Proc. 23rd Int. Workshop on Description Logics (DL2010), CEUR-WS 573, Waterloo, Canada, 2010.

Automata-Based Abduction for Tractable Diagnosis

Thomas M. Hubauer[1,2], Steffen Lamparter[2], and Michael Pirker[2]

[1] Institute Software, Technology, and Systems (STS)
Hamburg University of Technology, Germany
[2] Autonomous Systems Group
Siemens AG, Germany

Abstract. Abductive reasoning has been recognized as a valuable complement to deductive inference for tasks such as diagnosis and integration of incomplete information despite its inherent computational complexity. This paper presents a novel, tractable abduction procedure for the lightweight description logic $\mathcal{EL}$. The proposed approach extends recent research on automata-based axiom pinpointing (which is in some sense dual to our problem) by assuming information from a predefined abducible part of the domain model if necessary, while the remainder of the domain is considered to be fixed. Our research is motivated by the need for efficient diagnostic reasoning for large-scale industrial systems where observations are partially incomplete and often sparse, but nevertheless the largest part of the domain such as physical structures is known. Technically, we introduce a novel pattern-based definition of abducibles and show how to construct a weighted automaton that commonly encodes the definite and abducible part of the domain model. We prove that its behavior provides a compact representation of all possible hypotheses explaining an observation, and is in fact computable in PTime.

1 Introduction

Abductive reasoning is a method for generating hypotheses that explain an observation based on a model of the domain, typically in the presence of incomplete data. Its non-monotonicity and explorative nature make abduction a promising candidate for the interpretation of potentially incomplete information – a task which is much harder to accomplish using established monotonic inference methods such as deduction or the more elaborate axiom pinpointing. The applications of abductive inference are diverse, ranging from text interpretation [1] to plan generation and analysis [2], and interpretation of sensor [3] or multimedia data [4]. Our research on abductive inference is motivated by industrial applications in Ambient Assisted Living and assistive diagnosis for complex technical devices. In these scenarios we found the underlying models being typically large, though not overly complex in their structure. The main consideration is therefore scalability with respect to the size of the domain model; to effectively support humans or to avoid consequential damage to machinery, information processing is subject to soft realtime constraints.

Our proposed solution to this challenge is based upon logic-based abduction which is not the only, but probably the best-studied notion of this type of inference (see [5] for a survey). In logic-based reasoning, model, observations and hypotheses are represented and manipulated using formal logics; description logics were chosen here as a representation language due to their decidability. Since logic-based abduction is known to be at least as hard as deduction [6], the underlying description logic obviously has to be polynomial for subsumption checking. As we found existential quantification to be of greater importance than universal quantification in both scenarios considered so far, we decided to base our approach on the lightweight description logic $\mathcal{EL}$. Choosing a lightweight description logic, however, does not necessarily guarantee tractability of abduction since the so-called support selection task common to all forms of goal-directed reasoning renders hypotheses generation NP-hard even for Horn-theories [7]. It was shown in [8] that this hardness result can only be alleviated if the number of hypotheses is bounded polynomially, allowing (under certain conditions) to generate a single preferred hypothesis in PTIME for $\mathcal{EL}$ and $\mathcal{EL}^+$ knowledge bases [9].

The remainder of this paper is structured as follows: We first recall some basics on description logics and abduction, relating the proposed approach to existing work in this field. Sect. 3 introduces the formalism and justifies its tractability, followed by Sect. 4 where we show how it can be applied to elegantly solve a diagnosis problem. We conclude by summing up the results and giving an outlook on ongoing work.

2 Preliminaries

Description logics are a family of logic-based knowledge representation formalisms designed to ensure decidability of standard reasoning tasks. A concrete description logic is characterized by its admissible concept constructors and axiom types, typically constituting a tradeoff between expressivity and computational complexity. The $\mathcal{EL}$ family of lightweight description logics [10] was tailored specifically to tractability, resulting in a language combining PTIME decidability of standard reasoning tasks with adequate expressivity for modeling e. g. the biomedical ontology SNOMED CT. Table 1 summarizes the constructs available in $\mathcal{EL}$ for defining concepts and axioms based on the sets N_{C} and N_{R} of concept names and role names, respectively. To simplify presentation we will assume for the remainder of this paper that the knowledge base $\mathcal{T}$ is in normal form, containing only general concept inclusion axioms of the form $A_1 \sqcap A_2 \sqsubseteq B$, $A_1 \sqsubseteq \exists r.B$ and $\exists r.A_1 \sqsubseteq B$, where $r \in N_{\mathrm{R}}$, $A_1, A_2, B \in N_{\mathrm{C}} \cup \{\top\}$. For the complete $\mathcal{EL}$ family, normalization of an axiom set is linear in the number of axioms both concerning the time required and the number of new axioms generated [11].

Axiom pinpointing, which provides a basis for the approach presented in this paper, can be seen to extend subsumption checking by determining sets $S \sqsubseteq \mathcal{T}$ of axioms from such that the axioms in each set provide a justification for a given subsumption $C \sqsubseteq D$ (i. e. $S \models C \sqsubseteq D$). While this non-standard

Table 1. $\mathcal{EL}$ syntax & semantics

Syntax	Semantics
$\top$	$\Delta^{\mathcal{I}}$
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
$\exists r.C$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} : (x,y) \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
$C \equiv D$	$C^{\mathcal{I}} = D^{\mathcal{I}}$

inference task provides useful information in case $\mathcal{T} \models C \sqsubseteq D$, it necessarily fails if $\mathcal{T} \not\models C \sqsubseteq D$. In this latter situation, abductive inference offers a solution by determining sets of hypotheses $\mathcal{H}$ compatible with $\mathcal{T}$ that justify the observation if added to the knowledge base (formally, $\mathcal{T} \cup \mathcal{H} \not\models \bot$ and $\mathcal{T} \cup \mathcal{H} \models C \sqsubseteq D$). Due to the restriction of $\mathcal{EL}$ to terminological information we focus our attention on TBox abduction, where both observations and hypotheses are represented by concept inclusion axioms.

In this respect, our work is closely related to the framework of concept abduction [12] which determines, given a knowledge base $\mathcal{T}$ and two concepts C and D, a concept H such that $\mathcal{T} \not\models C \sqcap H \equiv \bot$ and $\mathcal{T} \models C \sqcap H \sqsubseteq D$. This approach as well as the more elaborate notion of structural abduction [13] employ a tableaux-based calculus for finding a single, $\sqsubseteq$-optimal explanation. The authors do not address computational complexity; due to the underlying description logic and the tableau-based approach, we presume that their approach is at least ExpTime-hard. Regarding ABox abduction, [4] presents an approach for $\mathcal{SHIQ}$ knowledge bases extended with non-recursive DL-safe rules. Abduction is implemented as an iterative query answering process that returns a single optimal solution subject to a quality criterion which rewards using asserted information while penalizing assumptions. The approach was successfully implemented in a media interpretation framework, its ExpTime worst-case complexity however is prohibitive in the scenario under consideration. Various aspects of abductive inference have also been studied in the context of logic programming, where resolution is most commonly employed for hypotheses generation. This integrates abductive reasoning tightly with the general setting of logic programming but also poses new questions, for example regarding the interaction of abduction with negation as failure used in most logic programming systems. The interested reader is referred to [14] for a thorough introduction to the field of abductive logic programming. [15] examines the relationships between abductive inference and filtering, a process of model selection similar to conditioning in Bayesian networks. Filtering has successfully been applied in performance-critical applications, proving that it can be implemented efficiently. Under certain conditions abduction can indeed be implemented as a process of filtering, yet in in the general case (and especially for unrestricted propositional and first-order theories) filtering is equivalent only to so-called weak abduction.

In order to obtain a tractable algorithm for abductive reasoning within description logics, we resort to recent work on automata-based axiom pinpointing for $\mathcal{EL}$ [16, 17]. The proposed method is based on encoding the model into a weighted Büchi automaton whose accepting runs (called behavior) represent all derivations of the observation from domain knowledge and abducible information, the latter of which is defined compactly using patterns. A hypothesis formula encoding this set of explanations can be determined in PTime with respect to the size of the knowledge base. The upcoming section presents the details of our approach.

3 Automata-Based Abduction for EL

We start by introducing the abductive framework this paper builds on. It differs from other approaches presented above in that both the observation we want to explain and the abducibles are general concept inclusion axioms, which is actually the only way to express relationships between domain elements in $\mathcal{EL}$ due to the absence of individuals. As mentioned before, we assume that the knowledge base $\mathcal{T}$ is in normal form.

Definition 1 (Axiom pattern; instantiation). *Let $\mathcal{T}$ be an $\mathcal{EL}$ TBox over concept names N_{C} and role names N_{R}, $\mathcal{V}^{\mathrm{C}}$ a set of concept variables V_i^{C}, and $rng : \mathcal{V}^{\mathrm{C}} \rightarrow \mathcal{P}(N_{\mathrm{C}} \cup \{\top\})$ a complete function mapping each concept variable to a set of concept names (possibly including $\top$), called its range. The range extends by subsumption to $rng^*(V_i^{\mathrm{C}}) = \{C \in N_{\mathrm{C}} \cup \{\top\} \mid \exists D \in rng(V_i^{\mathrm{C}}) : \mathcal{T} \models C \sqsubseteq D)\}$ (with $rng(V_i^{\mathrm{C}}) \subseteq rng^*(V_i^{\mathrm{C}})$ since $\mathcal{T} \models C \sqsubseteq C$ trivially holds). An* axiom pattern *is an axiom as defined in Table 1 (not necessarily in normal form), where concept descriptions may contain concept variables from $\mathcal{V}^{\mathrm{C}}$. An* instantiation *of a pattern is an axiom derived from the pattern by replacing each of its concept variables V_i^{C} with an element of $rng^*(V_i^{\mathrm{C}})$.*

Definition 2 (Abduction problem). *Let $\mathcal{T}$ be an $\mathcal{EL}$ TBox over concept names N_{C} and role names N_{R}, $A_0 \sqsubseteq B_0$ a general concept inclusion in normal form such that $A_0, B_0 \in N_{\mathrm{C}}$ (called the* observation*), and Pat a set of axiom patterns over $\mathcal{V}^{\mathrm{C}}$ whose size is bounded polynomially by the number of concept names in N_{C}, and rng a range function. The tuple $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{\mathrm{C}}, rng)$ is called an* abduction problem.

Concept patterns and range function allow for a very fine-grained definition of the parts of the domain which may be assumed. This proves valuable in large-scale applications where typically most of the domain is considered to be fixed (and assumptions most presumably contradict reality), while only certain types of axioms are likely to represent missing information. As an example, compositional (`partOf`) hierarchies of technical systems are completely known to the constructor, whereas the set of observations about such a system is much more likely to be incomplete. Furthermore, explanations are typically required to be non-trivial [5], in particular a piece of information must not be explained by itself. This can be achieved easily here by selecting appropriate axiom patterns and concept variable ranges. As a side-effect, restricting the set of abducibles cuts the search space and the number of hypotheses generated and may therefore increase efficiency. Note that the limitation of the size of *Pat* in Definition 2 is required to ensure a polynomial worst-case complexity of the algorithm, yet it never posed a severe limitation for domain experts in practice.

Definition 3 (Abducible). *Given* $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{\mathrm{C}}, rng)$*, the* set of abducibles $Abd_{\mathcal{AP}}$ *contains all axioms generated by normalizing the elements of* Pat *and instantiating them with concept names from* rng*, omitting axioms already contained in* $\mathcal{T}$ *. Let* $\mathrm{N}_{\mathrm{C}'}$ *denote the set of concept names* N_{C} *extended with the new concept names introduced during normalization.*

Definition 4 (Labeling function). *Let* $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{\mathrm{C}}, rng)$ *be an abduction problem. Assume that each axiom* ax *in* $\mathcal{T}$ *and each abducible* abd *in* $Abd_{\mathcal{AP}}$ *is labeled with a unique propositional variable* l_{ax} *and* l_{abd}*, respectively, such that the sets of axiom labels and abducible labels are disjoint. The* labeling function lab *then assigns a label to each general concept inclusion* gci *as follows: If* gci *is an axiom (abducible), then* $lab(gci)$ *is the predefined propositional variable* l_{ax} (l_{abd})*. Otherwise, if* gci *is a tautology of the form* $A \sqcap A \sqsubseteq A$ *or* $A \sqcap A \sqsubseteq \top$*, we set* $lab(gci) = \top$*; in all other cases* $lab(gci) = \bot$*. We finally denote by* $lab(\mathcal{AP})$ *the set of all labels occurring in the abduction problem.*

To simplify notation we identify a propositional valuation $\mathcal{V}$ with the set of variables it assigns to be true, and let $\mathcal{A}_{|\mathcal{V}} = \{ax \in \mathcal{A} \mid lab(ax) \in \mathcal{V}\}$ denote the restriction of an axiom set $\mathcal{A}$ to the axioms made true by $\mathcal{V}$. We extend this definition to axiom problems by letting $\mathcal{AP}_{|\mathcal{V}} = (\mathcal{T} \cup Abd_{\mathcal{AP}})_{|\mathcal{V}}$.

Definition 5 (Hypotheses formula). *A* hypotheses formula *for an abduction problem* $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{\mathrm{C}}, rng)$ *is a monotone Boolean formula* $\eta_{\mathcal{AP}}$ *over* $lab(\mathcal{AP})$ *such that for all valuations* $\mathcal{V} \subseteq lab(\mathcal{AP})$ *it holds that* $\mathcal{V} \models \eta_{\mathcal{AP}}$ *iff* $\mathcal{AP}_{|\mathcal{V}} \models A_0 \sqsubseteq B_0$.

Abductive inference on the original knowledge base $\mathcal{T}$ can now be expressed as a pinpointing problem in the extended problem space $\mathcal{T} \cup Abd_{\mathcal{AP}}$. To this end, we define an abductive automaton employing the approach proposed in [17].

Definition 6 (Abductive automaton; behavior). *An* abductive automaton *for an abduction problem* $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{\mathrm{C}}, rng)$ *is a weighted Büchi automaton* $\mathcal{A}_{\mathcal{AP}} = \{Q, wt, in, F\}$ *over binary trees with*

- $Q = \{(A, B), (A, r, B) \mid A, B \in N_{\mathrm{C}'} \cup \{\top\}, r \in N_{\mathrm{R}}\}$,
- $\forall A, B, B_1, B_2 \in N_{\mathrm{C}'} \cup \{\top\}, \forall r \in N_{\mathrm{R}}$
  - $wt((A, B), (A, B_1), (A, B_2)) = lab(B_1 \sqcap B_2 \sqsubseteq B)$,
  - $wt((A, r, B), (A, B_1), (A, A)) = lab(B_1 \sqsubseteq \exists r.B)$,
  - $wt((A, B), (A, r, B_1), (B_1, B_2)) = lab(\exists r.B_2 \sqsubseteq B)$,
  - $wt(q_1, q_2, q_3) = \bot$ *for all other* $q_1, q_2, q_3 \in Q$,
- $in(q) = \top$ *iff* $q = (A_0, B_0)$*, otherwise* $in(q) = \bot$ *, and*

– $F = \{(A,A) \mid A \in N_{C'} \cup \{\top\}\}$,

where Q denotes the set of states, $F \subseteq Q$ the set of terminal states, in the initial distribution, and wt the transition weights of $\mathcal{A}_{\mathcal{AP}}$.

We extend the definition of wt to a complete run $\overrightarrow{r} = q_1 \cdots q_n$ as $wt(\overrightarrow{r}) = wt(q_1) \wedge \cdots \wedge wt(q_n)$, and let $succ(q)$ be the set of all successful runs of $\mathcal{A}_{\mathcal{AP}}$ starting in q. The behavior *of $\mathcal{A}_{\mathcal{AP}}$ is defined by $\bigwedge_{q \in Q}(in(q) \wedge \bigvee_{\overrightarrow{r} \in succ(q)} wt(\overrightarrow{r}))$.*

As there is exactly one state q having $in(q) \neq \bot$, namely (A_0, B_0), the behavior of $\mathcal{A}_{\mathcal{AP}}$ is the disjunction of the weights of all its successful runs starting in (A_0, B_0). Due to the specification of the transition weights, each run corresponds to a derivation of $A_0 \sqsubseteq B_0$. Intuitively, wt attributes triples (q_1, q_2, q_3) of states with provenance information regarding the derivation of q_1 from q_2 and q_3: Trivial derivation steps (such as $q_1 = (A, \top)$ or $q_1 = q_2 = q_3$) are labeled with the symbol $\top$ due to Definition 4; the weight of a non-trivial step is the label of an axiom / abducible such that q_1 can be deduced from q_2 and q_3 using this axiom / abducible (or $\bot$ if none exists). As an example, the definition $wt((A,B),(A,B_1),(A,B_2)) = lab(B_1 \sqcap B_2 \sqsubseteq B)$ expresses that, given $A \sqsubseteq B_1$ and $A \sqsubseteq B_2$, we can derive $A \sqsubseteq B$ if we know $B_1 \sqcap B_2 \sqsubseteq B$.

Theorem 1. *Given an abduction problem $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{C}, rng)$, the behavior of the abductive automaton $\mathcal{A}_{\mathcal{AP}}$ is a hypotheses formula for the observation $A_0 \sqsubseteq B_0$.*

This result carries over from [17]. In fact, if we set $Pat = \emptyset$, the abductive automaton and hypotheses formula defined before coincide with the notions of pinpointing automaton and pinpointing formula due to the empty space of abducibles. If $Abd_{\mathcal{AP}}$ is nonempty, the automaton $\mathcal{A}_{\mathcal{AP}}$ can be interpreted as a pinpointing automaton for TBox $\mathcal{T}' = \mathcal{T} \cup Abd_{\mathcal{AP}}$ as noted before. Due to space limitations the reader is referred to [16, 18] for details on how to compute the behavior of such an automaton effectively. In the setting introduced above this can even be done efficiently, as the following theorem claims.

Theorem 2. *Given an abduction problem $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{C}, rng)$, computing the hypotheses formula $\eta_{\mathcal{AP}}$ takes polynomial time in the size of $\mathcal{T}$.*

Proof. Given $\mathcal{AP} = (\mathcal{T}, A_0 \sqsubseteq B_0, Pat, \mathcal{V}^{C}, rng)$, we denote by N_C and N_R the sets of concept and role names in $\mathcal{T}$, and by $N_{C'}$ the extended set of concept names including the new names generated during normalization of the axiom patterns in Pat. As motivated before we can regard $\mathcal{A}_{\mathcal{AP}}$ as a pinpointing automaton for the extended problem space $\mathcal{T} \cup Abd_{\mathcal{AP}}$, whose behavior can be computed with an algorithm that is polynomial in the number of states of the automaton as shown in [17]. Following the construction given in Definition 6, $\mathcal{A}_{\mathcal{AP}}$ has $\binom{|N_{C'}|+1}{2}$ states of type (A,B) plus $\binom{|N_{C'}|+1}{2} * \binom{|N_R|}{1}$ states of type (A,r,B), which is polynomial in $N_{C'}$ and N_R. To complete the proof, we therefore have to show that N_C does not grow too fast during normalization of Pat, more concretely we require that $|N_{C'}| = poly(|N_C|)$ (normalization of $\mathcal{EL}$ axiom patterns introduces no new role names at all). To this end, observe that the number of new concept names introduced by normalizing a set of axioms is linear in the number of axioms in the set [11]. Therefore, $|N_{C'}| \leq |N_C| + c * |Pat|$ for some constant c which can be chosen independently of N_C. Since the number of axiom patters is bounded polynomially by the size of N_C in Definition 2, this proves the polynomial bound on the size of $N_{C'}$ and therefore also on the size of $\mathcal{A}_{\mathcal{AP}}$. Also note that the size of the abductive automaton and thus the complexity of the proposed approach are independent of the number of concept variables used since variables cannot induce new states in $\mathcal{A}_{\mathcal{AP}}$. □

In assistive diagnosis, it is often convenient to be able to compare explanations of different, competing diagnoses (called a differential diagnosis in medicine). The abduction method proposed here naturally meets this demand, as the only part of the automaton that depends on the observation $A_0 \sqsubseteq B_0$ is the initial distribution *in*. To derive the hypotheses formula for a different observation $A_1 \sqsubseteq B_1$, the complete automaton $\mathcal{A}_{\mathcal{AP}}$ can be re-used without any modification to determine the successful runs starting in (A_1, B_1).

To conclude this section, we give an intuition of how the hypotheses formula generated by $\mathcal{A}_{\mathcal{AP}}$ can be interpreted. $\eta_{\mathcal{AP}}$ compactly encodes all possible derivations of $A_0 \sqsubseteq B_0$ w. r. t. $\mathcal{T}$ and $Abd_{\mathcal{AP}}$. An explicit representation of the set of hypotheses can be derived in a straightforward manner by transforming it $\eta_{\mathcal{AP}}$ into disjunctive normal form, each clause representing a single hypothesis. This approach is obviously not optimal since it may lead to an exponential blowup [16], a real-world system should therefore directly present, interpret and manipulate the compact representation $\eta_{\mathcal{AP}}$ whenever possible. Note that the hypotheses formula carries information on both necessary assumptions and axioms required to justify $A_0 \sqsubseteq B_0$. The proposed approach can therefore be seen to integrate and complement axiom pinpointing by allowing to infer reasons for unwanted entailments to hold as well as for expected subsumptions not to hold. This provides additional capabilities which may be useful among others for ontology debugging and refactoring. If one is only interested in determining necessary assumptions but not in their interactions with the axioms from the domain model, the approach can easily be adapted by adding only labels for abducibles to the hypotheses formula, leading to a much more compact $\eta_{\mathcal{AP}}$.

4 Industrial Scenario

This section illustrates the proposed approach by applying it to a use case in industrial diagnosis. Real-world models in this scenario typically consist of thousands of components and subcomponents, for most of which one can observe certain symptoms indicating possible failure states of the system. More often than not, the causal structure of the domain is at least partially unknown, models for diagnosis therefore have to be built on experience, relating sets of symptoms to diagnoses determined by a technician checking the system. We focus on assistive diagnosis, where sensor data and observations made by maintenance personnel are used to interactively diagnose the system by actively requesting missing observations.

For our necessarily simplified scenario, we consider a CNC lathe with two components surveyed by sensors: the axle motor, and the oil pump of the motor cooling system. Sensors mounted at the axle motor can recognize vibrations and increased temperature, the monotored parameters for the oil pump include the current voltage. We assume that the measurements of these sensors are enough to recognize two different failure states, namely an untrue axle (characterized by vibrations and high axle motor temperature) and a power failure in the axle cooling system (defined by an overheating motor and low oil pump voltage). A system having an axle cooling failure, for example, can be represented by the following $\mathcal{EL}$ axiom:

$$\begin{gathered}\exists hasComp.(AxleMotor \sqcap \exists hasSymp.HiTemp) \ \sqcap \\ \exists hasComp.(OilPump \sqcap \exists hasSymp.LowVoltage) \sqsubseteq \\ \exists hasDiag.AxleCoolFail\end{gathered}$$

Normalizing the axiom results in the normal form axioms

$$Has^{\text{Comp}}_{\text{HotAM}} \sqcap Has^{\text{Comp}}_{\text{DeadOP}} \sqsubseteq System_{\text{ACF}} \tag{1}$$

$$\exists hasComp.HotAM \sqsubseteq Has^{\text{Comp}}_{\text{HotAM}} \tag{2}$$

$$AxleMotor \sqcap Has^{\text{Symp}}_{\text{HiTemp}} \sqsubseteq HotAM \tag{3}$$

$$\exists hasSymp.HiTemp \sqsubseteq Has^{\text{Symp}}_{\text{HiTemp}} \tag{4}$$

$$\exists hasComp.DeadOP \sqsubseteq Has^{\text{Comp}}_{\text{DeadOP}} \tag{5}$$

$$OilPump \sqcap Has^{\text{Symp}}_{\text{LowVoltage}} \sqsubseteq DeadOP \tag{6}$$

$$\exists hasSymp.LowVoltage \sqsubseteq Has^{\text{Symp}}_{\text{LowVoltage}} \tag{7}$$

where $System_{\text{ACF}}$ is a new concept name defined by

$$System_{\text{ACF}} \equiv \exists hasDiag.AxleCoolFail$$

An untrue axle, the second diagnosis considered in this example, can defined and normalized analogously, leading to the following additional $\mathcal{EL}$ axioms in normal form:

$$Has^{\text{Comp}}_{\text{HotAM}} \sqcap Has^{\text{Comp}}_{\text{VibratAM}} \sqsubseteq System_{\text{UA}} \tag{8}$$

$$\exists hasComp.VibratAM \sqsubseteq Has^{\text{Comp}}_{\text{VibratAM}} \tag{9}$$

$$AxleMotor \sqcap Has^{\text{Symp}}_{\text{Vibrations}} \sqsubseteq VibratAM \tag{10}$$

$$\exists hasSymp.Vibrations \sqsubseteq Has^{\text{Symp}}_{\text{Vibrations}} \tag{11}$$

Having specified general (terminological) knowledge about the dependencies of certain symptoms and diagnoses, we now formalize the concrete system under consideration denoted by $System_{\text{Obs}}$, for which we have measured both an increased axle temperature and low voltage in the system for pumping the oil used to cool the axle motor:

$$System_{\text{Obs}} \sqsubseteq \exists hasComp.AxleMotor_{\text{Obs}} \tag{12}$$
$$System_{\text{Obs}} \sqsubseteq \exists hasComp.OilPump_{\text{Obs}} \tag{13}$$
$$AxleMotor_{\text{Obs}} \sqsubseteq AxleMotor \tag{14}$$
$$AxleMotor_{\text{Obs}} \sqsubseteq \exists hasSymp.HiTemp \tag{15}$$
$$OilPump_{\text{Obs}} \sqsubseteq OilPump \tag{16}$$
$$OilPump_{\text{Obs}} \sqsubseteq \exists hasSymp.LowVoltage \tag{17}$$

Assume that the maintenance personnel wants to compare explanations for the diagnoses untrue axle and axle cooling failure to decide on further diagnostic or corrective steps. We then have two target states $q_0 = (System_{\text{Obs}}, System_{\text{ACF}})$ and $q_1 = (System_{\text{Obs}}, System_{\text{UA}})$ for which the hypotheses formula may be determined independently using the same abductive automaton $\mathcal{A}_{\mathcal{AP}}$ (with a modified definition of *in*). Regarding the space of abducibles, we regard the physical structure of the system as fixed and only allow for symptoms to be assumed. This can be done by defining $Pat = \{V_{\text{Comp}} \sqsubseteq \exists hasSymp.V_{\text{Symp}}\}$, where $rng(V_{\text{Comp}}) = Component$ and $rng(V_{\text{Symp}}) = Symptom$. The number of concept inclusions in $Abd_{\mathcal{AP}}$ is too large for an extensive listing even in this simple case, so we limit our presentation to one axiom in $Abd_{\mathcal{AP}}$ required to form a hypothesis for the diagnosis of an untrue axle:

$$AxleMotor_{\text{Obs}} \sqsubseteq \exists hasSymp.Vibrations \tag{18}$$

For the same reason, we cannot present the complete automaton $\mathcal{A}_{\mathcal{AP}}$ here. Figure 1 depicts an excerpt containing one successful run for each diagnosis under consideration. These two runs actually correspond to the most natural hypotheses in terms of requiring the least number of assumptions to be made. Regular/ input/ terminal states are drawn as light/ medium/ dark rectangles, and light/ medium/ dark circles represent axiom labels, the tautology label $\top$, or the labels of abducibles, respectively. To keep the representation compact, we merge identical subtrees.

The weights of the runs from the two input nodes $(System_{\text{Obs}}, System_{\text{ACF}})$ and $(System_{\text{Obs}}, System_{\text{UA}})$ to the terminal (leaf) nodes represent two partial hypotheses formulas for the diagnoses *AxleCoolingFailure* and *UntrueAxle*:

$$\begin{aligned}
\eta^{\text{part}}_{\text{ACF}} &= 1 \wedge (5 \wedge 13 \wedge (6 \wedge (16 \wedge \top) \wedge (7 \wedge 17))) \\
&\quad \wedge (2 \wedge 12 \wedge (3 \wedge (14 \wedge \top) \wedge (4 \wedge 15))) \\
\eta^{\text{part}}_{\text{UA}} &= 8 \wedge (2 \wedge 12 \wedge (3 \wedge (14 \wedge \top) \wedge (4 \wedge 15))) \\
&\quad \wedge (9 \wedge 12 \wedge (10 \wedge (14 \wedge \top) \wedge (11 \wedge \mathbf{18})))
\end{aligned}$$

Comparing the two hypotheses, it shows that neither of them is clearly better than the other: On the one hand, an axle cooling failure is justified by the observations alone (requiring no assumptions to be made), yet it postulates faults

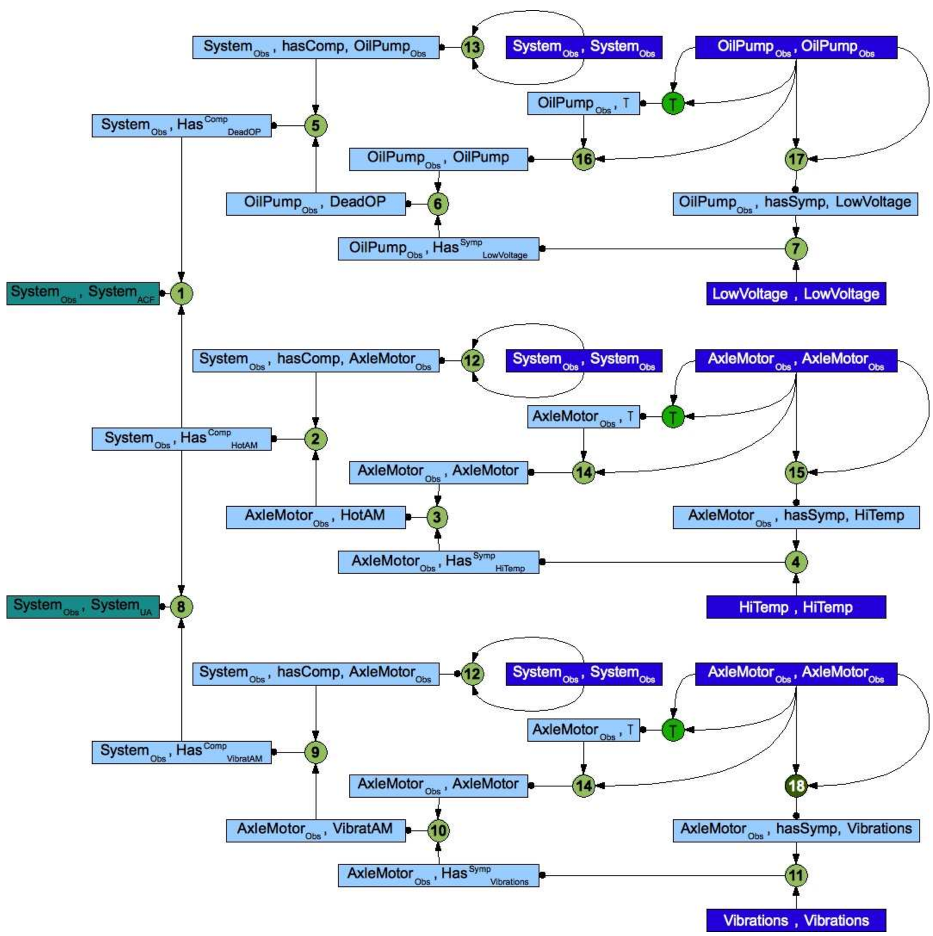

Fig. 1. Automaton for the diagnosis example (compacted excerpt)

in two distinct components. On the other hand, an untrue axle can be diagnosed locally for one component, it however requires the assumption of general concept inclusion axiom 18.

5 Conclusions and Future Work

We have presented a PTime procedure for TBox abduction in the lightweight description logic $\mathcal{EL}$ based on a novel reduction to axiom pinpointing, and demonstrated its applicability in an industrial diagnosis scenario. Given a knowledge base and a concept inclusion representing the observation to be explained, the procedure determines a hypotheses formula that compactly encodes all explanations with respect to a pattern-based representation of the abducible part of the domain model; the remainder of the model is considered to be fixed in accordance with our scenario. The proposed reduction of abductive inference to axiom pinpointing exploits the duality of the two tasks: whereas the latter addresses the problem of explaining why a certain unwanted subsumption is entailed by the ontology, our method determines the reason for an expected subsumption not to hold, expressed in terms of additions to the domain model necessary to actually make it hold.

We are currently working on extending the approach presented in this paper in several ways: Since role inclusion axioms and nominals are frequently used in diagnostic models, it is favorable to extend the logical expressivity as much as possible without sacrificing tractability. Additionally, including quantitative information into the model allows for weighting hypotheses and can eventually be used as a criterion for guiding hypothesis generation. Finally, extending minimality criteria for single hypotheses to sets of hypotheses compactly represented by a hypothesis formula will allow us to efficiently infer common effects (as proposed e. g. in [15]).

References

1. Hobbs, J.R., Stickel, M.E., Appelt, D.E., Martin, P.A.: Interpretation as abduction. Artificial Intelligence **63**(1-2) (1993) 69–142
2. Appelt, D.E., Pollack, M.E.: Weighted abduction for plan ascription. In: User Modeling and User-Adapted Interaction. Volume 2., Springer (1992) 1–25
3. Shanahan, M.: Perception as abduction: Turning sensor data into meaningful representation. Cognitive Science **29**(1) (2005) 103–134
4. Peraldi, I.S.E., Kaya, A., Melzer, S., Möller, R., Wessel, M.: Towards a media interpretation framework for the semantic web. In: Proceedings of the 2007 IEEE / WIC / ACM International Conference on Web Intelligence (WI 2007), Silicon Valley, California, USA, IEEE Computer Society (November 2-5 2007)
5. Paul, G.: Approaches to abductive reasoning: An overview. Artificial Intelligence Review **7**(2) (1993) 109–152
6. Eiter, T., Gottlob, G.: The complexity of logic-based abduction. Journal of the ACM **42**(1) (1995) 3–42

7. Selman, B., Levesque, H.J.: Abductive and default reasoning: A computational core. In: Proceedings of the 8th National Conference on Artificial Intelligence (AAAI 1990), Boston, Massachusetts, USA, AAAI Press / The MIT Press (July 29 - August 3 1990) 343–348
8. Eiter, T., Makino, K.: On computing all abductive explanations. In: Proceedings of the 21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference (AAAI 2006 + IAAI 2006), Edmonton, Alberta, Canada, AAAI Press (July 28 - August 1 2002) 62–67
9. Bienvenu, M.: Complexity of abduction in the $\mathcal{EL}$ family of lightweight description logics. In Brewka, G., Lang, J., eds.: Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR 2008), Sydney, Australia, AAAI Press (September 16-19 2008) 220–230
10. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. In Kaelbling, L.P., Saffiotti, A., eds.: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005), Edinburgh, Scotland, UK, Professional Book Center (July 30 - August 5 2005) 364–369
11. Baader, F., Brandt, S., Lutz, C.: Pushing the $\mathcal{EL}$ envelope. LTCS-Report LTCS-05-01, Institute for Theoretical Computer Science, TU Dresden (2005)
12. Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F.M., Mongiello, M.: Concept abduction and contraction in description logics. In Calvanese, D., De Giacomo, G., Franconi, E., eds.: Proceedings of the 16th International Workshop on Description Logics (DL 2003). Volume 81 of CEUR Workshop Proceedings., Rome, Italy, CEUR-WS.org (September 5-7 2003)
13. Di Noia, T., Di Sciascio, E., Donini, F.M.: A tableaux-based calculus for abduction in expressive description logics: Preliminary results. In Grau, B.C., Horrocks, I., Motik, B., Sattler, U., eds.: Proceedings of the DL Home 22nd International Workshop on Description Logics (DL 2009). Volume 477 of CEUR Workshop Proceedings., Oxford, UK, CEUR-WS.org (July 27-30 2009)
14. Kakas, A.C., Kowalski, R.A., Toni, F.: Abductive logic programming. Journal of Logic and Computation **2**(6) (December 1992) 719–770
15. Baral, C.: Abductive reasoning through filtering. Artificial Intelligence **120**(1) (2000) 1–28
16. Baader, F., Peñaloza, R.: Automata-based axiom pinpointing. In Armando, A., Baumgartner, P., Dowek, G., eds.: Proceedings of the 4th International Joint Conference on Automated Reasoning (IJCAR 2008). Volume 5195 of Lecture Notes in Artificial Intelligence., Sydney, Australia, Springer (August 12-15 2008) 226–241
17. Peñaloza, R.: Using tableaux and automata for pinpointing in $\mathcal{EL}$. In Goranko, V., ed.: Proceedings of the TABLEAUX 2009 Wokshop on Tableaux versus Automata as Logical Decision Methods (AutoTab 2009). (2009)
18. Baader, F., Peñaloza, R.: Automata-based axiom pinpointing. Journal of Automated Reasoning **Special Issue: IJCAR 2008** (2009)