Intentional communication: computationally easy or difficult?
van Rooij, I.; Kwisthout, J.; Blokpoel, M.; Szymanik, J.K.; Wareham, T.; Toni, I.

Published in:
Frontiers in Human Neuroscience

DOI:
10.3389/fnhum.2011.00052

Citation for published version (APA):
Human intentional communication is marked by its flexibility and context sensitivity. Hypothesized brain mechanisms can provide convincing and complete explanations of the human capacity for intentional communication only insofar as they can match the computational power required for displaying that capacity. It is thus of importance for cognitive neuroscience to know how computationally complex intentional communication actually is. Though the subject of considerable debate, the computational complexity of communication remains so far unknown.

In this paper we defend the position that the computational complexity of communication is not a constant, as some views of communication seem to hold, but rather a function of situational factors. We present a methodology for studying and characterizing the computational complexity of communication under different situational constraints. We illustrate our methodology for a model of the problems solved by receivers and senders during a communicative exchange. This approach opens the way to a principled identification of putative model parameters that control cognitive processes supporting intentional communication.

Keywords: communication, computational complexity, computational modeling, intractability, Bayesian modeling, goal inference
Situational constraints that render communication tractable under the given models are then candidate explanations of the speed of everyday human communications. Crucially, these candidate explanations can then be empirically tested, assessing whether the same relation between situational constraints and computational demands predicted by the model fit with the behavior and cerebral processes observed in participants in the lab while solving the same communicative problems. This type of model-driven approach offers several benefits for cognitive neuroscience. First, it identifies putative model parameters that control the cognitive processes supporting intentional communication. Second, it provides a rigorous ground for empirical tests of those models. Third, it offers the possibility to test the neurophysiological plausibility and cognitive relevance of the model by comparing the predicted dynamics of relevant model parameters with the observed dynamics of communicators’ internal states (i.e., cerebral signals measured with neuroimaging methods).

OVERVIEW OF THE PAPER

We will illustrate our proposed approach using one particular model of intentional communication as a case study. The model is an extension of the Bayesian inverse planning (BIP) model of goal inference proposed by Baker et al. (2007, 2009). In Section “A Probabilistic Model of Intentional Communication,” we will describe in detail the original BIP model of goal inference and our adaptations of this model in the form of a Receiver model and Sender model. In Section “Computational Complexity Results,” we will study the computational complexity of these Receiver and Sender models, observing that both are computationally intractable unless properly constrained. We then set out to identify constraints that render the models tractable. In Section “Implications and Predictions for Cognitive Neuroscience” we will discuss how the conditions of (in)tractability reported in Section “Computational Complexity Results” can inform critical empirical tests in cognitive neuroscience. In Section “Open Theoretical Questions and Future directions” we...
present a set of open questions and suggestions for future theoretical research. We close, in Section “Conclusion,” with a reflection on the general relevance of this analytic methodology for cognitive neuroscience of communication and other forms of social interaction.

**A PROBABILISTIC MODEL OF INTENTIONAL COMMUNICATION**

Our model of communication builds on an existing model of how humans infer *instrumental* goals from observing someone’s actions, called the BIP model (Baker et al., 2007, 2009; see also Blokpoel et al., 2010). We extend this model to the domain of communication by making it apply to *communicative* goals as well. We first explain the BIP model (see Preliminaries: The BIP Model of Goal Inference), and then we explain how it can be adapted to the domain of intentional communication (see A BIP Model of Sender and Receiver).

**PRELIMINARIES: THE BIP MODEL OF GOAL INFERENCE**

According to the BIP model, observers assume that actors are “rational” in the sense that they tend to adopt those actions that best achieve their goals. Here “best” may, for instance, be defined in terms of (expected or believed) efficiency of a set of actions for achieving a given (combination of) goal(s). Say, a person has a single goal of tying his shoe laces. Then this person could make the necessary moves of the fingers, or he could start the finger movements, pause to scratch his chin, and then continue making the finger movements until his shoe laces are tied. If “rationality” is defined in terms of efficiency then the latter sequence of actions would be considered less rational for the goal of “tying one's laces” than the former sequence of actions. Be that as it may, the latter sequence of actions can of course be rational for a different goal, e.g., if the actor has two simultaneous goals “to tie the shoe laces” and “to get rid of that itch on the chin” (and an observer of the latter sequence of action will also likely attribute this combination of goals, over a single goal to the actor).

Given the assumption of rationality, (probabilistic) knowledge of the world and how actions are affected by it, and a measure of relative rationality of action–goal pairs, one can compute the probability that an agent performs an action given its goals, denoted

\[
Pr(\text{action}_1, \ldots, \text{action}_n | \text{goal}_1, \ldots, \text{goal}_m, \text{context})
\]

or in shorter format:

\[
Pr(\text{a}_1, \ldots, \text{a}_i | \text{g}_1, \ldots, \text{g}_m, c)
\]

An important insight of researchers such as Baker et al. (2007, 2009) is that observers can use this probabilistic model of planning to make inferences about the most likely goals that actors are pursuing. The reason is that the probability in Eq. 1a/b can be inverted using Bayes’ rule to compute the probability of a given combination of goals, given observations of actions that an actor performs:

\[
Pr(\text{g}_1, \ldots, \text{g}_m | \text{a}_1, \ldots, \text{a}_i, c) \propto \Pr(\text{a}_1, \ldots, \text{a}_i | \text{g}_1, \ldots, \text{g}_m, c) \Pr(\text{g}_1, \ldots, \text{g}_m | c)
\]

Of all the possible combinations of goals that an observer can (or does) entertain, the goal combination that maximizes the probability in Eq. 2 best explains why the observed actions were performed, and according to the BIP model it is this goal combination that an observer will infer.

In sum, goal inference under the BIP model equals the computation of the following input–output mapping, informally stated.

**Goal Inference (informal)**

**Input:** A representation of the probabilistic dependencies between actions, goals, and states and how these dependencies change over time, and a sequence of observed actions and world states.

**Output:** A combination of goals that best explains the sequence of actions and world states against the background of the probabilistic dependencies between actions, goals, and world states and how these dependencies change over time.

To be able to analyze the computational complexity of the Goal Inference problem we need to define formal counterparts of all the notions and constructs introduced in the informal problem definition above. Following Baker et al. (2007, 2009), we use BIP-Bayesian networks to formally model the input representations in the Goal Inference problem (for a mathematical definition see Preliminaries from Bayesian Modeling in the Appendix). In these networks, actions, goals, and world are modeled as value assignments to nodes in the network, and probabilistic dependencies are indicated by directed arcs connecting such nodes (see Figure 2 for an illustration). For each node in the network there is an associated prior probability of the node taking on a value (e.g., “true” or “false”). In addition, each node has an associated probability distribution, coding the conditional probability of the node taking on a value as a function of the probability of different possible value assignments for the nodes that are connecting to it (called, its “parents”). A set of observed actions and world states is modeled by a value assignment to a sequence of action and state nodes. A combination of goals is modeled as a truth assignment for the goal nodes. A combination of goals is said to “best” explain the observations of actions and world states if it maximizes the probability defined in Eq. 2.

Having formalized all the relevant notions in the Goal Inference problem, we can now state the formal counterpart of the problem (for explanation of notation and mathematical concepts, refer to Preliminaries from Bayesian Modeling in Appendix).

**Goal Inference (formalized)**

**Input:** A BIP-Bayesian network \( B = (D, G) \), where \( D \) is a dynamic Bayesian network with time slices \( D_1, \ldots, D_n \), where each time slice contains an action variable \( A \) and a state variable \( S \), and \( G = G_1 \ldots G_k \) denotes a set of instrumental goals. Further, a set of observed actions and states \( a \cup s \).

**Output:** The most probable joint value assignment \( g \) to the goals in \( G \), i.e., \( \arg \max G \Pr(\text{G} = g | A = a, S = s) \), or \( \emptyset \), if \( \Pr(\text{G} = g | A = a, S = s) = 0 \) for all joint value assignments \( g \) (here the output \( \emptyset \) can be read as meaning “no plausible goal attribution can be made”).

---

1 In other words, in the BIP model, goal inference is conceptualized as a form of probabilistic inference to the best explanation, a.k.a. abduction (Charniak and Shimony, 1990). As this model is situated at Marr’s (1982) computational level, the theory is in principle consistent with a variety of hypotheses about the architecture that implements the postulated computations, ranging for instance from predictive coding or forward models (Miall and Wolpert, 1996; Orto et al., 2005; Kilner et al., 2007a,b) to neural network models (Paine and Tani, 2004) or random sampling models (Vul et al., 2009).
The computational complexity of this model has previously been analyzed by Blokpoel et al. (2010). Specifically, these authors presented a mathematical proof that the problem is computationally intractable (NP-hard) if no constraints are imposed on the input representations. They furthermore found that the problem is tractably computable under the constraint that the set of candidate goals \( G \) is small and/or the probability of the most probable goal assignment \( p \) is large. Here, we will use the Goal Inference problem as an inspiration for similarly explicit characterizations of the Sender and Receiver problems solved by human communicators. As we will see, some (though not all) of the computational complexity results for Goal Inference apply also to the new Sender and Receiver problems.

**A BIP MODEL OF SENDER AND RECEIVER**

If indeed – as the BIP-model postulates – observers of actions infer goals from actions by means of an inference to the best explanation, then senders can use this knowledge to predict how a receiver will interpret their actions ahead of time. For instance, they could engage an internal simulation of a receiver’s inferences to the best explanation when considering the suitability of candidate actions (Noordzij et al., 2009). Such a simulation subroutine could be called multiple times during the planning of instrumental and communicative behaviors, allowing a sender to converge on an action sequence which is both efficient and is likely to lead a receiver (if properly simulated) to attribute the correct (i.e., intended) communicative goal to the sender. This conceptualization of the computational bases of sender signal creation can be summarized by the following informal input–output mapping:

\[
S_1 \rightarrow \ldots \rightarrow S_T \quad A_1 \rightarrow \ldots \rightarrow A_{T-1} \quad G_1 \rightarrow \ldots \rightarrow G_K
\]

\[
\text{Goal Inference} \quad \text{Input:} \quad \text{A representation of the probabilistic dependencies between actions, goals, and states and how these dependencies change over time, and one or more communicative and instrumental goals.}
\]

\[
\text{Output:} \quad \text{A sequence of actions that will lead to the achievement of the instrumental goals and will lead a receiver to attribute the correct communicative goals to the sender.}
\]

Note that our model allows for the possibility that senders can have simultaneous instrumental and communicative goals as seems required for ecological validity. After all, in everyday settings communicative behaviors are typically interlaced with several sequences of instrumental actions – think, for instance, of a car driver signaling to another driver at night that his lights are off, while at the same time trying to drive safely and stay on route.

Building on the formalisms used in the Goal Inference model, the informal Sender model yields the following formal input–output mapping (see Figure 3 for an illustrative example of a BIP-network for communicative goals).

**Sender (informal)**

**Input:** A representation of the probabilistic dependencies between actions, goals, and states and how these dependencies change over time, and one or more communicative and instrumental goals.

**Output:** A sequence of actions that will lead to the achievement of the instrumental goals and will lead a receiver to attribute the correct communicative goals to the sender.

---

3We remark that the computational-level Sender and Receiver models that we propose are in one respect a simplification and in another an enrichment of the Shafto and Goodman (2008) model (see also Frank et al., 2009). Our models are a simplification in the sense that they do not assume infinite recursive reasoning by sender and receiver about each other, in contrast to the Shafto and Goodman model. Our model is an enrichment in the sense that we adopt the Bayesian network structure of the Baker et al. (2007, 2009) BIP model to have a more general characterization of the assumed dependencies between states, goals, and actions.
a sender with the aim of making her communicative goals interpretable by a receiver, while concurrently trying to achieve one or more instrumental goals unknown to the receiver. This difference between the Goal Inference problem and the Receiver problem can have non-trivial consequences for the computational complexity of the Receiver problem, which is why we need to analyze it anew in this paper.

For completeness and clarity, we state the Receiver problem in both informal and formal format below.

**Receiver (informal)**

**Input:** A representation of the probabilistic dependencies between actions, goals, and states and how these dependencies change over time, and a sequence of observed actions and world states.

**Output:** A combination of communicative goals that best explains the sequence of actions and world states against the background of the probabilistic dependencies between actions, goals, and world states and how these dependencies change over time.

**Receiver (formalized)**

**Input:** A BIP-Bayesian network \( B = (D, G) \) as in the Sender problem; a set of observed actions and states \( a \cup s \).

**Output:** The most probable joint value assignment \( g_c \) to the communicative goals in \( G_c \), i.e., \( \text{argmax}_{g_c} \Pr(G_c = g_c | A = a, S = s) \), or \( \emptyset \), if \( \Pr(G_c = g_c | A = a, S = s) = 0 \) for all joint value assignments \( g_c \).

Having defined both the generic Sender and Receiver problems, we are now in a position to analyze their computational complexity. We report on our analyses in the next section.

**COMPUTATIONAL COMPLEXITY RESULTS**

In this section we state all computational complexity results. Readers interested in full details on the mathematical proofs are referred to the Section “Computational Models and Computational Complexity Analyses” in the Appendix. This section assumes familiarity with concepts and methods from the field of computational complexity analysis (for a primer see Preliminaries from Complexity Theory in the Appendix). For instance, contrary to the suggestion that computationally intractable functions can be approximately computed efficiently (e.g., Love, 2000; Chater et al., 2003, 2006), computational complexity theory (cf. Kwisthout et al., 2011) has clearly shown that many intractable (e.g., NP-hard) functions cannot be efficiently approximated (Arora, 1998; Ausiello et al., 1999), and almost all are intractable to approximate if only a constant sized error is allowed (Garey and Johnson, 1979; van Rooij et al., 2010). Accordingly – although the analyses we provide work under the assumption that the models exactly predict the outputs of human computations – we conjecture that the obtained results also apply under the assumption that the models approximately predict the outputs of human computations. Here “approximate” means that the approximate prediction does not differ from the exact prediction by more than some small constant factor. For ideas on how approximation can be explicitly taken into account in fixed-parameter tractability analyses we refer the reader to Hamilton et al. (2007) and Marx (2008).

3We are aware of computer simulation studies in the cognitive science literature that seem to suggest that Bayesian models can be efficiently approximated (Vul et al., 2009; Sanborn et al., 2010). However, the intractability results for approximating Bayesian computations available in the computer science literature show that these “simulation-based approximability findings” cannot generalize to all types of inputs, but only to some that have special properties (cf. “parameters”) that can be exploited for tractable (exact or approximate) computation (Kwisthout et al., in 2011).
INTRACTABILITY OF GENERIC COMMUNICATION

Our two main computational intractability results are:

Result 1. The Receiver problem is NP-hard.

Result 2. The Sender problem is NP-hard.

These results establish that both Receiver and Sender problems are computationally as hard to compute as a whole class of problems that are strongly conjectured to be intractable, called the NP-complete problems, and possibly even harder. This means that Sender and Receiver problems cannot be computed efficiently (more precisely, in polynomial-time) unless all NP-complete problems can be (Garey and Johnson, 1979; Aaronson, 2005; Fortnow, 2009). As there is both theoretical and empirical evidence that no NP-complete problem allows for an efficient solution procedure we conclude that Receiver and Sender problems, as we have defined them, are intractable. We note that our Results 1 and 2 are, to the best of our knowledge, the first formal proofs that are consistent with intuitive claims of intractability in the communication literature (Levinson, 1995, 1996; Pickering and Garrod, 2004; Barr and Keysar, 2005).

IDENTIFYING CONDITIONS FOR TRACTABILITY

Importantly, our analyses do not stop at the intractability results (Results 1 and 2). On the contrary, we view such results as merely the fruitful starting point of rigorous analyses of the sources of complexity in human communication. For these further analyses we adopt a method for identifying sources of intractability in cognitive models developed by van Rooij and Wareham (2008; see also van Rooij, 2008; van Rooij et al., 2008). The method builds on concepts and techniques from the mathematical theory of parameterized complexity (Downey and Fellows, 1999), and works as follows.

First, one identifies a set of potentially relevant problem parameters $K = \{k_1, k_2, \ldots, k_m\}$ of the problem $P$ under study (for us, the Receiver and Sender problems). Then one tests if it is possible to solve $P$ in a time that can grow excessively fast (more precisely: exponential or worse) as a function of the elements in the set $K = \{k_1, k_2, \ldots, k_m\}$ yet slowly (polynomial) in the size of the input$^1$. If this is the case, then $P$ is said to be fixed-parameter (fp-) tractable for parameter set $K$, and otherwise it is said to be fp-intractable for $K$. Observe that if a parameter set $K$ is found for which $P$ is fp-tractable then the problem $P$ can be solved quite efficiently, even for large inputs, provided only that the members of $K$ are relatively small. In this sense the “unbounded” nature of $K$ can be seen as a reason for the intractability of the unconstrained version of $P$. Therefore, we also call $K_a source of intractability of P$.

The Receiver and Sender models have several evident parameters, each of which may be a source of the intractability inherent in the general problems postulated by these models. Table 1 gives an overview of the parameters that we consider here. Using the abovementioned methodology for fp-tractability analysis we have been able to derive a set of fp-(in)tractability results, which are summarized in Table 2. We will discuss these results and what they imply for the tractability of the Receiver and Sender under different situational constraints.

$^1$More formally, this would be a time on the order of $f(k_1, k_2, \ldots, k_m)n^c$, where $f$ is an arbitrary computable function, $n$ is a measure of the overall input size, and $c$ is a constant.

We start by considering the fp-(in)tractability of the Receiver problem. We have two main fp-tractability results:

Result 3. The Receiver problem is fp-tractable for parameter set $\{|G_i|, |G_j|\}$.

Result 4. The Receiver problem is fp-tractable for parameter set $\{|G_i|, 1 - p\}$.

Note that, by definition, a problem is fp-tractable for a parameter set $K$ then it is also fp-tractable for any superset $K' \supseteq K$. Hence, Results 3 and 4 imply all other fp-tractability results listed in Table 2 for the Receiver problem, including the fp-tractability of the problem for the combined parameter set $\{|G_i|, |G_j|, 1 - p\}$.

Result 3 establishes that it is possible to compute the input–output mapping defined by the Receiver problem in a time which grows fast (read: exponential) only in the two parameters $|G_i|$ and $|G_j|$, and slow (read: polynomial) in the remainder of the input, regardless the size of other input parameters. Result 4 establishes that it is

| Parameter Description | $|G_i|$ | $|G_j|$ | $|G_i|, |G_j|$ |
|-----------------------|--------|--------|----------------|
| $G_i$                  | The size of the set of communicative goals |
| $G_j$                  | The size of the set of instrumental goals |
| $|A|$                  | The size of the set of action nodes |
| $|S|$                  | The size of the set of states |
| $T$                   | The number of time slices |
| $1 - p$               | Here $p$ is the probability of the most probable communicative goal attribution |

Table 2 | Overview of complexity results and open questions for Receiver problem (top) and Sender problem (bottom).

| Parameter Description | $|G_i|$ | $|G_j|$ | $|G_i|, |G_j|$ |
|-----------------------|--------|--------|----------------|
| Receiver               |        |        |                |
| $|A|$                  | NP-hard | fp-intractable | fp-intractable | fp-tractable |
| $1 - p$               | fp-intractable | fp-intractable | fp-intractable | fp-tractable |
| Sender                |        |        |                |
| $|A|$                  | NP-hard | fp-intractable | fp-intractable | fp-tractable |
| $1 - p$               | fp-intractable | fp-intractable | fp-intractable | fp-tractable |

The table lists all possible combinations of parameters from Table 1: each cell stands for the parameter set consisting of its row and column labels. In the special case of the empty parameter set classical (non-parameterized) complexity applies, in this case NP-hardness. Note that in our model $|A| = |S| = T$. Hence for all listed results including parameter $|A|$ in the parameter set the other two parameters $|S|$ and $T$ may be replaced (or even added) with no change to the listed result. Full details on the proofs of theorems and explanations of corollaries and conjectures can be found in Section “Computational Models and Computational Complexity Analyses” in the Appendix.

$^2$Technically, as a consequence of probabilities being constrained to a value between 0 and 1, these results are not fp-(in)tractability results in the formal sense of the term. Yet, we do have formal proofs that show $1 - p$ is (or respectively, is not) a source of complexity in the same sense that the formal notion of fp-(in)tractability intends to capture.
also possible to compute the input–output mapping defined by the Receiver problem in a time which grows fast only in the two parameters \(1 - p\) and \(|\mathbf{G}_1|\), and slow in the remainder of the input, regardless the size of other input parameters (including \(|\mathbf{G}_0|\)). Informally, this means that the inference task modeled by the Receiver problem can be performed fast, even for large networks of beliefs, if the number of possible instrumental goals that the receiver believes the sender may have (\(|\mathbf{G}_1|\)) is relatively small and at the same time either the number of communicative goals is relatively small or the probability of the most likely communicative goal attribution (\(p\)) is relatively large.

These mathematical results lead to a clear prediction. A receiver is able to quickly attribute communicative intentions to a sender’s actions if the number of instrumental goals that the receiver assumes that the sender is pursuing in parallel to her communicative goals is small and the sender does not have many communicative goals she wishes to convey simultaneously, or otherwise she was able to construct a sequence of actions that leads to a signal of low ambiguity (captured by small \(1 - p\)).

Importantly and perhaps counter-intuitively, low ambiguity of the signal is by itself not sufficient for tractability, as we have the following intractability result.

**Result 5.** The Receiver problem is \(fp\)-intractable for parameter set \([1 - p]\).

Similarly, a small number of communicative goals is by itself not sufficient for tractability, as we also have the following \(fp\)-intractability result.

**Result 6.** The Receiver problem is \(fp\)-intractable for parameter set \(|\mathbf{G}_0|\).

However, if senders were to focus solely on communicating (and forgetting for the moment about any other instrumental goals they may also want to achieve), then \(|\mathbf{G}_1| = 0\), and hence low ambiguity (or small number of communicative goals) would by itself suffice to make the inference task easy for the receiver. The reason that low ambiguity (or small number of communicative goals) is not enough for tractability for \(|\mathbf{G}_1| > 0\) is that, in order to compute the probability of the most probable communicative goal assignment, the receiver needs to do this computation against the background of all possible instrumental goals a sender may have, as having one or more of those can affect the probability of the target communicative goals. The number of possibilities that the receiver needs to be taken into account grows, in this case, exponential in \(|\mathbf{G}_1|\), and therefore the computation is too resource demanding to be done efficiently for large \(|\mathbf{G}_1|\).

Having seen that neither large \(p\) nor small \(|\mathbf{G}_1|\) are by itself sufficient for tractability of the Receiver problem, but that combining either with small \(|\mathbf{G}_1|\) does yield tractability, one may ask if there are other combinations of constraints that are sufficient for tractability of the Receiver problem. As can be seen from the overview of \(fp\)-(in)tractability results in Table 2, no combination of parameters excluding \(|\mathbf{G}_0|\), \(|\mathbf{G}_1|\) or \(|\mathbf{G}_1|, 1 - p\) as subsets yields \(fp\)-tractability for the Receiver problem. This means that these two sets seem to fully characterize what makes the Receiver problem difficult or easy (at least, relative to the total set of parameters that we consider here and listed in Table 1).

As the parameter \(|\mathbf{G}_0|\), \(|\mathbf{G}_1|\) and \(1 - p\) all seem to be, to some extent, under the control of the sender, Results 3 and 4 raise the question of how computationally complex it is for senders to ensure these parameters (or pairs of them) are small, and thereby make the intention attribution task tractable for receivers. We will discuss a set of results that are relevant to this question, starting with the following:

**Result 7.** The Sender problem is \(fp\)-intractable for parameter set \(|\mathbf{G}_1|, \mathbf{G}_0|\).

In other words, the Receiver problem being tractable is by itself not sufficient for the Sender problem to be tractable (compare Result 3 and Result 7). Additional constraints need to be assumed to explain the tractability of the sender’s task. Result 8 shows that a low upper-bound on the length of the sequence of actions that needs to be planned can help make the sender’s task easier.

**Result 8.** The Sender problem is \(fp\)-tractable for parameter set \(|\mathbf{A}|, |\mathbf{G}_1|, |\mathbf{G}_0|\).

Result 8 can be understood as a consequence of the ability of a sender to use the same \(fp\)-tractable algorithm that the receiver can use to tractably infer the sender’s communicative goals to predict the receiver’s interpretation of a given action sequence, and then search the space of action sequences (which is exponential only in \(|\mathbf{A}|\)) for a sequence that yields the right receiver inference. This algorithmic strategy does not yield \(fp\)-tractability for the Sender problem for the parameter set \(|\mathbf{A}|, |\mathbf{G}_1|, 1 - p|\), because contrary to \(|\mathbf{G}_1|\), which is constant for all action sequences the sender may consider, the value of \(1 - p\) is different for different candidate action sequences. As a result, \(1 - p\) cannot generally be assumed to be small for all possible action sequences. Whether or not a different strategy exists that can exploit this combination of parameters for solving the Sender problem in fixed-parameter tractable time is currently unknown:

**Open question.** Is the Sender problem is \(fp\)-tractable for parameter set \(|\mathbf{A}|, |\mathbf{G}_1|, 1 - p|\)?

What we do know is that no other combination of parameters excluding \(|\mathbf{A}|, |\mathbf{G}_0|, |\mathbf{G}_1|\) as a subset suffices to render the Sender problem \(fp\)-tractable. If future research were to reveal that the Sender problem is \(fp\)-tractable for parameter set \(|\mathbf{A}|, |\mathbf{G}_1|, 1 - p|\), then this would indicate that given that the Receiver problem is tractable (e.g., Result 3 and 4) parameter a bound on \(|\mathbf{A}|\) suffices for tractability of the Sender problem. If, on the other hand, it would turn out that the Sender problem is \(fp\)-intractable for parameter set \(|\mathbf{A}|, |\mathbf{G}_1|, 1 - p|\), then this would underscore the high complexity of the Sender problem, as it is impossible for the input to the Sender problem to be large when all three parameters \(|\mathbf{A}|, |\mathbf{G}_0|, |\mathbf{G}_1|\) are small at the same time.

**DISCUSSION**

We analyzed the complexity of two models of senders and receivers in a communicative exchange. In Section “Implications and Predictions for Cognitive Neuroscience” we will discuss the implications of our findings for cognitive neuroscience, including a set of empirical predictions testable by cognitive neuroscience methods. All predictions made in this section must, of course, be understood relative to the models that we studied, which inevitably make some basic assumptions about properties of the mental representations of receivers and senders in the communication task. In Section "Open Theoretical Questions and Future directions" we will explicitly discuss these assumptions and give pointers for possible extensions or
adaptsions of the models. As such new model variants need not
inhibit all the (in)tractability results from our Sender and Receiver
models they naturally yield a set of open questions for future theo-
retical research.

**IMPLICATIONS AND PREDICTIONS FOR COGNITIVE NEUROSCIENCE**

We presented two new models of the tasks engaging senders and
receivers during a communicative exchange: see the Receiver and
Sender models in Section “A BIP model of Sender and Receiver.” As
these models are situated at Marr’s (1982) computational level, they
are not bound to particular assumptions about a specific algorithm
that human brains would use to perform the postulated computa-
tions, nor about how these algorithms are actually implemented in
neural mechanisms. They merely state the hypothesized input–out-
put mappings assumed to underlie the receiver and sender tasks in
communication respectively. This high level of abstraction has the
benefit of generalizability, i.e., the sources of intractability identified
in the models are inherent in the problems that they describe, and
not specific to particular algorithms for solving those problems
or implementations thereof. As a consequence, our intractability
results are proofs of impossibility: no algorithmic implementation of
any type, in any neural mechanism, can ever compute the problems
quickly if the input domain is unconstrained.

It could be argued that, since human communication is evi-
dently tractable for real humans in the real world, the present results
are useless. However, the mismatch between model and reality is
informative for improving formal models of human communica-
tion, and for rejecting claims that intentional communication is
inherently computationally impossible. This high level of abstraction
has the benefit of generalizability, i.e., the sources of intractability identified
in the models are inherent in the problems that they describe, and
not specific to particular algorithms for solving those problems
or implementations thereof. As a consequence, our intractability
results are proofs of impossibility: no algorithmic implementation of
any type, in any neural mechanism, can ever compute the problems
quickly if the input domain is unconstrained.

The fp-tractability results are possibility proofs in the sense that
there exist fp-algorithms that can perform the sender and receiver
tasks quickly if certain situational parameters are constrained to
take small values. Specifically, we proved the existence of an fp-
tractable algorithm for the Sender problem which exploits small
values for parameters \( |A| \) (the number of to be planned sender
actions), \( |G_j| \) (the number of sender instrumental goals), and \( |G_j| \)
(the number of sender communicative goals). Furthermore, we
proved the existence of an fp-tractable algorithm for the Receiver
problem that can exploit the same set of parameters, or even a
smaller set which does not include \( |A| \). In addition, we proved the
existence of an fp-tractable algorithm for the Receiver problem that
can exploit the parameters \( |G_j| \) and \( 1 - p \) (where \( p \) is the probability of
the most likely communicative goal attribution). Whether or
not human brains actually use and implement such fp-tractable
algorithms is an empirical question. Note that posing this question
is only made possible by the type of complexity analysis described
in this paper. We believe this is a relevant question, as it stems
from a principled analysis rather than occasional observations.

The cognitive neuroscience literature is replete with examples of
cerebral responses being attributed to “task difficulty” on the basis
of introspections. Here, we show how task difficulty could be for-
mally parameterized. This approach opens the way for using cogni-
tive neuroscience methods to test whether and how human brains
exploit similar fp-tractable computations to efficiently solve com-
municative problems. The prediction to be tested is that resource
demands during human communication should be extremely sen-
sitive to the parameters identified as critical for tractability of
the model. Recall that an fp-tractable algorithm for parameters \( \{k_1, k_2, \ldots, k_m\} \) runs fast if and only if all the parameters \( k_1, k_2, \ldots, k_m \)
are constrained to sufficiently small values in the current input (i.e., situation). If, in the lab, one were to create communication tasks in which the critical parameters can be systematically manipulated, then two critical predictions follow from the hypothesis that the senders use, say, the fp-tractable algorithm underlying Result 8, and receivers use the fp-tractable algorithm underlying Result 3 or 4:

1. As long as \(|A|, |G_c|, 1 - p|\) remain small the task of communication should proceed efficiently.
2. As soon as two parameters in the set \(|G_c|, 1 - p|\) take on large values, the task of communication would start to consume an excessive amount of computational resources (for sender and/or receiver, depending in which the relevant parameters are large), which may be reflected in a decrease of speed, increase of communicative errors, and/or increased brain activities in certain brain areas underlying the relevant computations.

If these predictions were to be confirmed, then this result would provide evidence for both cerebral use of fp-tractable computations, and for the Receiver and Sender models as defined above. A natural next question would then be how these algorithms are neurally implemented. If, on the other hand, the predictions were to be disconfirmed, then two interpretations are open. Either, the Receiver and Sender models are not valid, and this would motivate new modeling directions; or Receiver and Sender models are valid, but humans use different fp-tractable algorithms for computing them, exploiting parameters other than the ones that we have studied here. This latter option is not excluded, as there may be parameters of the problems outside the set described in Table 1. Constraining these additional parameters may render computation of these problems tractable as well. Both research directions appear highly informative.

We end this section by remarking that complexity results, such as we have derived here, also expand the ways in which computational-level models can be tested for input–output equivalence with human communicators. This form of testing involves presenting a participant with a situation that can be described as a particular input in the model’s input domain, computing the output that according to the model corresponds to that input, and testing if the responses of the participant match the predicted outputs. This test is quite common in cognitive science, and was also the method used by Baker et al. (2007, 2009) to test their computational-level model of (instrumental) goal inference. However, this methodology is often used within particular “toy-scenarios,” where the tested situations need to be exceedingly simple, such that inputs can be explicitly modeled and output computation is tractable for the researcher (see e.g., Oztop et al., 2005; Guipers et al., 2006; Erlhagen et al., 2006; Baker et al., 2007; Yoshida et al., 2008). Without a clear estimate of the complexity of the problem at hand, the results found within the boundaries of those toy-scenarios might easily fail to scale up to more complex and realistic domains. For instance, those computational models of goal inference that seem to work (i.e., that make plausible goal inferences without running into tractability issues) severely restrict the possible contexts and the number of possible actions and goals, keeping their application domain far removed from realistically complex situations. In contrast, the present results enable communication researchers to move away from trivial “toy domains,” using the fp-tractable algorithms identified here to expand the domains in which to compute predicted input–output mappings (for details on the algorithms see the Appendix). In cases where it is difficult or impossible to get an explicit model of the entire communicative setting, our proposed methodology of testing sources of tractability in the computational-level models of interest can be used instead.

OPEN THEORETICAL QUESTIONS AND FUTURE DIRECTIONS

For our case study, we opted for analyzing the Receiver and Sender models as defined in Section “A BIP model of Sender and Receiver.” By doing so, we made several commitments that may have affected our complexity results. We will briefly highlight those commitments known to affect the results, and those which we believe are unlikely to have had an effect. We will distinguish between model-specific commitments (such as the nature of the connectivity of the input networks and assumptions about the amount and type of information available to senders and receiver) and our choice of formalism (i.e., a probabilistic formalism).

The models that we defined make the strong commitment that there is no direct (probabilistic) dependency between states and goals, as all such dependencies are assumed to be mediated by actions (see Figures 2 and 3). This assumption seems problematic if one considers that states of the world may make certain goals of an actor more likely (e.g., if a cup is present and within reach, then a grasping action toward that cup is more likely to have as goal “grasp the cup,” as compared to a situation where the cup is not present, or out of reach). Although this commitment strictly speaking reduces the validity of the model, we know it has no effect on our complexity results (they would apply equally well to a model that allows for direct state-goal dependencies). The reason is that the model makes an additional assumption, namely the assumption that all relevant states and actions are observable by the receiver. This brings us to the point of validity of this assumption.

Although complete observability may hold true in some situations (and in those situations the models may thus apply without problem), it is likely that in many real-world situations not all relevant actions and states are observable. Think, for instance, of the possibility that part of a visual scene is occluded by nearby objects or even an eye blink. Then some actions and state changes may be observed directly, while others need to be inferred by an observer. Humans are often able to do this, and models of receivers should be able to explain how they can do so. Because our fp-tractability results reported in Section “Computational Complexity Results” heavily depend on the assumption of complete observability, we must admit that our analyses have not yet shed light on how communication can proceed efficiently under conditions of partial observability. Future research on the computational complexity of model variants that include the assumption of partial observability are thus called for.

Another important property of our models is that they assume that the sender and receiver know everything that is relevant to the context of the communicative exchange, including which action and goal variables are relevant, and their probabilistic interdependencies. Clearly, this assumption sweeps a considerable amount of computational complexity under the rug; after all, computing the
set of relevant information itself may well be computationally very resource demanding. Be that as it may, our analyses apply in those cases where humans do have complete knowledge of everything relevant to the communication task, and this is a condition that can certainly be met in laboratory settings used to investigate sources of complexity in communication as suggested in Section “Implications and Predictions for Cognitive Neuroscience.” Also, our analyses show that even if there are not an indefinite number of intentions that senders and receivers may entertain (see our Introduction), communication under simpler models can still be intractable. In contrast, even in the case where sender and receiver have complete common ground on the set of possible sender goals, the planning biases of the sender (i.e., the probabilistic dependencies between sender’s goals and actions, and world, as well as how those dependencies change over time) and all events that are relevant to the interpretation of the sender’s actions, still the task of communication is computationally intractable. Future research may address the additional computational complexities introduced by establishing this common ground in the first place. This could be done by building an explicit model of that process and submitting it to the same sort of complexity analyses as we have performed here for the Sender and Receiver models.

Last, we reflect on our choice of formalism: we opted for Bayesian networks as a formal representation of communicators’ situational knowledge and we used the mathematical notion of posterior probability to define the intuitive notion of “inference to the best explanation.” This choice of probabilistic formalism aligns our models with the current trend in cognitive neuroscience, which assumes that the brain implements its cognitive functions by means of probabilistic computations (Wolpert and Ghahramani, 2000; Ma et al., 2006), including the cognitive function of inferring people’s goals and intentions from their actions (Cuijpers et al., 2006; Baker et al., 2007, 2009; Kilner et al., 2007a, 2007b). Importantly, our complexity results should not be seen as bound to this formalism. It is well known that knowledge structures can be represented using other formalisms, such as (non-classical) logics or as constraint networks (van Ditmarsch et al., 2007; Russell and Norvig, 2009). In those cases, “Inference to the best explanation” could be defined in terms of logical abduction (e.g., Bylander et al., 1991; Eiter and Gottlob, 1995; Nordh and Zanuttini, 2005, 2008) or constraint satisfaction (e.g., Thagard and Verbeurgt, 1998; Thagard, 2000). The present results would apply to those alternative formalisms, provided that the same basic structural information is encoded in the representations of those alternative formal languages. If models identical or similar to those analyzed in this paper get expressed in the different formalisms, chances are that all such formal model variants can be transformed into each other via mathematical reductions not unlike the one we used to prove that the Sender and Receiver models are of the same computational complexity as the known intractable logic problem called Satisfiability (see Computational Models and Computational Complexity Analyses). One would not want complexity predictions to depend on the formal language chosen to express a model, in the same way that one would not want a verbal theory to lead to different predictions if it is expressed in English or in Dutch.

CONCLUSION

Recent accounts of human communication have focused on the motoric abilities of individual agents (Jacoboni et al., 2005), as if the meaning conveyed by a pointing finger were an intrinsic property of that action. Others have correctly highlighted the fact that, when unconstrained, human communication is intractable (Levinson, 1995, 2006). Here we have tried to move our knowledge forward, showing that those positions are case limits of a continuum as a function of the constraints applicable to the inferences involved in human communication. We believe the approach exemplified in this paper has the potential to formally define sources of computational complexity, opening the way to tackle some hard problems in intentional communication. Acquiring this knowledge appears fundamental to generate a neurobiologically grounded formal account of the computational mechanisms supporting our communicative abilities, which is a necessary step to understand the biological and cognitive bases of human sociality.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their valuable comments on an earlier version of this article. Johan Kwisthout was supported by the OCTOPUS project under the responsibility of the Embedded Systems Institute. Mark Blokpoel was supported by a DCC PhD grant awarded to Iris van Rooij and Ivan Toni. Todd Wareham was supported by NSERC Personal Discovery Grant 228104. Jakub Szymanki was supported by NWO (VICI grant #277-80-001). Ivan Toni was supported by NWO (VICI grant #453-08-002).

REFERENCES


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 July 2010; accepted: 16 May 2011; published online: 30 June 2011.


Copyright © 2011 van Rooij, Kwisthout, Blokpoel, Szynarik, Wareham and Toni. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.

Intentional communication: easy or difficult?
APPENDIX

In this Appendix we explain general concepts, notation and terminology that are used in the main text. In particular, we introduce the mathematical notions of dynamic Bayesian Networks and a special case of such networks, called Bayesian Inverse Planning (BIP-)Bayesian Networks (Baker et al., 2007, 2009; Blokpoel et al., 2010). Further, we explain the conceptual bases of computational complexity analysis, including definitions and proof techniques. Last, we present formal proofs for all results in the main text.

PRELIMINARIES FROM BAYESIAN MODELING

Bayesian or probabilistic networks are tools for modeling uncertain knowledge (Jensen and Nielsen, 2007). A Bayesian network \( B \) is a graphical structure that models a set of stochastic variables, the (in-)dependencies among these variables, and a joint probability distribution over these variables. \( B \) includes a directed acyclic graph \( N \), modeling the variables and (in-)dependencies in the network, and a set of parameter probabilities \( \Gamma \) in the form of conditional probability tables (CPTs), capturing the strengths of the relationships between the variables. A CPT is associated with each variable in the network and contains the probability distribution of that variable \( X \) for each joint value assignment to its parents \( Y_1 \ldots Y_n \), denoted by \( \Pr(X | Y_1 \ldots Y_n) \). The network models a joint probability distribution \( \Pr(V) \) over its variables \( V \), where the arcs denote dependencies between variables and lack of arcs between variables denote independencies in the probability distribution. Figure A1 illustrates an example network that models the dependencies between the variables WetGrass (\( G \)), SprinklerOn (\( S \)), Rain (\( R \)) and PowerFailure (\( P \)). The grass can be wet both due to rain and to the sprinkler being on. Whether the sprinkler is on depends on whether there is a power failure. To model abduction there are two types of Bayesian inference we use in our proofs. For completeness we introduce their formal input-output mappings here.

Most probable explanation

Input: A probabilistic network \( B = (N, \Gamma) \), where \( N = (V, \Lambda) \) and \( V \) is partitioned into a set of evidence nodes \( E \) with a joint value assignment \( e \) and an explanation set \( M \), such that \( E \cup M = V \).

Output: What is the most probable joint value assignment \( m \) to the nodes in \( M \) given evidence \( e \)?

Partial maximum a-posteriori probability

Input: A probabilistic network \( B = (G, \Gamma) \), where \( V \) is partitioned into a set of evidence nodes \( E \) with a joint value assignment \( e \), a set of intermediate nodes \( I \neq \emptyset \) and an explanation set \( M \), such that \( E \cup I \cup M = V \).

Output: What is the most probable joint value assignment \( m \) to the nodes in \( M \) given evidence \( e \)?

While Bayesian networks denote static knowledge, they can be made dynamic by incorporating a discrete notion of time. In dynamic Bayesian networks (Grahamani, 1998), sequences of variables are used that are indexed by a time stamp; variables with the same time stamp \( t \) form a time slice. In addition to the static dependencies between variables in the same time slice, dynamic dependencies may then be modeled by arcs between variables in different time slices. For example, if nothing happens, wet grass will dry up eventually; a power failure will probably be solved somewhere in the future; if it’s raining now, chances are high that it will be raining as well in the near future. Dynamic dependencies can be between similar or different variables in each time slice; for example the temperature in time slice \( t + 1 \) will depend on both the temperature and the state of the thermostat in slice \( t \), and vice versa. In Figure A2 the example network is enhanced with dynamic dependencies.

A BIP-Bayesian network (BIPBN) is a dynamic Bayesian network with specific connectivity, that models goal inference in the form of inverse planning (Baker et al., 2007, 2009). In a BIPBN \( D \) each slice consists of a state variable \( S_{t} \) and action variable

![Figure A1](Image)  
**Figure A1** | A Bayesian network denoting static knowledge. This BN contains four variables: \( P, S, R, \) and \( G \). These variables are dependent on each other as denoted by the arcs: \( S \) depends on \( P \), \( G \) depends on \( S \) and \( R \).

![Figure A2](Image)  
**Figure A2** | Bayesian network denoting dynamic knowledge. This BN contains four variables: \( P, S, R, \) and \( G \). These variables exist multiple times, once each time slice. As in Figure 1 their dependencies are denoted by arcs, but in addition there are also dependencies between various variables in different time slices: \( P_t \) depends on \( P_{t-1} \), etcetera. However, \( S_{t+1} \) is independent of \( S_t \) in this example.
A is A. Additionally there is a set of goal variables G that contains an arbitrary number of variables that encode the goal(s). In this framework, at a particular time t, the action A depends on the current state S and on (at least one) goal variable in G. State variables S depend on the previous state S and action variable A. See Figure A3 for a graphical illustration.

A BIPBN is based on the assumption that agents choose actions that maximize the probability that their goals are achieved. In particular, a BIPBN assumes that agents solve a Markov Decision Problem (Bellman, 1957) to achieve their goals, i.e., they pick exactly these set of actions A, which Pr(A|G) is maximal. Given this assumption, a prior probability distribution on the goals, and Bayes's rule, we can infer the probability of the agent's goals, given an observation of its actions, since Pr(G|A) = Pr(A|G) Pr(G) according to Bayes's rule.

Some proofs require the degradation of dependencies, i.e., when the dependencies are per definition required in the instance but should not have an effect on the Bayesian inference. The following lemma states how to achieve this:

**Lemma 1.** Let X be a variable and let P denote the parents of X, i.e. the set of variables on which X depends. To degrade the dependencies between X and O ≤ P we set the conditional probabilities as follows, where P' = P ∪ O:

\[\forall x \in O \text{, } Pr(X|O = o, P) = Pr(X|P)\]

**Preliminaries from complexity theory**

In computational complexity analyses one studies the amount of computational resources required to compute (or solve) a problem P: I → R, mapping inputs in the range I to the domain of outputs R. Our focus here is on the resource time. We express the time complexity of a problem as a function of the size of the input, using the Big-Oh notation O(.). A function f(x) is said to be O(g(x)), if there are constants c > 0 and a red number x such that f(x) ≤ cg(x) for all x ≥ x. Note how the function O(.) describes an asymptotic upper-bound, as it ignores constants and low-order function terms. For this reason O(.) is also called the order of magnitude.

We are interested in the time complexity of models in terms of the size of the input. The input i of a problem has size n = |i| which is the number of symbols used in a typical encoding (usually in binary). A problem P is said to be solvable in time O(g(n)) if there exists at least one algorithm that solves P in time O(g(n)). The time complexity of a problem P is measured by the fastest algorithm that solves P. We will say that a problem P is computationally intractable (for all but small inputs) if the time required to compute it grows excessively fast as a function of input size.

To make precise what we mean by “excessively fast”, we adopt a definition that is widely used in both computer science and cognitive science:

**Definition 1.** Classical (in)tractability. A problem P is said to be tractable if it can be computed in polynomial-time, i.e., time O(|i|\(^\alpha\)), where \(\alpha\) is a constant. If an problem requires super-polynomial time, e.g., exponential-time O(|i|\(^\alpha\)), where \(\alpha\) is a constant, then P is intractable.

To see why this definition has merit, compare the speed with which polynomial functions (say, \(2n, n^2, n^3\)) and an exponential functions (say, \(2^n\)) grow as a function of input size \(n\). Table A3 shows how for small \(n\), the numbers \(n^2\) and \(2^n\) do not differ much, but \(2^n\) is even smaller than \(n^3\), as \(n\) grows, \(2^n\) rockets up so fast that is is no longer plausible that a resource limited mind can perform that number of computations in a reasonable time. As reference points, consider that the number of neurons in a human brain is estimated to be 10\(^{12}\), and 10\(^{27}\) is about the number of seconds that have past since the birth of the universe. It seems highly unlikely that a human mind could perform this number of operations in only a couple of minutes (which is a generous upper bound on the time scale of most cognitive processes of interest). Even if a human mind could perform

---

**Table A1** | Complexity results for **RECEIVER**.

<table>
<thead>
<tr>
<th>RECEIVER</th>
<th>G</th>
<th>G</th>
<th>G</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>−</td>
<td>NP-hard</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>[A]</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>1 − p</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>[A], 1 − p</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
</tbody>
</table>

Cells without theorem or corollary are implied by Lemma 2.

**Table A2** | Complexity results for **SENDER**.

<table>
<thead>
<tr>
<th>SENDER</th>
<th>G</th>
<th>G</th>
<th>G</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>−</td>
<td>NP-hard</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>[A]</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>1 − p</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
<tr>
<td>[A], 1 − p</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-intractable</td>
<td>fp-tractable</td>
</tr>
</tbody>
</table>

Cells without theorem or corollary are implied by Lemma 2.
from parameterized complexity theory. This mathematical theory
complexity assessment is done using the tools and proof techniques
Definition 1, but also fixed-parameter tractability. The latter type of
computational problems.

NP-hard problems are problems that cannot\(^1\) be solved in poly-
nomial time, and hence are computationally intractable according
to above definition. A problem \(P_1\) can be proven NP-hard, by taking
a known NP-hard problem \(P_2\) and polynomial reducing it to the
problem \(P_1\). A polynomial time reduction from \(P_2\) to \(P_1\) involves the
construction of a transformation from \(P_2\) to \(P_1\) such that any
solution for the latter also implies a solution for the former. Further,
the transformation must be performable by an algorithm that runs
in polynomial time. Given such reduction we can conclude that \(P_1\)
is solvable in polynomial time only if \(P_2\) is as well. However, because
\(P_2\) is not solvable in polynomial time neither can \(P_1\) be.

The reductions in the proofs in Sections “Computational Models
and Computational Complexity Analyses” and “Parameterized
Computational Complexity Analyses” use the following NP-hard
computational problems.

\textbf{Clique}

\textbf{Input:} An undirected graph \(G = (V,E)\) where \(V\) is ordered and
\(k \in \mathbb{N} > 0\).

\textbf{Output:} Does there exist a subset \(V' \subseteq V\) such that \(|V'| = k\) and
\(\forall_{u,v \in V'} \neg((u,v) \in E)\)?

\textbf{3-Satisfiability (3SAT)}

\textbf{Input:} A tuple \((U,C)\), where \(C\) is a set of clauses on Boolean
variables in \(U\). Each clause is a disjunction of at most three
variables.

\textbf{Output:} Does there exist a truth assignment to the variables in
\(U\) that satisfies the conjunction of all clauses in \(C\)?

Our analyses not only consider (classical) tractability as in
Definition 1, but also fixed-parameter tractability. The latter type of
complexity assessment is done using the tools and proof techniques
from parameterized complexity theory. This mathematical theory
is motivated by the observation that many NP-hard problems can
be computed by algorithms whose running time is polynomial in
the overall input size \(|i|\) and non-polynomial only in one or more
small aspects of the input. These aspects are called parameters. As
the main part of the input contributes to the overall complexity in
a “good” way, and only the parameters contribute to the overall
complexity in a “bad” way, the problem is well-solved even for
large inputs provided only that the parameters remain small. This
intuitive characterization is captured by the formal notion of fixed-
parameter tractability (see also Downey and Fellows, 1999).

\textbf{Definition 2. Fixed-parameter (in)tractability. A problem \(P\) is said to
be fixed-parameter (fp)tractable for parameter set \(K = \{k_1, k_2, \ldots, k_m\}\)
if there exists at least one algorithm that computes \(P\) for any input of
size \(n\) in time \(f(k_1, k_2, \ldots, k_m) n^\alpha\) where \(f(\cdot)\) is an arbitrary computable
function and \(\alpha\) is a constant. If no such algorithm exists then \(P\) is said
to be fixed-parameter (fp)intractable for parameter set \(K\).

Proving fixed-parameter tractability is conceptually straightforward:
It suffices to produce just one algorithm that computes the
problem in fixed-parameter tractable time (see, e.g., Sloper and Telle,
2008, for a review of generic techniques for building such algorithms).

\(W[1]\)-hard problems are problems, including a set of param-
ters, that cannot\(^2\) be solved in fixed parameter time even when all
parameters in their set are small, and hence are fixed-parameter
intractable according to above definition. A problem \(K_1-P_1\) can be
proven \(W[1]\)-hard, by taking a known \(W[1]\)-hard problem \(K_2-P_2\)
and parameterized reducing it to the problem \(K_1-P_1\). A param-
eterized reduction from \(K_1-P_1\) to \(K_2-P_2\) involves the construction of a
transformation from \(K_1-P_1\) to \(K_2-P_2\) such that any solution
for the latter also implies a solution for the former. Further, the
transformation must be performable by an algorithm that runs
in fixed parameter tractable time, and all parameters in \(K_2\) must
be a function of a parameter in \(K_1\). Given such reduction we can
consider that \(K_2-P_2\) is solvable in fixed parameter tractable time
only if \(K_1-P_1\) is as well. However, because \(K_1-P_1\) is not solvable in
fixed parameter tractable time neither can \(K_2-P_2\) be.

The proofs in the Sections “Computational Models
and Computational Complexity Analyses” and “Parameterized
Computational Complexity Analyses” use the following \(W[1]\)-hard
parameterized problem.

\textbf{k-Clique}

\textbf{Input:} A undirected graph \(G = (V,E)\) where \(V\) is ordered and
\(k \in \mathbb{N} > 0\).

\textbf{Parameters:} \(k\), the size of the desired clique.

\textbf{Output:} Does there exist a subset \(V' \subseteq V\) such that \(|V'| = k\) and
\(\forall_{u,v \in V'} \neg((u,v) \in E)\)?

Finally, the following Lemma is used to propagate fp-(in)tractability
results derived for one parameter set \(K\) to another \(K'\).

\textbf{Lemma 2. Let \(P\) be a problem that is fp-intractable for the param-
ter set \(K = \{k_1, \ldots, k_m\}\), then \(P\) is also fp-intractable for any subset
\(K' \subseteq K\). Conversely, let \(P\) be a problem that is fp-tractable for
the parameter set \(K = \{k_1, \ldots, k_m\}\), then \(P\) is also fp-tractable for any
super-set \(K' \supseteq K\).

\textsuperscript{2}This is true, assuming that FPT \(\neq W[1]\) (Downey and Fellows, 1999).

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{n} & \textbf{2n} & \textbf{n²} & \textbf{n³} & \textbf{2^n} \\
\hline
2 & 4 & 4 & 8 & 4 \\
5 & 10 & 25 & 125 & 32 \\
10 & 20 & 100 & 1000 & 1024 \\
20 & 40 & 400 & 8000 & 1048576 \\
50 & 100 & 2500 & 125000 & \textgreater 10^{15} \\
100 & 200 & 10000 & 1000000 & \textgreater 10^{15} \\
200 & 400 & 40000 & 8000000 & \textgreater 10^{15} \\
\hline
\end{tabular}
\caption{Illustration of how polynomial growth rates \((n, 2n, n^2, n^3)\)
compare with an exponential growth rate \(2^n\).}
\end{table}

---

\textsuperscript{1}This is true, assuming that the class of problems solvable in polynomial time (P) is not equal to the class of problems whose solutions can be verified in polynomial time (NP). This “P \(\neq\) NP” conjecture is believed by most living mathematicians, both on theoretical and empirical grounds (for more details see Garey and Johnson, 1979; Aaronson, 2005; Fortnow, 2009).
Theorem A. Receiver is NP-hard.

Proof. Given an instance \( (G = (V, E)) \) of Clique, construct an instance \( (B, a, s) \) of Receiver as follows (see Figure A5 for an example):

1. Assume an arbitrary order on the vertices in \( V \) such that \( V = V_1 < V_2 < \ldots < V_{|V|} \).
2. Assume the basic structure of \( B \) as in the definition of Receiver and create \( 1 + (k - 1) + (k(k - 1)/2) \) Boolean state variables \( S_{0,1}, \ldots, S_{k-1}\), \( k\) Boolean action variables \( A_1, \ldots, A_k \), and \( k \) Boolean goal variables \( G = G_1, \ldots, G_{k|V|} \) such that \( G_1 = \emptyset \) and \( G_{i+1} = G_i \cup \{v_i\} \), where \( v_i \) is an arbitrary conditional probability distribution over \( |V| \) given \( \emptyset \) and \( G_i \) \( \emptyset \).
3. Set \( S_0 = \text{true} \) and for \( 1 \leq i \leq (k - 1) + (k(k - 1)/2) \), let \( S_i \) depend on \( S_{i-1} \) and \( A_{i-1} \) and have the following conditional probability:

\[
\Pr(S_i | S_{i-1}, A_{i-1}) = \begin{cases} 1 & \text{if } S_{i-1} = \text{true} \text{ and } A_{i-1} = \text{true} \\ 0 & \text{otherwise} \end{cases}
\]

These state variables effectively function as conjunctions which ensure that there is some assignment \( g \) to \( G \) such that \( \Pr(g) > 0 \) iff. all action variables are set to \( \text{true} \).

4. All dependencies between \( S_i \) and \( A_i \) are degraded as defined in Lemma 1.
5. For \( 1 \leq i < k \), let \( A_i \) depend on \( B_i, B_{i+1}, \) and \( S_i \) and have the following conditional probability:

\[
\Pr(A_i = \text{true} | B_i, B_{i+1}, S_i) = \begin{cases} 1 & \text{if } v(B_i) < v(B_{i+1}) \text{ and } S_i = \text{true} \\ 0 & \text{otherwise} \end{cases}
\]

These action variables ensure that a joint value assignment \( g \) to \( G \) has \( \Pr(g) > 0 \) iff. the values encoded in the Boolean blocks are distinct.
6. For \( k \leq i \leq (k(k-1)/2) \), let \( A_i \) depend on a distinct pair of goal-variable blocks \( (B_{p_i}, B_{q_i}) \) and \( S_i \), and have the following conditional probability:

\[
Pr\left( A_i = \text{true} \mid B_{p_i}, B_{q_i}, S_i \right) = \begin{cases} 
1 & \text{if } (v(B_{p_i}) \neq v(B_{q_i})) \in E \text{ and } S_i = \text{true} \\
0 & \text{otherwise}
\end{cases}
\]

These action variables ensure that a joint value assignment \( g \) to \( G \) has \( Pr(g) > 0 \) iff all pair of values encoded in the Boolean blocks is an edge in \( E \) from the clique instance.

7. Make the prior probability distribution for each goal variable uniform.

8. Set \( a \) and \( s \) such that all action and state variables are assigned \text{true}.

As the number of conditional probability tables that are constructed by the reduction is proportional to the total number of variables \( G \) plus \( \log |V| \) and each table involves at most \( 1 + \log_2 |V| \) Boolean variables (resulting in a maximum of \( 2^\log_2 |V| \) Boolean variables in the size of the given instance of Clique). Moreover, note that in this instance of \( Receiver |A| = (k - 1) + \lfloor k(k-1)/2 \rfloor \) and \( |G| = 0 \).

To prove that the construction above is a valid reduction, we must show that the answer to the given instance of Clique is “Yes” iff there exists a solution to the constructed Receiver instance unequal to \( \emptyset \). We will show this we proved both directions of this implication separately:

- If the answer to the given instance of Clique is “Yes”, then there exists a subset \( V' \subseteq V \) such that \( |V'| = k \) and \( \forall_{u \neq v} \forall_{(u, v) \in E} \).

Let \( g \) be the assignment to \( G \) corresponding to the vertices in \( V' \) and \( V \). As the vertices in \( V' \) are distinct, action variables \( A_1, \ldots, A_{k-1} \) will all be \text{true} with probability \( 1 \) relative to \( G \). Moreover, as \( V' \) is a \( k \)-clique and there is an edge between each pair of distinct vertices in \( V' \), action variables \( A_1, \ldots, A_{k-1} \) will all be \text{true} with probability \( 1 \) relative to \( G \). Given the structure of \( B \), this implies that \( Pr(g) > 0 \), which means that the answer to the constructed instance of Receiver is not empty.

- If the answer to the constructed instance of Receiver is not empty, then there is an assignment \( g \) to \( G \) such that \( Pr(g) > 0 \). Given the dependencies in and the conditional probabilities in \( B \), this can only happen if all action variables have value \text{true} with probability \( 1 \). Hence, the values of the goal variables in \( g \) are not only distinct (by the conditional probabilities in (5) above), but the values encoded in the Boolean blocks also correspond to a set of vertices such that each distinct pair of these vertices is connected by an edge in \( G \) (by the conditional probabilities in (6) above), which means that the answer to the given instance of Clique is “Yes.”

\[ \square \]

**Theorem B.** Sender is NP-hard.

**Proof.** Observe that in the Receiver instance constructed in the proof of Theorem A, \( Pr(G_c = g_c \mid A = a, s) \) for any joint value assignment to \( a \) to \( A \) and \( s \) to \( S \) other than the assignment that sets all variables in \( A \) and \( S \) to \text{true}. If the output of Sender is not \( \emptyset \), then there exists a joint value assignment \( a \) and \( s \) such that Receiver \( (B, a, s) \neq \emptyset \) for \( g_c \). By definition of the reduction \( g_c \), then corresponds to a clique in \( G \). This reduces Clique to Sender and thus Sender is NP-hard.

**PARAMETERIZED COMPUTATIONAL COMPLEXITY ANALYSES**

**Theorem C.** Receiver is \( fp \)-tractable for parameter set \( \{|A|, |G|\} \).

**Proof.** To calculate the output for for any instance of Receiver one can try out all possible joint value assignments to \( G \cup G_c \) in a time that is only exponential in \( |G| \) and \( |G_c| \) (viz. time \( O(|G| \log |G|) \), where \( v \) is the maximum number of possible values per goal variable). As the values of \( A \) and \( S \) are given, the conditional probability for each goal value assignment can be computed in a time polynomial in the total number of variables, because all variables are observed. The computational complexity of Receiver is exponential only in the number of instrumental \( G \) and communicative goals \( G_c \), thus Receiver is \( fp \)-tractable for the parameter set \( \{|A|, |G|\} \).

**Corollary A.** Receiver is \( fp \)-intractable for parameter set \( \{|A|, |G|\} \).

**Proof.** To prove that Receiver is \( fp \)-intractable for parameter set \( \{|A|, |G|\} \), it suffices to provide a parameterized reduction from a \( \text{W}[1] \) \(-hard problem, namely \( k \)-Clique, to \( \{|A|, |G|\} \)-Receiver. The reduction is exactly the same as the proof of Theorem A. That reduction runs in polynomial time and thus also in fixed-parameter tractable time. Furthermore, \( |A| \) is a function of the size \( k \) of the requested clique and \(|G| = 0\). As \( k \)-Clique is \( \text{W}[1] \)-hard, \( \{|A|, |G|\} \)-Receiver is also \( \text{W}[1] \)-hard and thus Receiver is \( fp \)-intractable for parameter set \( \{|A|, |G|\} \).

**Theorem D.** Receiver is \( fp \)-intractable for parameter set \( \{|A|, |G|\}, 1 - \rho \).

**Proof.** In this proof we show that Receiver remains NP-hard even if we allow the probability of the most probable explanation to be as high as \( 1 - \rho \) for arbitrarily small values of \( \rho \). In this proof \(|A| \) is a function of of the size \( k \) of the requested clique and \(|G| = 0\). This proves Receiver is \( fp \)-intractable for the parameter set \( \{|A|, |G|\}, 1 - \rho \), where \( 1 - \rho \) is the probability of the most probable explanation for \( a \) and \( s \) a communicator considers.

Consider a variant of the reduction in the proof of Theorem A. In addition to \( G_1, \ldots, G_{\lfloor \log |V| \rfloor} \), we create an extra Boolean goal variable \( G_x \). In this variant \( G_i = G_{i-1}, \ldots, G_{\lfloor \log |V| \rfloor} \) and \( G_c = \{G_x\} \). As before, \( G_i \) is divided into \( k \) blocks of \( \lceil \log |V| \rceil \) Boolean goal variables \( B_1, \ldots, B_k \). The conditional probabilities of the action variables are as follows, where \( \alpha = (1/|A|) \cdot \varepsilon \):

- For \( 1 \leq i < k \) let \( A_i \) depend on \( B_i \) \( B_{i+1} \) and \( S_i \) and have the following conditional probability:

\[
Pr(A_i = \text{true} \mid B_{i}, B_{i+1}, S_i, G_x) = \begin{cases} 
1 & \text{if } (v(B_i) \neq v(B_{i+1}), S_i = \text{true} \text{and } G_x = \text{true} \\
0 & \text{if } (v(B_i) = v(B_{i+1}), S_i = \text{true} \text{and } G_x = \text{false} \\
\alpha & \text{if } (v(B_i) = v(B_{i+1}), S_i = \text{true} \text{and } G_x = \text{true} \\
\alpha & \text{if } (v(B_i) \neq v(B_{i+1}), S_i = \text{true} \text{and } G_x = \text{false} \\
0 & \text{otherwise}
\end{cases}
\]

\[ \square \]
These action variables ensure that a joint value assignment \( g \) to \( G \) has \( \Pr(g) > 0 \) iff. the values encoded in the Boolean blocks are distinct.

- For \( k \leq i \leq [k(k - 1)/2] \), let \( A_i \) depend on a distinct pair of goal-variable blocks \( (B_{p_i}, B_{q_i}) \) and \( S_i \) and have the following conditional probability:

\[
\Pr(A_i = \text{true} | B_{p_i}, B_{q_i}, S_i, G_x) = \\
\begin{cases} 
1 & \text{if } (v(B_{p_i}), v(B_{q_i})) \in E, \quad S_i = \text{true and } G_x = \text{true} \\
0 & \text{if } (v(B_{p_i}), v(B_{q_i})) \in E, \quad S_i = \text{true and } G_x = \text{false} \\
\alpha & \text{if } (v(B_{p_i}), v(B_{q_i})) \notin E, \quad S_i = \text{true and } G_x = \text{true} \\
0 & \text{if } (v(B_{p_i}), v(B_{q_i})) \notin E, \quad S_i = \text{true and } G_x = \text{false} 
\end{cases}
\]

These action variables ensure that a joint value assignment \( g \) to \( G \) has \( \Pr(g) > 0 \) iff. the values encoded in the Boolean blocks is an edge in \( E \) from the clique instance.

If \( G \) has a \( \tilde{k} \)-clique and all action and state variables are observed to be true, then \( \Pr(G_x = \text{true}) = 1 - 1|A| = 1 - |A|/|A| = 1 - \varepsilon. \) If \( G \) does not have a \( \tilde{k} \)-clique, then \( \Pr(G_x = \text{false}) = 1. \) Hence, even if the probability of the most probable value of \( G_x \) is at least \( 1 - \varepsilon, \) it is still NP-hard to decide on the most probable value. \( \Box \)

**Corollary B.** **Receiver** is fp-tractable for the parameter set \( |I|_G|, 1 - p]. \)

**Proof.** The **Receiver** problem is in fact a special case of the more general **PARTIAL MAP** problem, where the input is an arbitrary Bayesian network \( B \), in which the set of variables is partitioned into a set of observed variables (called the evidence set \( E \)), a set of variables for which the most probable explanation is sought (called the explanation set \( M \)), and a set of variables that are neither in the evidence nor explanation set (called the intermediate set \( I \)). In the Bayesian network in **Receiver**, the action and state variables \( A \) and \( S \) are observed and form the evidence set \( E \), the explanation set \( M \) consists of the communicative goal variables \( G_x \), and the instrumental goal variables form the intermediate set \( I \). **PARTIAL MAP** is NP-hard in general and remains so under severe constraints on the structure of the network; however, **PARTIAL MAP** is tractable when both the probability of the most probable explanation is high, and the number of variables in the intermediate set is low (Park and Darwiche, 2004).

Because of the inheritance from the **PARTIAL MAP** problem, the **Receiver** problem is tractable when both the probability \( p \) of the most probable joint value assignment to the goal variables is high and when \( |I|_G| \), the number of instrumental goal variables, is bounded, i.e., **Receiver** is fixed parameter tractable for \( |I|_G|, 1 - p]. \)

**Observation 1.** If **Receiver** is intractable, then **Sender** is also intractable.

**Corollary C.** **Sender** is fp-intractable for parameter set \( |I|_A, |I|_G| \).

**Proof.** Follows from **Corollary A** and Observation 1.

**Theorem E.** **Sender** is fp-tractable for parameter set \( |I|_A, |I|_G|, |I|_C \).

**Proof.** To calculate the output of **Sender** we can try out all possible joint value assignments \( a \) to \( A \). For all \( a \) we have to calculate the **Receiver** \( (B, a, s) \) output \( g_c \). The correct output is that \( a \) for which \( \Pr(A = a | G_1 = g_1) = 1 \) is maximal. This takes time \( O(|A|) \), where \( v \) is the maximum number of possible values per goal variable, because **Receiver** is fp-tractable for parameter set \( |I|_A, |I|_G| \) (Theorem C). The computational complexity of **Sender** is only exponential in the number of actions \( |A| \) when \( |I|_G| \) and \( |I|_C \) are small, thus **Sender** is fp-tractable for the parameter set \( |I|_A, |I|_G|, |I|_C \).

**Corollary D.** **Sender** is fp-intractable for parameter set \( |I|_A, |I|_G|_1, 1 - p \).

**Proof.** Follows from **Theorem D** and Observation 1.

**Theorem F.** **Sender** is fp-intractable for parameter set \( |I|_G| \), \( |I|_G|_1, 1 - p \).

**Proof.** The proof of this theorem uses a variant of the proof construction by Kwisthout (2009), using a reduction from 3SAT to prove NP-hardness of **PARAMETER TUNING** restricted to polytrees; which on its turn was inspired by a similar construction by Park and Darwiche (2004), who proved NP-hardness of **PARTIAL MAP** restricted to polytrees. The proof uses a network construction as in **Figure A6**, in which all \( A_i \) model the variables of the 3SAT formula, all \( S_i \) model the clauses and \( S_j \) acts as a clause selector. The conditional probabilities are constructed such that Bayesian inference on the network solves 3SAT.

Let \( m = |I|_C \) denote the number of clauses of a 3SAT formula and let \( n = |I|_A \) denote the number of variables. The conditional probabilities of the variables \( S_1, ..., S_m \) are such that \( \Pr(S_{i+1} = \text{true}) = m/n \) if and only if there is a joint value assignment to the variables \( A_i, ..., A_m \) that corresponds to a satisfying truth assignment to the 3SAT formula, or \( \Pr(S_{i+1} = \text{true}) \leq (m - 1)/n \)

\( ^{A \text{polytree is a directed acyclic graph for which there are no undirected cycles when the arc direction is dropped.}} \)
otherwise. We add binary variables $S_{n+1}$, $G$, $G_c$, and $A_i$, dummy variables $A_{i1}$ and $S_n$, with the following probability distributions (see Figure A7):

- $\Pr(S_{n+1} = true \mid S_n = true) = 1 - \varepsilon + \left(1 - \frac{m - 1/2}{n}\right) \times \alpha$

- $\Pr(S_{n+1} = true \mid S_n = false) = 1 - \varepsilon - \frac{m - 1/2}{n} \times \alpha$

where $\alpha$ is a sufficiently small number to guarantee that all probabilities are in $[0,1]$. It plays no further role in the proof, so we fix $\alpha = \varepsilon^2$.

- $\Pr(G = true) = 1$ and $\Pr(A \mid G) = \text{uniformly distributed}$;

- $\Pr(G_c = true) = \Pr(G = false) = 1/2$;

- $\Pr(A_c = true \mid S_{n+1} = true, G_c = false) = 1$, $\Pr(A_c = true \mid S_{n+1} = false, G_c = false) = \varepsilon$, and $\Pr(A_c = true \mid S_{n+1} = true, G_c = true) = 1/2$;

- Dependencies between $G, S_n, S_{n+1}$, and $A_i$ are degraded;

- Dependencies between $S_{n+1}, A_i$, and $S_n$ are degraded;

- All dependencies between $S_i$ and $A_i$ are degraded as defined in Lemma 1.

Using this reduction from 3SAT instances to SENDER we will prove that any joint value assignment $a$ to $A$ that maximizes the probability of an arbitrary joint value assignment $g$ to $G_1 \cup G_2$ (i.e. a solution to SENDER) also is a solution for the 3SAT instance, even if the probability of that joint value assignment is at least $1 - \varepsilon$.

The following now holds:

$$
\Pr(S_{n+1} = true) = \Pr(S_{n+1} = true \mid S_n = true) \times \Pr(S_n = true)
+ \Pr(S_{n+1} = true \mid S_n = false) \times \Pr(S_n = false)
= \left(1 - \varepsilon + \left(1 - \frac{m - 1/2}{n}\right) \times \alpha\right) \times \Pr(S_n = true)
+ \left(1 - \varepsilon - \frac{m - 1/2}{n} \times \alpha\right) \times (1 - \Pr(S_n = true))
= 1 - \varepsilon + \Pr(S_n = true) \times \alpha - \frac{m - 1/2}{n} \times \alpha
$$

Recall that $\Pr(S_n = true) = mn/n$ if the 3SAT formula is satisfiable, and at most $(m-1)/n$ otherwise. Hence, $\Pr(S_{n+1} = true) > 1 - \varepsilon$ if the 3SAT formula is satisfiable, and $\Pr(S_{n+1} = true) < 1 - \varepsilon$ if it is not satisfiable. As we fixed $\alpha = \varepsilon^2$ and given that $m \leq n$ by definition, we have in particular that:

$$1 - 2\varepsilon < \Pr_{\text{UNSAT}}(S_{n+1} = true) < 1 - \varepsilon < \Pr_{\text{SAT}}(S_{n+1} = true) < 1$$

Where $\Pr_{\text{SAT}}(S_{n+1} = true)$ and $\Pr_{\text{UNSAT}}(S_{n+1} = true)$ denote the probability that $S_{n+1} = true$ given that the 3SAT formula is respectively satisfiable or not. Observe that the posterior probability of $G_i$ is independent of the value assignment of the action variables $A_1...A_n$. Given this independence, if $A_c = true$ then $\Pr(G_c = true) > 1 - \varepsilon$ independent of $\Pr(S_{n+1} = true)$. Also if $A_c = false$ and $\Pr(S_{n+1} = true) > 1 - \varepsilon$ then $\Pr(G_c = false) > 1 - \varepsilon$; and if $A_c = false$ and $\Pr(S_{n+1} = true) < 1 - \varepsilon$ then $\Pr(G_c = false) < 1 - \varepsilon$. Observe that in each of these cases the most probable value assignment to $G_i$ has a probability which is larger than $1 - \varepsilon$, however, the most probable assignment to $G_i$ flips from true to false depending on the satisfiability of the 3SAT instance.

Thus, if there exists a value assignment $a$ to $A = \{A_1, ..., A_n\} \cup A_c$ such that $g = \{G_1 = true, G_c = false\}$ is the most probable explanation to $\{G,G_c\}$, then the 3SAT instance is satisfiable (viz. those variables in $U$ corresponding to the $A_1...A_n$ are true iff. $A_c = true$ and false otherwise). Likewise, if the 3SAT instance is satisfiable then there is a value assignment to $A$, with $A_c = false$, such that $g = \{G_1 = true, G_c = false\}$ is the most probable explanation to $\{G_1, G_c\}$.

This proves SENDER $\text{FP}$-intractable for the parameter set $\{|G_1|, |G_c|, 1 - \rho\}$, where $1 - \rho$ is the minimum probability of all most probable explanations for all possible $a$ and $s$ a communicator considers. □

REFERENCES