Games for Learning

A Sabotage Approach

Nina Gierasimczuk, Lena Kurzen and Fernando R. Velázquez-Quesada
Institute for Logic, Language and Computation
Universiteit van Amsterdam
Email: {N.Gierasimczuk | L.M.Kurzen | F.R.VelazquezQuesada}@uva.nl

Abstract—**In formal approaches to inductive learning, the ability to learn is understood as the ability to single out a correct hypothesis from a range of possibilities. Although most of the existing research focuses on the characteristics of the learner, in many paradigms the significance of the teacher's abilities and strategies is in fact undeniable. Motivated by this observation, in this paper we highlight the interactive nature of learning by proposing a game-theoretical and logical approach. We consider learning as a *sabotage-type* game between *Teacher* and *Learner*, and present different variants based on the level of cooperativeness and the actions available to the players. We characterize the existence of a winning strategy in such games by formulas of *Sabotage Modal Logic*, analyzing also their complexity. Our work constitutes the first step towards a unified game-theoretical and logical approach to formal learning theory.**

I. Introduction

The objective of this paper is to investigate how logics for interaction in multi-agent systems can be used to reason about strategic abilities and information flow during the *learning* process. Formal learning theory (see e.g. [4]) is concerned with the process of inductive inference: it formalizes the process of inferring general conclusions from partial, consecutively given information, as in the case of language learning (inferring grammars from sentences) and scientific inquiry (drawing general conclusions from partial experiments). We can think of this general process as a game between two players: *Learner* and *Teacher*. The game starts with a class of possible worlds from which Teacher chooses the actual one, and Learner has to find out which one it is. Teacher provides information about the world in an inductive manner, and whenever Learner receives a piece of information, he picks a conjecture from the initial class, indicating which one he thinks is the case. Several conditions can be defined for the success of the learning process: we can require that Learner arrives at a correct hypothesis (finite identification), or that the sequence of Learner's conjectures converges to a correct hypothesis (identification in the limit) [3].

We give a high-level analysis of the described process. First, we treat learning as a procedure of singling out one correct hypothesis from a range of possibilities. Second, we see this procedure not as a one-move choice; instead, we allow many steps of update before the conclusion is reached. These two properties make our notion of learning different from the concept of learning formalized as epistemic update in *Dynamic Epistemic Logic* (see e.g. [2]), where the word "learning" is often used as a synonym of "getting to know" and is usually represented as a one-step epistemic update. Moreover, in our approach we pay attention to the strategies for teaching, highlighting the fact that restricted power and knowledge of the learner can be compensated by additional insights and intentions of the teacher.

The paper is structured as follows. Section II introduces the framework of learning as Sabotage Games, shows how sabotage modal logic can express the existence of winning strategies in three different versions of Sabotage Learning Games and gives complexity results for them. Section III analyzes Sabotage Learning Games in which the players do not need to move in alternation. Section IV presents a refined interactive view on teaching based on existing learning algorithms. Section V concludes.

II. Learning as a Sabotage Game

Our work is motivated by the *learning from queries and counterexamples* model [1]. In that paradigm, the goal of Learner is to recognize an initially unknown language $\mathcal{L}$. In order to do this, he is allowed to ask Teacher two types of questions: about the membership of a certain string to $\mathcal{L}$, and about the equivalence of his conjecture (another language) to $\mathcal{L}$. When answering those questions, Teacher does not have any freedom — her responses are restricted by $\mathcal{L}$. However, a negative answer to the second question is accompanied by a counterexample, which plays the role of a hint for Learner. This is the only point of the procedure in which Teacher has a relative freedom of choice, and in fact, the informativeness of the string given as a counterexample influences the effectiveness of the learning process. We want to focus on this aspect of learning and show that the "profile" of Teacher is relevant for the learning process. We consider several possible scenarios — we describe games in which Teacher is either helpful or unhelpful, and Learner is either eager or unwilling to learn.

Let us consider a very simple "classroom" situation with one teacher and one learner. From our high-level

TABLE I
CORRESPONDENCE WITH LEARNING MODEL

Learning Model	Sabotage Games
hypotheses	states
correct hypothesis	goal state
possibility of a mind change from hypothesis a to hypothesis b	edge from state a to b
a mind change from hypothesis a to hypothesis b	transition from state a to b
giving a counterexample that eliminates the possibility of a mind change from a to b	removing a transition between a and b

perspective, learning is a step-by-step process through which Learner changes his information state, and the process is successful if he eventually reaches a state representing the goal. The information Teacher provides can be seen as feedback about Learner's current conjecture, allowing him to rule out possible changes of mind. We can represent the situation as a graph whose vertices represent Learner's possible information states and edges stand for transitions between them. During the learning process, Learner can change his information state by moving along the edges and Teacher can cut off edges, thereby preventing Learner from making certain transitions. One state is associated with the learning goal: if Learner reaches it, we say that the learning process has been successful. The correspondence between the learning model from formal learning theory and our proposal is described in Table I.

Observe that in learning from queries and counterexamples, a counterexample results in the absolute removal of some initially possible hypothesis. Our setting generalizes this idea: the removal of a transition need not make the target vertex unreachable.

A. Sabotage Games

Our perspective on learning leads naturally to the framework of Sabotage Games [6], [11]. A sabotage game is played in a directed multi-graph, with two players, *Runner* and *Blocker*, moving in alternation with Runner being the first. Runner moves by making a single transition from the current vertex; Blocker moves by deleting a *single* edge from *any* part of the graph. We begin by defining the structure in which a Sabotage Game takes place.

Definition 2.1 ([6]): A *directed multi-graph* is a tuple $G = (V, E)$ where V is a set of vertices and $E : V \times V \to \mathbb{N}$ is a function indicating the number of edges between any two vertices.

The Sabotage Game is defined as follows.

Definition 2.2 ([6]): A *Sabotage Game* $SG = \langle V, E, v, v_g \rangle$ is given by a directed multi-graph (V, E) and two vertices $v, v_g \in V$. Vertex v represents the position of Runner and v_g represents the goal state.

Each match is played as follows: the initial position $\langle E_0, v_0 \rangle$ is given by $\langle E, v \rangle$. Round $k+1$ from position $\langle E_k, v_k \rangle$ consists of Runner moving to some v_{k+1} such that $E(v_k, v_{k+1}) > 0$, and then Blocker removing an edge (v, v') such that $E_k(v, v') > 0$. The new position is $\langle E_{k+1}, v_{k+1} \rangle$, where $E_{k+1}(v, v') := E_k(v, v') - 1$ and, for every $(u, u') \neq (v, v')$, $E_{k+1}(u, u') := E_k(u, u')$. The match ends if a player cannot make a move or if Learner reaches the goal state, which is the only case in which he wins.

Remark 1: It is easy to see that Sabotage Games have the *history-free determinacy property*: if one of the players has a winning strategy then she has a winning strategy that depends only on the current position. Then, each round can be viewed as a transition from a Sabotage Game $SG = \langle V, E_k, v_k, v_g \rangle$ to another Sabotage Game $SG' = \langle V, E_{k+1}, v_{k+1}, v_g \rangle$, since previous moves become irrelevant. We will use this fact through the whole paper. Also, by edges and vertices of $SG = \langle V, E, v, v_g \rangle$, we will mean edges and vertices of its underlying directed multi-graph (V, E).

In this definition of the Sabotage Game, Blocker removes an edge between two states v, v' by decreasing the value of $E(v, v')$ by 1. As we will see later, this definition of the game based on the above definition of multi-graphs can lead to some technical problems when transforming such a graph into a Sabotage Model. Therefore, we will now present an alternative definition, which we later show (Theorem 1) to be equivalent with respect to the existence of a winning strategy.

Definition 2.3: Let $\Sigma = \{a_1, \ldots a_n\}$ be a finite set of labels. A *directed labelled multi-graph* is a tuple $G^\Sigma = (V, \mathcal{E})$ where V is a set of vertices and $\mathcal{E} = (\mathcal{E}_{a_1}, \ldots, \mathcal{E}_{a_n})$, where $\mathcal{E}_{a_i} \subseteq V \times V$ for each $a_i \in \Sigma$.
In this definition, labels from Σ are used to represent multiple edges between two vertices; $\mathcal{E}$ is simply an ordered collection of binary relations on V with labels from Σ. Then, the definition of the game is as follows.

Definition 2.4: A *Labelled Sabotage Game* $SG^\Sigma = \langle V, \mathcal{E}, v, v_g \rangle$ is given by a directed labelled multi-graph $(V, \mathcal{E})$ and two vertices $v, v_g \in V$. Vertex v represents the position of Runner and v_g represents the goal state.

Each match is played as follows: the initial position $\langle \mathcal{E}^0, v_0 \rangle$ is given by $\langle \mathcal{E}, v \rangle$. Round $k+1$ from position $\langle \mathcal{E}^k, v_k \rangle$ with $\mathcal{E}^k = (\mathcal{E}^k_{a_1}, \ldots, \mathcal{E}^k_{a_n})$, consists of Runner moving to some v_{k+1} such that $(v_k, v_{k+1}) \in \mathcal{E}^k_{a_i}$ for some $a_i \in \Sigma$, and then Blocker removing an edge $((v, v'), a_j)$, where $(v, v') \in \mathcal{E}^k_{a_j}$ for some $a_j \in \Sigma$. The new position is $\langle \mathcal{E}^{k+1}, v_{k+1} \rangle$, where $\mathcal{E}^{k+1}_{a_j} = \mathcal{E}^k_{a_j} \setminus \{(v, v')\}$ and $\mathcal{E}^{k+1}_{a_i} = \mathcal{E}^k_{a_i}$ for all $i \neq j$. The match ends if a player cannot make a move or if Runner reaches the goal state, with him winning only in the last case.
What is said in Remark 1 also holds for Labelled Sabotage Games.

In this definition of the game, it is easy to see that when Blocker removes an edge from v to v', it is irrelevant what is the label of the removed edge; what matters for the existence of a winning strategy is the number of edges from v to v' that are left.

Observation 1: Let $SG^{\Sigma} = \langle V, \mathcal{E}, v_0, v_g \rangle$ and $SG'^{\Sigma} = \langle V, \mathcal{E}', v_0, v_g \rangle$ be two Labelled Sabotage Games that differ only in the labels of their edges, that is,

$$\forall (v, v') \in V \times V : |\{\mathcal{E}_{a_i} \mid (v, v') \in \mathcal{E}_{a_i}\}| = |\{\mathcal{E}'_{a_i} \mid (v, v') \in \mathcal{E}'_a\}|,$$

where $|\cdot|$ stands for cardinality. Then Runner has a winning strategy in SG^{Σ} iff he has a winning strategy in SG'^{Σ}.

We will now show that the problems of deciding whether Runner has a winning strategy in each of the Sabotage Games SG and SG^{Σ} are polynomially equivalent. We start by formalizing the problems.

Definition 2.5: The decision problem SABOTAGE is defined as follows.

- **INPUT**: A Sabotage Game $SG = \langle V, E, v_0, v_g \rangle$.
- **QUESTION**: Does Runner have a winning strategy in SG?

Definition 2.6: The decision problem Σ-SABOTAGE is defined as follows.

- **INPUT**: A Sabotage Game on a labelled multi-graph $SG^{\Sigma} = \langle V, \mathcal{E}, v_0, v_g \rangle$.
- **QUESTION**: Does Runner have a winning strategy in SG^{Σ}?

Theorem 1: SABOTAGE and Σ-SABOTAGE are polynomially equivalent.

Proof: We show that the problems can be polynomially reduced to each other.

First we show that SABOTAGE can be reduced to Σ-SABOTAGE. Given a Sabotage Game $SG = \langle V, E, v_0, v_g \rangle$, let $m := \max\{E(u, u') \mid (u, u') \in (V \times V)\}$. Define the Labelled Sabotage Game $f(SG) := \langle V, \mathcal{E}, v_0, v_g \rangle$ where $\mathcal{E} := (\mathcal{E}_1, \ldots, \mathcal{E}_m)$ and each $\mathcal{E}_i$ is given by $\mathcal{E}_i := \{(u, u') \in V \times V \mid E(u, u') \geq i\}$.

We show that Runner has a winning strategy (w.s.) in SG iff he has one in $f(SG)$. The proof is by induction on $n = \sum_{(v,v') \in V \times V} E(v, v')$, which is the number of edges of SG. Note that by definition of f, $n = \sum_{i=1}^{i=m} |\mathcal{E}_i|$, that is, $f(SG)$ has the same number of edges.

The base case is straightforward since in both games Runner has a w.s. iff $v_0 = v_g$. For the inductive case, from left to right, suppose Runner has a w.s. in the game $SG = \langle V, E, v_0, v_g \rangle$ with $n+1$ edges. Then, there is some $v_1 \in V$ such that $E(v_0, v_1) > 0$ and Runner has a w.s. for all games $SG' = \langle V, E', v_1, v_g \rangle$ that result from Blocker removing any edge (u, u') with $E(u, u') > 0$. Note that all such games SG' have just n edges, so by induction hypothesis Runner has a w.s. in $f(SG')$. But then, by Observation 1, Runner has also a w.s. in all games $f(SG)'$ that result from removing an arbitrary edge from $f(SG)$, because for any removed edge (u, u'), the only possible difference between $f(SG')$ and $f(SG)'$ is in the labels of the edges between u and u' (in $f(SG')$ the removed label was the largest, in $f(SG)'$ the removed label is any). Now, by definition of f, choosing v_1 is also a legal move for Runner in $f(SG)$ and, since he can win every $f(SG)'$, he has a w.s in $f(SG)$.

From right to left, Runner having a w.s. in $f(SG)$ means that he can choose some v_1 with $(v_0, v_1) \in \mathcal{E}_i$ for some $i \leq m$ such that he has a w.s. in all games $f(SG)'$ resulting from Blocker's move. Choosing v_1 is also a legal move of Runner in SG. Suppose that Blocker replies by choosing (v, v'). Let us call the resulting game SG'. By assumption and Observation 1, Runner also has a w.s. in the game $f(SG')$ which is the result from Blocker choosing $((v, v'), E(v, v'))$. Since $f(SG)' = f(SG')$, we can apply the inductive hypothesis.

Let us see now how SG^{Σ} can be polynomially reduced to SG. Given $SG^{\Sigma} = \langle V, \mathcal{E}, v, v_g \rangle$ with $\Sigma = \{a_1, \ldots a_m\}$, define $f'(SG^{\Sigma}) := \langle V, E, v, v_g \rangle$, where $E(v, v') := |\{\mathcal{E}_{a_i} \mid (v, v') \in \mathcal{E}_{a_i}\}|$.

Showing that Runner has a w.s. in SG^{Σ} iff he has one in $f(SG^{\Sigma})$ is straightforward, and can be done by induction on $n := \sum_{a \in \Sigma} |\mathcal{E}_a|$. Both f and f' are polynomial. ■

B. Sabotage Learning Games

Based on the Sabotage Games framework, we define Sabotage Learning Games as follows.

Definition 2.7: A *Sabotage Learning Game (SLG)* is a Labelled Sabotage Game played by *Learner* (L, taking the role of Runner) and *Teacher* (T, taking the role of Blocker). We distinguish between three different versions, *SLGUE*, *SLGHU* and *SLGHE*, differing in the winning conditions (given in Table II).

TABLE II
Sabotage Learning Games

Game	Winning Condition
SLGUE	Learner wins iff he reaches the goal state. Teacher wins otherwise.
SLGHU	Teacher wins iff Learner reaches the goal state. Learner wins otherwise.
SLGHE	Both players win iff Learner reaches the goal state. Both lose otherwise.

The different winning conditions correspond to different levels of Teacher's helpfulness and Learner's willingness to learn. We can have an *unhelpful* teacher and an *eager* learner (*SLGUE*), but there is also the possibility of a *helpful* teacher and an *unwilling* learner (*SLGHU*). The cooperative case corresponds to the version with a *helpful* teacher and an *eager* learner (*SLGHE*).

We now show how Sabotage Modal Logic can be used for reasoning about Learner's and Teacher's strategic power in the learning games previously defined.

C. Sabotage Modal Logic

Sabotage Modal Logic (SML) has been introduced in [11]. Besides the standard modalities, it also contains "transition-deleting" modalities for reasoning about model change that occurs when a transition is removed. To be more precise, we have formulas of the form $\diamondminus\phi$, expressing that it is possible to delete a pair from the accessibility relation such that ϕ holds in the resulting model at the current state.

Definition 2.8 (Sabotage Modal Language [11]): Let PROP be a countable set of propositional letters and let Σ be a finite set. Formulas of the language of Sabotage Modal Logic are given by

$$\phi ::= p \mid \neg\phi \mid \phi \vee \phi \mid \Diamond_a\phi \mid \diamondminus_a\phi$$

with $p \in$ PROP and $a \in \Sigma$. We write $\Diamond\phi$ for $\bigvee_{a\in\Sigma} \Diamond_a\phi$ and $\diamondminus\phi$ for $\bigvee_{a\in\Sigma} \diamondminus_a\phi$.

Definition 2.9 ([7]): Given a countable set of propositional letters PROP and a finite set $\Sigma = \{a_1, \ldots, a_n\}$, a *Sabotage Model* is a tuple $M = \langle W, (R_{a_i})_{a_i\in\Sigma}, Val\rangle$ where W is a non-empty set of worlds, each $R_{a_i} \subseteq W \times W$ is an accessibility relation and Val : PROP $\to \mathcal{P}(W)$ is a propositional valuation function. The pair (M, w) with $w \in W$ is called a *Pointed Sabotage Model*.

For the semantics of SML, we first define the model that results from removing an edge.

Definition 2.10: Let $M = \langle W, R_{a_1}, \ldots R_{a_n}, Val\rangle$ be a Sabotage Model. The model $M^{a_i}_{(v,v')}$ that results from removing the edge $(v, v') \in R_{a_i}$ is defined as

$$M^{a_i}_{(v,v')} := \langle W, R_{a_1}, \ldots R_{a_{i-1}}, R_{a_i} \setminus \{(v, v')\}, R_{a_{i+1}}, \ldots R_{a_n}, Val\rangle.$$

Definition 2.11: Given a Sabotage Model $M = \langle W, (R_a)_{a\in\Sigma}, Val\rangle$ and a world $w \in W$, atomic propositions, negations, disjunctions and standard modal formulas are interpreted as usual. For the case of "transition-deleting" formulas, we have

$$M, w \models \diamondminus_a\phi \text{ iff } \exists\, v, v' \in W : (v, v') \in R_a \;\&\; M^a_{(v,v')}, w \models \phi,$$

and $\boxminus_a\phi$ is defined to be equivalent to $\neg\diamondminus_a\neg\phi$.

Theorem 2 ([7]): Combined complexity of model checking for SML is PSPACE-complete.
Note that "combined complexity" means that both the formula and the model are taken as input.

D. Sabotage Learning Games in Sabotage Modal Logic

For any given Sabotage Learning Game SG^Σ we can construct a Pointed Sabotage Model $M(SG^\Sigma)$ in a straightforward way.

Definition 2.12: Let $SG^\Sigma = \langle V, \mathcal{E}, v_0, v_g\rangle$ be a Sabotage Game with $\mathcal{E} = (\mathcal{E}_a)_{a\in\Sigma}$. We define the Pointed Sabotage Model $(M(SG^\Sigma), v_0)$ over the set of atomic propositions PROP := $\{goal\}$ with

$$M(SG^\Sigma) := \langle V, \mathcal{E}, Val\rangle,$$

where $Val(goal) := \{v_g\}$.

In the light of this construction, SML becomes useful for reasoning about players' strategic power in *SLG*s. For each winning condition in Table II, we can define a formula of SML that characterizes the existence of a winning strategy, that is, the formula is true in a given Pointed Sabotage Model if and only if the corresponding player has a winning strategy in the game represented by the model.

First we look at the game *SLGUE* (the standard Sabotage Game of [11]), with Learner trying to reach the goal state and Teacher trying to prevent him from doing so. Inductively, we define:

$$\gamma_0^{UE} := goal, \qquad \gamma_{n+1}^{UE} := goal \vee \Diamond\boxminus\gamma_n^{UE}.$$

Our following result is a variation of Theorem 7 of [7], rephrased for Labelled Sabotage Games. We provide a detailed proof to show how our "labelled" definition avoids a technical issue present in the original proof.

Theorem 3: Learner has a winning strategy in the *SLGUE* game $SG^\Sigma = \langle V, \mathcal{E}^0, v_0, v_g\rangle$ iff $M(SG^\Sigma), v_0 \models \gamma_n^{UE}$, for $n := \sum_{a\in\Sigma} |\mathcal{E}^0_a|$ (the number of edges in $(V, \mathcal{E}^0)$).

Proof: The proof is by induction on n.

Base case

($\Rightarrow$) L having a w.s. in SG^Σ implies that $v_0 = v_g$. Thus, $M(SG^\Sigma), v_0 \models goal$ and hence, $M(SG^\Sigma), v_0 \models \gamma_0^{UE}$.

($\Leftarrow$) $M(SG^\Sigma), v_0 \models \gamma_0^{UE}$ means that $M(SG^\Sigma), v_0 \models goal$. Thus $v_0 = v_g$. Hence, L wins SG^Σ immediately.

Inductive case

($\Rightarrow$) Suppose that SG^Σ has $n+1$ edges, and assume that L has a w.s. There are two possibilities. (1) v_0 is the goal state; then $M(SG^\Sigma), v_0 \models goal$ and hence $M(SG^\Sigma), v_0 \models \gamma_{n+1}^{UE}$. (2) v_0 is not the goal state. Since L has a w.s., there is some $v_1 \in V$ such that $(v_0, v_1) \in \mathcal{E}^0_{a_i}$ for some $a_i \in \Sigma$ and no matter which pair $((u, u'), a_j) \in (V \times V) \times \Sigma$ with $(u, u') \in \mathcal{E}^0_{a_j}$ T chooses, L has a w.s. in the resulting game $SG'^\Sigma = \langle V, \mathcal{E}^1, v_1, v_g\rangle$, with $\mathcal{E}^1 = (\mathcal{E}^0_{a_1}, \ldots \mathcal{E}^0_{a_{j-1}}, \mathcal{E}^0_{a_j} \setminus \{u, u'\}, \mathcal{E}^0_{a_{j+1}}, \ldots \mathcal{E}^0_{a_{|\Sigma|}})$. Now, SG'^Σ has n edges and thus by inductive hypothesis, $M(SG'^\Sigma), v_1 \models \gamma_n^{UE}$. This implies $M(SG^\Sigma), v_0 \models \Diamond\boxminus\gamma_n^{UE}$ and thus $M(SG^\Sigma), v_0 \models \gamma_{n+1}^{UE}$. ($\Leftarrow$) $M(SG^\Sigma), v_0 \models goal \vee \Diamond\boxminus\gamma_n^{UE}$ implies that v_0 is the goal state (so L wins immediately) or else there is v_1 accessible from v_0 such that $M(SG^\Sigma), v_1 \models \boxminus\gamma_n^{UE}$, that is, $M(SG^\Sigma)^{a_i}_{(v,v')}, v_1 \models \boxminus\gamma_n^{UE}$ for any $((v, v'), a_i) \in (V \times V) \times \Sigma$. By inductive hypothesis, this gives L a w.s. at v_1 in a game that results from removing any edge from SG^Σ, and hence a w.s. at v_0 in the game SG^Σ. ■

They key observation for the left-to-right direction of this proof is that the model that results from removing an edge from $M(SG^\Sigma)$ is always a model that results from transforming a Labelled Sabotage Game into a model. With the original definition of a Sabotage Game, this is not the case: after removing an edge between v and v' with label k, the resulting model does not need to be the image of a multi-graph because the label of the removed edge does not need to be the biggest of them. Another way to look at it is the following: the multiple edges of the original multi-graph can be seen as implicitly labelled by numbers, and the existence of an edge labelled with k implies the existence of edges labelled with $1, \ldots, k-1$. This property is not preserved when Teacher removes an edge with an arbitrary label from the model $M(SG)$.

Consider now the game *SLGHU*, with Teacher trying to force Learner to reach the goal state. Inductively, define

$$\gamma_0^{HU} := goal, \qquad \gamma_{n+1}^{HU} := goal \vee (\Diamond\top \wedge \Box\diamondminus\gamma_n^{HU}).$$

Now, we can show that this formula corresponds to the existence of a winning strategy for Teacher. Note that in order to win, Teacher has to make sure that Learner does not get stuck before he has reached the goal state. This is why we need the conjunct $\Diamond\top$ in the formula.

Theorem 4: Teacher has a winning strategy in the *SLGUE* game $SG^{\Sigma} = \langle V, \mathcal{E}^0, v_0, v_g\rangle$ iff $M(SG^{\Sigma}), v_0 \models \gamma_n^{HU}$, for $n := \sum_{a\in\Sigma} |\mathcal{E}_a^0|$.

Proof: Similar to the proof of Theorem 3. ■

Finally, consider *SLGHE*, with Teacher and Learner winning iff Learner reaches the goal state. The corresponding formula is defined as follows

$$\gamma_0^{HE} := goal, \qquad \gamma_{n+1}^{HE} := goal \vee \Diamond\diamondminus\gamma_n^{HE}.$$

Theorem 5: Teacher and Learner have a joint winning strategy in the *SLGHE* game $SG^{\Sigma} = \langle V, \mathcal{E}^0, v_0, v_g\rangle$ iff $M(SG^{\Sigma}), v_0 \models \gamma_n^{HE}$, for $n := \sum_{a\in\Sigma} |\mathcal{E}_a^0|$.

Proof: Note that L and T have a joint w.s. iff there is a path from v_0 to v_g. From left to right this is obvious. From right to left, if there is such path, then there is also one without cycles; then, there is a joint w.s. that follows the path and at each step removes the edge that has just been used. The Theorem follows by observing that γ_n^{HE} expresses the existence of such path. ■

The previous results are summarized in Table III.

TABLE III
Winning Conditions for *SLG* in SML

Game	Winning Condition in SML	Winner
SLGUE	$\gamma_0^{UE} := goal,\ \gamma_{n+1}^{UE} := goal \vee \Diamond\boxminus\gamma_n^{UE}$	Learner
SLGHU	$\gamma_0^{HU} := goal,\ \gamma_{n+1}^{HU} := goal \vee (\Diamond\top \wedge (\Box\diamondminus\gamma_n^{HU}))$	Teacher
SLGHE	$\gamma_0^{HE} := goal,\ \gamma_{n+1}^{HE} := goal \vee \Diamond\diamondminus\gamma_n^{HE}$	Both

E. Complexity of Sabotage Learning Games

Intuitively, some versions of the Sabotage Learning Game are simpler than others. With a helpful teacher and an eager learner, the learning process should be easier than with an unhelpful teacher or a unwilling learner. This is indeed reflected in the computational complexity of deciding in a given game whether the winning condition is satisfied.

We have shown that our three winning conditions (Table III) can be expressed in SML, and Theorem 2 (proved in [7]) tells us that model checking of SML is PSPACE-complete. This gives us PSPACE upper bounds for the complexity of the problems of deciding whether each winning condition is satisfied in a given game. For two of the winning conditions (*SLGUE* and *SLGHE*), we can also give tight lower bounds.

For *SLGUE* – the standard Sabotage Game – PSPACE-hardness is shown by reduction from QBF [7].

Theorem 6 ([7]): *SLGUE* is PSPACE-complete.

As mentioned above, for *SLGHU* we obtain a PSPACE upper bound.

Theorem 7: *SLGHU* is in PSPACE.

Proof: Follows from Theorem 2 and Theorem 4. ■

It remains to be shown whether *SLGHU* is also PSPACE-hard. Whereas at first sight, *SLGHU* and *SLGUE* might seem to be duals of each other, the relationship between them is more complex due to the different nature of the players' moves (Learner moves locally by choosing an accessible state, whereas Teacher moves globally, manipulating the structure in which Learner moves).hus, a reduction from *SLGUE* to *SLGHU* is not straightforward. Let us now look at *SLGHE*. This game is of a different nature than the two previous ones. It is cooperative, and a winning strategy is a joint strategy for both players. Such a strategy does not need to take into account all possible moves of the opponent. This suggests that this version should be less complex than *SLGUE* and *SLGHU*.

The following result shows that at least for the comparison of *SLGUE* and *SLGHE*, this is indeed the case: for an eager learner, learning with a helpful teacher is easier than learning with an unhelpful one. This follows from the fact that the winning condition of *SLGHE* is satisfied iff the goal vertex is reachable from the initial vertex (note that Learner moves first). Thus, determining whether Teacher and Learner can win *SLGHE* is equivalent to solving the REACHABILITY (st-CONNECTIVITY) problem, which is known to be nondeterministic logarithmic space (NL)-complete [9].

Theorem 8: *SLGHE* is NL-complete.

Proof: Polynomial equivalence of *SLGHE* and REACHABILITY follows from the argument given in the proof of Theorem 5. ■

Table IV summarizes the complexity results for the different versions of *SLG*.

TABLE IV
Complexity Results for Sabotage Learning Games

Game	Winning Condition	Complexity
SLGUE	Learner wins iff he reaches the goal state, Teacher wins otherwise	PSPACE-complete.
SLGHU	Teacher wins iff Learner reaches the goal state, Learner wins otherwise.	PSPACE
SLGHE	Both players win iff Learner reaches the goal state. Both loose otherwise.	NL-complete

In the case of an eager Learner, the complexity results agree with our intuitions when comparing the cooperative version of the Sabotage Game (*SLGHE*) with the non-cooperative one (*SLGUE*). The easiest way to learn for an eager Learner is when the Teacher is helpful.

III. Relaxing strict alternation

As mentioned above, Learner's moves in the graph are interpreted as changes of information states. Then, when Teacher removes an edge, she actually informs Learner which changes of information state should not be performed. In this perspective, Learner's moves can be seen as internal ones while Teacher's moves can be interpreted externally. Due to this asymmetry, each Learner's move does not in principle need to be followed by a teacher's move.

Definition 3.1: A *Sabotage Learning Game without strict alternation* (for Teacher) is a tuple $SLG^* = \langle V, \mathcal{E}, v_0, v_g \rangle$. Moves of Learner are as in the Sabotage Learning Game and, once he has chosen a vertex v_1, Teacher has a choice between removing an edge, in which case the next game is given as in *SLG*, and doing nothing, in which case the next game is $\langle V, \mathcal{E}, v_1, v_g \rangle$. We again distinguish between three versions, SLG^*UE, SLG^*HU and SLG^*HE, with winning conditions given as before.

Though we provide Teacher with an additional possible move, this does not change her winning abilities. In the rest of this section we show that, for the three variations of a Sabotage Learning Game, a player has a w.s. in SLG^* iff she has a w.s. in *SLG*.

Consider the case of an unhelpful teacher and an eager learner SLG^*UE. Before we go into the details, note that if Learner can win the game, he can do so in a finite number of rounds.

Theorem 9: Consider the *SLG* $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $(V, \mathcal{E})$ a directed labelled multi-graph and v, v_g vertices in it. If Learner has a winning strategy in the corresponding *SLGUE*, then he has a winning strategy in the corresponding SLG^*UE.

Proof: This can be shown by induction on the number of rounds. The idea is that in each round L "pretends" that T has removed some edge and then makes the move given by his strategy for *SLGUE*. ■

If L can win a SLG^*UE, then it is easy to see that he can also win the corresponding *SLGUE* by using his w.s. from SLG^*UE.

Corollary 1: Consider the tuple $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $(V, \mathcal{E})$ a directed labelled multi-graph and v, v_g vertices in it. Learner has a winning strategy in the corresponding SLG^*UE iff he has a winning strategy in the corresponding *SLGUE*.

The case of a helpful teacher and an unwilling learner is more interesting. One might expect that the additional possibility of an empty move gives more power to Teacher since it allows her to skip a move in cases when removing an edge would have made the goal unreachable from the current vertex. However, we can show that this is not the case. First, we state the following lemmas.

Lemma 1: For any SLG^*HU $\langle V, \mathcal{E}, v_0, v_g \rangle$, if there is a path from v_0 to v_g and there is no path from v_0 to a state from where v_g is not reachable, then T has a winning strategy.

Proof: By assumption, all states reachable from v_0 are on paths to v_g. Therefore, even if T will refrain from removing any edge, L will always be on some path to the goal. There are two possibilities: either the path to the goal does not include a loop or it does. If it does not then T can simply wait until L will arrive to the goal. If it does, in order to win T can remove the edges that lead into the loops in such a way that v_g is still reachable from any vertex. L will eventually have to move to v_g. ■

Lemma 2: Consider the SLG^*HU game $\langle V, \mathcal{E}, v_0, v_g \rangle$. If T has a winning strategy and there is some edge $(v, v') \in \mathcal{E}_a$ for some $a \in \Sigma$ such that no path from v_0 to v_g goes via the edge (v, v'), then T also has a winning strategy in $\langle V, \mathcal{E}', v_0, v_g \rangle$, where $\mathcal{E}'$ is the result of removing (v, v') from $\mathcal{E}_a$.

Proof: If v is not reachable from v_0, it is easy to see that the claim holds. Let us consider the case that v is reachable from v_0. Since there is no path to v_g visiting v, T's winning strategy should keep L away from it (otherwise L would win). Hence, T can also win if the edge (v, v') is not there. ■

Theorem 10: If Teacher has a winning strategy in the SLG^*HU $\langle V, \mathcal{E}, v_0, v_g \rangle$, then she also has a winning strategy in which she removes an edge in each round.

Proof: The proof proceeds by induction on the number of edges $n = \sum_{a \in \Sigma} |\mathcal{E}_a|$.

The base case is straightforward. For the inductive case, assume that T has a winning strategy in SLG^*HU $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $\sum_{a \in \Sigma} |\mathcal{E}_a| = n + 1$.

Then if $v_0 = v_g$, we are done. Thus, assume that $v_0 \neq v_g$. Then, since T can win, there is some $v_1 \in V$ such that $(v_0, v_1) \in \mathcal{E}_a$ for some $a \in \Sigma$ and for all such v_1 it holds that:

1) There is a path from v_1 to v_g, and
2) a) T can win $\langle V, \mathcal{E}, v_1, v_g \rangle$, or
   b) there is some $((v, v'), a) \in (V \times V) \times \Sigma$ such that $(v, v') \in \mathcal{E}_a$ and T can win $\langle V, \mathcal{E}', v_1, v_g \rangle$ where $\mathcal{E}'$ is the result from removing (v, v') from $\mathcal{E}_a$.

If 2b holds, since $\sum_{a \in \Sigma} |\mathcal{E}'_a| = n$, we are done — we can use the inductive hypothesis and conclude that T has a w.s. in which she removes an edge in each round (in particular, she chooses $((v, v'), a)$ in the first round). This $((v, v'), a)$ can be chosen in one of the following ways.

If there is some $(v, v') \in V \times V$ such that $(v, v') \in \mathcal{E}_a$ for some $a \in \Sigma$ and this edge is not part of any path from v_1 to v_g then by Lemma 2, T can remove this edge and 2*b* holds and we are done.

If every edge in $(V, \mathcal{E})$ belongs to some path from v_1 to v_g, from 1, there are two cases: either there is only one, or there are more than one paths from v_1 to v_g.

In the first case (only one path) (v_0, v_1) can be chosen since it cannot be part of the *unique* path from v_1 to v_g.

Assume now that there is more than one path from v_1 to v_g. Let $p = (v_1, v_2, \ldots, v_g)$ be the/a shortest path from v_1 to v_g. This path cannot contain any loops. Then, from this path take v_i such that i is the smallest index for which it holds that from v_i there is a path $(v_i, v'_{i+1}, \ldots v_g)$ to v_g that is at least as long as the path following p from v_i (i.e. $(v_i, v_{i+1}, \ldots, v_g)$). Intuitively, when following path p from v_1 to v_g, v_i is the first point at which one can deviate from p in order to take another path to v_g (recall that we consider the case where every vertex in the graph is part of some path from v_1 to v_g). Now it is possible for T to choose $((v_i, v'_{i+1}), a)$ such that $(v_i, v'_{i+1}) \in \mathcal{E}_a$. Let $\mathcal{E}'$ be the resulting set of edges after removing (v_i, v'_{i+1}) from $\mathcal{E}_a$. Then we are in the position $\langle V, \mathcal{E}', v_1, v_g \rangle$. Note that because of the way we chose the edge that has been removed, in the new graph it still holds that from v_0 there is no path to a vertex from which v_g is not reachable (this holds because from v_i the goal v_g is still reachable). Then by Lemma 1, T can win $\langle V, \mathcal{E}', v_1, v_g \rangle$, which then implies 2*b*.

Hence, we conclude that 2*b* has to be the case and thus using the inductive hypothesis, we conclude that T can win the game $\langle V, \mathcal{E}, v_0, v_g \rangle$ also by removing an edge in every round. ■

Corollary 2: Consider the tuple $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $(V, \mathcal{E})$ a directed labelled multi-graph and v_0, v_g vertices in it. Teacher has a winning strategy in the corresponding *SLG*HU*, iff she has a winning strategy in the corresponding *SLGHU*.

Finally, let us move to the case of a helpful teacher and an eager learner.

Theorem 11: Consider the tuple $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $(V, \mathcal{E})$ a directed labelled multi-graph and v_0, v_g vertices in it. If Learner and Teacher have a winning strategy in the corresponding *SLG*HE*, then they have a winning strategy in the corresponding *SLGHE*.

Proof: The proof of Theorem 5 provides the needed strategy. ■

Corollary 3: Consider the tuple $\langle V, \mathcal{E}, v_0, v_g \rangle$ with $(V, \mathcal{E})$ a directed labelled multi-graph and v, v_g vertices in it. Learner and Teacher have joint winning strategy in the corresponding *SLG*HE* iff they have a joint winning strategy in the corresponding *SLGHE*.

In this section we have shown that in Sabotage Learning Games, allowing Teacher to skip moves, does not change the winning abilities of the players. Using these results, both the complexity and definability results from the previous section also apply to the versions of the game in which Teacher can refrain from making a move.

IV. Refined view on teaching: learning algorithms

The perspective on learning that we have adopted is very general. To give a more refined view, let us go back to the *queries and counterexamples* paradigm (see [1]). In that approach, Learner is an algorithm that embodies a winning strategy in the game of learning (the learning procedure succeeds on *all* possible *true* data). Teacher can significantly influence the learning process by giving counterexamples, and the time needed for learning depends on her choices. Therefore, the game of teaching in such a setting can be formalized in extensive form as presented in Figure 1.

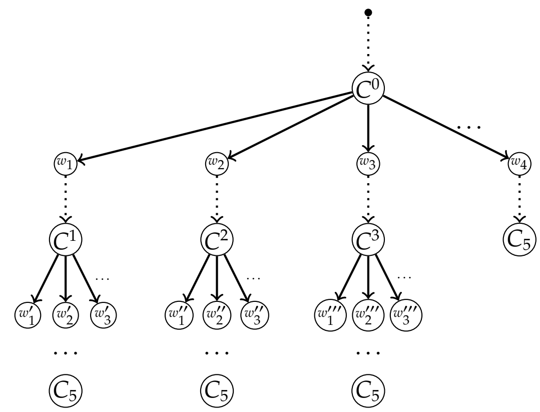

Fig. 1. The tree of the teaching game: dotted lines are Learner's moves, which are determined by his algorithm; solid lines are Teacher's moves; w_i are counterexamples given by Teacher; C_i are conjectures made by Learner; C_5 is the correct hypothesis.

There are many game-theoretical issues that arise when viewing the run of the learning algorithm as a game. We can for example consider the epistemic status of the players, introduce imperfect information and analyze payoff characteristics. Concerning the payoff characteristics and different classes of teachers such as (un)helpful teachers, we can define corresponding preference relations or payoffs: the *helpful* teacher may strictly prefer all shortest paths in the game tree, i.e. the paths in which the learner learns the fastest. The *unhelpful* teacher might strictly prefer all the longest paths in the game tree, i.e. the paths in which the learner learns slowly.

We can also provide a choice for Learner in this game. Firstly, we can allow that at each step the learner can choose from one or more procedures which are part of one algorithm. Secondly, in the beginning Learner can decide with which of the available algorithms he is going to proceed. Moreover, we can consider also another possibility that involves extending the traditional inductive inference paradigm. Usually, learnability of a class is interpreted as the existence of a learner that learns every element from the class independently of the behavior of Teacher — if we introduce the possibility of non-learnability to the game, we can view learning algorithms as winning strategies for an eager learner in the learning game. With the possibility of non-learnability, there are also paths in the game tree in which the learner never makes a correct conjecture. In this framework, a *helpful* teacher would also prefer all (shortest) paths ending in a position in which the learner makes a correct conjecture over all the other paths. An *unhelpful* teacher then prefers all the paths in which the learner does not learn over those in which he does learn.

V. Conclusions and further work

We have provided a game theoretical approach to learning that allows us to analyze different levels of cooperativeness between Learner and Teacher. We have defined *Sabotage Learning Games* with three variations, representing different didactic scenarios. Then, we have shown how *Sabotage Modal Logic* can be used to reason about these games and, in particular, we have identified certain formulas of the language with the existence of a winning strategy. We gave complexity results for the decision problems associated with each version of the game. These problems correspond to model checking for the associated formulas and models. Our complexity results support the intuitive claim that the cooperation of agents facilitates learning. We investigated an extension of the Sabotage Learning Games that relaxes the condition of the strict alternation of moves. Our results presented in Section III show that if we allow Teacher to skip a move, the winning abilities of the players do not change with respect to the original versions of the games. In the case of the helpful teacher and unwilling learner, this is quite surprising since it says that if Teacher can force Learner to learn in the game with nonstrict alternation, then even if she is forced to remove edges in each round she can do so without removing edges that are necessary for Learner to eventually reach the goal state.

From the perspective of Formal Learning Theory, several relevant extensions can be done. We have described the learning process as *changes in information states*, without going further into their epistemic and/or doxastic interpretation. A deeper analysis can give us insights about how the learning process is related to different notions of dynamics of information, such as belief revision or dynamic epistemic logic.

Moreover, it can be argued that in some natural learning scenarios, e.g. language learning, the goal of the learning process, e.g. a correct grammar, is concealed from Learner. In our approach we assume that Learner has at least some markers of the goal state, otherwise we could not introduce the two possible characteristics of Learner. The assumption of the absence of any ability to recognize the goal by Learner, leads to a model in which the Learner moves randomly. Some approaches that take into account randomness in Sabotage Games have already been introduced [8] and hopefully can be extended to deal with the aforementioned issue. What we can now hypothesize is that the complexity of the scenario with a random Learner and a helpful Teacher is bounded by the worst case scenario, in which Learner avoids the goal as long as possible, i.e. the game *SLGHU*.

In the introduction we described the concepts of *finite identification* and *identification in the limit*. Our work on *SLG*s is closer to the first one, as we understand learning as the ability to reach an appropriate information state, without taking into account what will happen after such a state has been reached. In particular, we are not concerned with the stability of the resulting belief. *Identification in the limit* extends *finite identification* by looking beyond reachability in order to describe "ongoing behaviour". Fixed-point logics, like the propositional μ-calculus [10], [5], can provide us with tools to express this notion of learnability. In this case, epistemic and doxastic interpretations of learning would involve notions of stable belief and a kind of operational, non-introspective knowledge as a result of the process.

References

[1] Dana Angluin. Learning regular sets from queries and counterexamples. *Information and Computation*, 75(2):87–106, 1987.

[2] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer Netherlands, 2007.

[3] E.M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[4] Sanjay Jain, Daniel Osherson, James S. Royer, and Arun Sharma. *Systems that Learn*. MIT Press, Chicago, 1999.

[5] Dexter Kozen. Results on the propositional mu-calculus. *Theoretical Computer Science*, 27:333–354, 1983.

[6] Christof Löding and Philipp Rohde. Solving the Sabotage Game is PSPACE-hard. In *Proceedings of the 28th International Symposium on Mathematical Foundations of Computer Science, MFCS '03*, volume 2474 of *LNCS*, pages 531–540. Springer, 2003.

[7] Christof Löding and Philipp Rohde. Solving the Sabotage Game is PSPACE-hard. Technical report, Aachener Informatik Berichte, Rwth Aachen, 2003.

[8] Dominik Klein, Frank G. Radmacher, and Wolfgang Thomas. The Complexity of Reachability in Randomized Sabotage Games. Proceedings of FSEN 2009, LNCS, 2009.

[9] Christos M. Papadimitriou. *Computational complexity*. Addison-Wesley, MA, 1994.

[10] Dana Scott and J. W. de Bakker. A theory of programs. Unpublished manuscript, IBM, Vienna, 1969.

[11] Johan van Benthem. An essay on sabotage and obstruction. In *Mechanizing Mathematical Reasoning, Essays in Honor of Jörg H. Siekmann on the Occasion of His 60th Birthday*, pages 268–276, 2005.

Acknowledgment

The authors would like to thank Jakub Szymanik and Johan van Benthem for their help and their useful comments on previous versions.

Fernando R. Velázquez-Quesada is supported by **Consejo Nacional de Ciencia y Tecnología** (**CONACyT**), México (scholarship # 167693).