F$.

- If the corresponding subformula $F$ is closed, then we will simply generate the information on whether it is true or not in the set $E$.
- If the subformula $F(a_1, \ldots, a_k)$ has free variables $a_1, \ldots, a_k$, then we will construct a list of all partial orders on the $k$-element set $\{a_1, \ldots, a_k\}$ for which this subformula is true in $E$.

At the end of this construction, when we get to the original closed formula, we will have its truth value.

Let us show how this works.

3°. For the elementary formula $a_1 \leq a_2$, we simply construct the set consisting of two partial orders:

- an order in which $a_1 \leq a_2$ and $a_2 \not\leq a_1$; and
• an order in which \( a_1 \leq a_2 \) and \( a_2 \leq a_1 \).

4°. Let us first consider the case when the next formula \( F(a_1, \ldots, a_k) \) is the negation of the previous one: \( F(a_1, \ldots, a_k) \overset{\text{def}}{=} \neg G(a_1, \ldots, a_k) \). In this case, for the formula \( G(a_1, \ldots, a_k) \) with \( k \) free variables for which we already have a list of all partial orders on the set \( \{a_1, \ldots, a_k\} \) for which this formula \( G \) is true in \( E \).

We need to construct a similar list for the negation \( \neg G(a_1, \ldots, a_k) \). For that, we simply generate all the possible partial orders of the set \( \{a_1, \ldots, a_k\} \) and then dismiss those from the given list (i.e., those for which \( G \) is true).

5°. Before we explain how to proceed with other propositional formulas and quantifiers, let us describe the following auxiliary operation.

In this auxiliary operation, we add a “dummy” variable \( a_{k+1} \) to the formula \( F(a_1, \ldots, a_k) \). In other words, we simply add this new variable to the list of variables, but the truth value of the formula \( F(a_1, \ldots, a_k) \) is not affected by this new variable.

From the viewpoint of logic, adding a dummy variable is a trivial construction, because it does not change any truth values. However, from the viewpoint of our proof, it is non-trivial:

• for the original formula, we have a list of all the partial orders on the \( k \)-element set \( \{a_1, \ldots, a_k\} \) which satisfy the property \( F(a_1, \ldots, a_k) \);

• for the new formula with a dummy variable added, we need the list of all partial orders on the \( (k+1) \)-element set \( \{a_1, \ldots, a_k, a_{k+1}\} \) which satisfy the property \( F(a_1, \ldots, a_k) \).

This new list can be constructed as follows:

• we consider all possible partial orders on the \( (k+1) \)-element set \( \{a_1, \ldots, a_k, a_{k+1}\} \);

• for each of these orders, we delete \( a_{k+1} \); if the resulting order on the \( k \)-element set is in the originally given list, we add the corresponding order on the \( (k+1) \)-element set to the new list.

In this manner, we can add any number of such dummy variables.

6°. Let us now describe how we can handle the case of a propositional connective \( F = G \& H \) or \( F = G \lor H \).

We assume that we already have lists corresponding to the subformulas \( G \) and \( H \), and we need to construct a new list corresponding to the formula \( F \).

First, if \( G \) has free variables which are not in \( H \), then we add these variables to \( H \) as dummy variables. Similarly, if \( H \) has free variables which are not in \( G \), then we add these variables to \( G \) as dummy variables.

As a result, we can assume that the formulas \( G \) and \( H \) have the exact same set of variables. Then, we proceed as follows:
• The formula $G \& H$ is true for a partial order if and only if both formulas $G$ and $H$ are true. Thus, the list corresponding to $G \& H$ is the intersection of the lists corresponding to $G$ and $H$.

• Similarly, the formula $G \lor H$ is true for a partial order if and only if one of the formulas $G$ and $H$ are true. Thus, the list corresponding to $G \lor H$ is the union of the lists corresponding to $G$ and $H$.

7°. To complete the description of our construction, we need to cover the quantifiers case, when the subformula $F(a_1, \ldots, a_k)$ has the form $\exists a_{k+1} G(a_1, \ldots, a_k, a_{k+1})$ or $\forall a_{k+1} G(a_1, \ldots, a_k, a_{k+1})$.

Due to de Morgan’s duality, the universal quantifier can be described in terms of the existential one:

$$\forall a_{k+1} G(a_1, \ldots, a_k, a_{k+1}) \iff \neg \exists a_{k+1} \neg G(a_1, \ldots, a_k, a_{k+1}).$$

Since we already know how to handle the negation, it is sufficient to describe the construction for the existential quantifier.

We assume that we already have a list of all the partial orders on the $(k+1)$-element set $\{a_1, \ldots, a_k, a_{k+1}\}$ for which the property $G(a_1, \ldots, a_k, a_{k+1})$ is true in $E$.

As we have mentioned earlier, the existence of such a list means, in effect, that the formula $\tilde{G}(a_1, \ldots, a_k, a_{k+1})$ is equivalent to a propositional formula of the type

$$(a_1 \leq a_2) \& (a_2 \not\leq a_1) \& \ldots) \lor (\ldots) \lor \ldots$$

which describes all the corresponding partial orders.

According to the above definition of the universal partially ordered set, if there is an element $a_{k+1}$ with the desired relation $\tilde{G}(a_1, \ldots, a_k, a_{k+1})$ in some partial ordered set, then such an element can be found in this universal set $E$ as well.

Thus, the formula $\exists a_{k+1} \tilde{G}(a_1, \ldots, a_k, a_{k+1})$ holds for the elements $(a_1, \ldots, a_k) \in E$ if and only if there exists a partial order on a $(k+1)$-element set $\{a_1, \ldots, a_k, a_{k+1}\}$ which extends the given partial order on the set $\{a_1, \ldots, a_k\}$ and for which the formula $\tilde{G}(a_1, \ldots, a_k)$ holds.

So, to construct the desired list for the formula $\exists a_{k+1} G(a_1, \ldots, a_k, a_{k+1})$, it is sufficient to consider all partial orders of the $(k+1)$-order set $\{a_1, \ldots, a_k, a_{k+1}\}$ from the $G$-list, and delete the element $a_{k+1}$ from each of these orders.

8°. We have shown that the above construction indeed leads to the desired lists – hence, at the end, this construction leads to the truth value of the original formula $F_0$. To complete the proof, we must show that this algorithm indeed finishes its computations in time $O \left( P(n) \cdot 2^{n^2} \right)$ for some polynomial $P(n)$.

Indeed, for each subformula $F$, the most time-consuming part is the enumeration of all the partial orders on a $k$-element set $\{a_1, \ldots, a_k\}$. To describe each such partial order, we must know, for each pair $(i, j)$ with $i = 1, 2, \ldots, k$
and $j = 1, 2, \ldots, k$, where $a_i \leq a_j$ or not. There are $k^2$ such pairs, so we need $k^2$ bits to describe an arbitrary order. Enumerating all such orders thus requires as many steps as there are sequences with $k^2$ bits, i.e., $2^{k^2}$ steps.

For each such sequence of bits, we also need to check whether the resulting relation is a partial order; this requires polynomial number of steps, so overall, we need $\leq p(k) \cdot 2^{k^2}$ steps for some polynomial $p(k)$.

For our algorithm, the input is the formula $F_0$ for which we want to know whether this formula holds or not. So, the length $n$ of the input is the length of the given formula $F_0$. On each intermediate step, the number of variables $k$ is bounded by the length $n$ of the formula: $k \leq n$. Thus, the computational complexity of each subformula is bounded by $p(n) \cdot 2^{n^2}$.

The number of subformulas is also bounded by the length $n$ of the input, so the overall computation time is bounded by $n \cdot p(n) \cdot 2^{n^2}$, where $P(n) \overset{\text{def}}{=} n \cdot p(n)$ is a polynomial. The proposition is proven.

5 Discussion: Is There Anything in Between?

Formulation of the problem. In the previous sections, we have analyzed two space-time models:

- the Minkowski space-time which (reasonably accurately) describes the current space-time, and
- the primordial space-time which seems to provide a reasonable description of the causality relation at the very beginning of the Universe.

A natural question is: is there anything radically different in between?

In this short section, we will provide an argument that probably we should not expect anything drastically different.

Symmetry violation as a physically natural way from the primordial space-time to current (locally Minkowski-like) one. From the physical viewpoint, one of the important features of different physical processes is their symmetries.

We have already mentioned that the universal partially ordered set has a lot of symmetries. These symmetries can be described in the following way.

**Definition 4.** A mapping $f : M \rightarrow M'$ between ordered sets is called a homomorphism if for every $a, b \in M$ for which $a \leq b$, we have $f(a) \leq f(b)$.

**Definition 5.** [17] Let $n > 0$ be an integer. We say that an ordered set $(M, \leq)$ is $n$-structurally homogeneous if for every isomorphism $f : S \rightarrow S'$ between two subsets $S, S' \subseteq M$ with $n$ or fewer elements, and for every $m \notin S$, there exists an element $m' \notin S'$ such that a mapping $f'$, extended to $S \cup \{m\}$ by setting $f'(m) = m'$, is a homomorphism $f' : S \cup \{m\} \rightarrow S' \cup \{m'\}$.
The universal partially order set is $n$-structurally homogeneous for every $n$. Since our real space-time is different, this means that these symmetries become violated, i.e., that, instead of $n$-homogeneity for all $n$, we only have structural homogeneity for some $n$.

In [17], we have shown that, in effect, the standard axioms for causality as presented in H. Busemann [8], E. H. Kronheimer and R. Penrose [18], and R. I. Pimenov [25], crudely speaking, correspond to 3-structural homogeneity.

Is there anything in between? To answer this questions, let us first briefly summarized what we have just discussed.

- The primordial space-time is $n$-structurally homogeneous for all $n$.
- After a while, these symmetries are violated, so we only have $n$-structural homogeneity for some $n$.
- The current space-time corresponds to $n = 3$.

According to this logic, we can, in principle, have intermediate stages corresponding to $n > 3$, e.g., to $n = 4$, $n = 5$, etc. Let us show that, in some reasonable sense, homogeneity for $n = 4$ implies homogeneity for all larger $n$. Thus, there is indeed nothing drastically different in between.

To explain this result in precise terms, we first need to clarify our brief description of the main result from [17].

First clarification: we restrict ourselves to kinematic causality. The first clarification is related to the fact that in the definition of a universal partial ordered set, we only considered countably many events. The justification for the possibility to restrict ourselves to countably many events came from the fact that measurements are never 100% accurate. So, a countable dense set of events is sufficient to describe all possible measurement results.

Therefore, to make a reasonable comparison between our space-time and this primordial space-time, we must restrict ourselves to a countable dense set of events for our space-time as well. In general, if we pick kind of “random” dense set, then for each even $a$, the “probability” to pick another events exactly on the boundary of the future cone of $a$ is 0. So, in the general case, for this dense set, the boundary of each future cone is empty, and the causality relation coincides with the interior of the future cone.

From the physical viewpoint, this interior corresponds to kinematic causality, i.e., causality via regular particles (which can be at rest). Descriptions from [8, 18, 25] explicitly require this type of causality.

Second clarification: we restrict ourselves to positive properties of causality. The properties which form $n$-structural homogeneity mean that if some events $a_1, \ldots, a_k$ are related to each other in some reasonable way, then there exists a new element $a$ which is related to these elements $a_i$ in a known way.
Each relation between the original events $a_i$ and $a_j$ and each relation between a new event $a$ and the old events $a_i$ can be either positive (i.e., of the type $a_i < a_j$) or negative (i.e., of the type $a_i \not< a_j$). In [17], it was shown that for the properties formed from the positive relations, 3-structural homogeneity is indeed naturally equivalent to their standard formulation in [8, 18, 25], while for negative properties, we had to add additional postulates.

Thus, it is reasonable to restrict ourselves only to positive relations, in which $a$ causally follows some events $a_i$ and causally precedes some other events $b_j$. Of course, for this to be possible, we must make sure that every “lower bound” $a_i$ causally precedes every “upper bound” $b_j$. The resulting properties can be classified depending on how many lower bounds and how many upper bounds we have. The property corresponding to $\ell$ lower bounds and $u$ upper bounds will be described as an $(\ell, u)$-property.

For example:

- a $(1,1)$-property means that if $a_1 < b_1$, then there exists an event $a$ such that $a_1 < a < b_1$;
- a $(0,2)$-property means that for every $b_1$ and $b_2$, there exists an event $a$ such that $a < b_1$ and $a < b_2$; and
- a $(2,0)$-property means that for every $a_1$ and $a_2$, there exists an event $a$ such that $a_1 < a < a_2$.

Let us describe this in precise terms.

**Definition 6.** Let $l$ and $u$ be given natural numbers. We say that a partially ordered set $E$ has a $(\ell, u)$-property if for every two tuples $a_1, \ldots, a_\ell$ and $b_1, \ldots, b_u$ for which $a_i < b_j$ for all $i$ and $j$, there exists an element $a$ for which $a_i < a < b_j$ for all $i$ and $j$.

**Definition 7.** Let $n$ be a positive integer. We say that a partially ordered set $E$ is positively $n$-homogeneous if it has the $(\ell, u)$-property for all $\ell$ and $u$ for which $\ell + u \leq n$.

**Proposition 4.** If a partially ordered set $E$ is positively 4-homogeneous, then it is positively $n$-homogeneous for all integers $n$.

**Proof.**

1°. Since $E$ is positively 4-homogeneous, it has, in particular, the following three properties:

- it has the $(0,2)$-property, according to which for every $b_1$ and $b_2$, there exists an event $a$ such that $a < b_1$ and $a < b_2$;
• it has the (1,2)-property, according to which for every \( a_1, b_1, \) and \( b_2 \) for which \( a_1 < b_1 \) and \( a_1 < b_2 \), there exists an \( a \) for which \( a_1 < a < b_1 \), and \( a < b_2 \); and

• it has the (2,2)-property: for every four events \( a_1, a_2, b_1, \) and \( b_2 \) for which \( a_1 < b_j \) for each \( i \) and \( j \), there exists an \( a \) for which \( a_1 < a < b_j \) for all \( i \) and \( j \).

Let us show that from these three properties, we can deduce an \((\ell, u)\)-property for all possible natural numbers \( \ell \) and \( u \).

2°. Let us first prove that for every \( \ell \leq 2 \), the \((\ell, u)\)-property holds for all \( u \).

To prove this property, we must show that if we have elements \( a_1, \ldots, a_\ell, b_1, \ldots, b_u \) for which \( a_i < b_j \) for all \( i \) and \( j \), then there exists an element \( a \) for which \( a_i < a < b_j \) for all \( i \) and \( j \).

We will prove this by induction over \( u \).

2.1°. The induction base is straightforward: For \( u = 2 \), the existence of the desired element comes directly from the \((0,2)\), \((1,2)\), and \((2,2)\)-properties.

2.2°. Let us prove the induction step. Let us assume that we have already proven the \((\ell, u)\)-property; based on this, we will prove the \((\ell, u + 1)\)-property.

Indeed, let \( a_1, \ldots, a_\ell, b_1, \ldots, b_u, b_{u+1} \) be any given elements for which \( a_i < b_j \) for all \( i \) and \( j \). By the \((\ell, u)\)-property, there exists an element \( a' \) for which \( a_i < a' \) for all \( i \), and \( a' < b_1, a' < b_2, \ldots, a' < b_u \).

Now, for \( a_1, \ldots, a_\ell, a' \) and \( b_{u+1} \), we have \( a_i < a' \) and \( a_i < b_{u+1} \). So, due to the \((\ell, 2)\)-property, there exists an element \( a \) for which \( a_i < a, a < a' \) and \( a < b_{u+1} \). Since \( a < a' \) and \( a' < b_j \) for all \( j \leq u \), by transitivity, we conclude that \( a < b_j \) for all \( j \leq u \); we also know that \( a < b_{u+1} \). Thus, \( a_i < a < b_j \) for all \( i \) and \( j \).

The induction step is proven, and so is the statement.

3°. Now, for every \( u \), we can use a similar induction – this time over \( \ell \) – to prove that for every \( \ell \), we have a \((\ell, u)\)-property.

3.1°. The induction base is straightforward: For \( \ell \leq 2 \), the existence of the desired element \( a \) was proven in Part 2 of this proof.

3.2°. Let us prove the induction step. Let us assume that we have already proven the \((\ell, u)\)-property; based on this, we will prove the \((\ell + 1, u)\)-property.

Indeed, let \( a_1, \ldots, a_\ell, a_{\ell+1}, b_1, \ldots, b_u \) be any given elements for which \( a_i < b_j \) for all \( i \) and \( j \). By the \((\ell, u)\)-property, there exists an element \( a' \) for which \( a_1 < a', \ldots, a_\ell < a' \) and \( a' < b_j \) for all \( j \).

Now, for \( a', a_{\ell+1}, b_1, \ldots, b_u \), we have \( a' < b_1 \) and \( a_{\ell+1} < b_1 \). Thus, due to the \((2, u)\)-property, there exists an element \( a \) for which \( a' < a, a_{\ell+1} < a \), and \( a < b_j \) for all \( j \). Since \( a' < a \) and \( a_i < a' \) for all \( i \leq \ell \), by transitivity, we conclude that \( a_i < a \) for all \( i \leq \ell \); we also know that \( a_{\ell+1} < a \). Thus, \( a_i < a < b_j \) for all \( i \) and \( j \).

The induction step is proven, and so is the statement. The proposition is proven.
6 Conclusions

Causality is one of the most fundamental notions of physics. It is therefore important to be able to decide which statements about causality are correct in different models of space-time. In this paper, we analyze the computational complexity of the corresponding decision problems.

We start this analysis by finding a relation between Tarski’s elementary theory of geometry and the first order theory of causality for Minkowski space-time. Due to this relation, the decision problem for Minkowski space-time is as difficult as Tarski’s decision problem for elementary geometry – i.e.,

• requires at least exponential time, and
• can be solved in doubly exponential time.

We hope that the new relation between elementary geometry and causality theory will lead to even more efficient decision algorithms for physically important classes of causality formulas. Indeed, M. Ben-Or, D. Kozen, J. Reif, J. Canny, and J. Renegar have shown that many fragments of elementary geometry have a much faster decision procedure. In particular,

• the universal fragment has an exponential-time procedure (J. Renegar), and
• the entire existential fragment is in PSPACE (J. Canny) – hence, in practice, even more efficient than exponential time.

It is desirable to find physically interesting classes of causality formulas that correspond to such fragments – e.g., to the existential fragment.

After analyzing Minkowski space-time,

• we show that for every non-trivial space-time model, the corresponding decision problem is PSPACE-hard, and
• we produce a reasonable model of primordial space-time for which the decision problem is in PSPACE.

In the final section, we provide arguments that from the viewpoint of computational complexity, a general space-time should be similar either to Minkowski space-time or to the above primordial model.

Acknowledgments. This work was supported in part by NSF grants EAR-0225670 and EIA-0080940, by Texas Department of Transportation grant No. 0-5453, by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006-08, and by the Max Planck Institut für Mathematik.

The authors are thankful to A. Bovykin, D. Grigoriev, and G. Mints for valuable discussions, and to the anonymous referees for important suggestions.
References


