# A Theory-based Logic of Inductive Generalisation

Panu A. Kalliokoski

31.5.2007

**Abstract**

This paper presents a logic of generalisation **TLIG**: an algorithm for constructively deriving correct generalisations from a set of observations, i.e. an induction procedure. **TLIG** is heavily based on the concept of theory, which is a set of propositions, and the operations between theories. Applications of **TLIG** include artificial intelligence, especially machine learning, insight into the operation of inductive processes, and ways to apply logic to real-world phenomena.

Keywords: induction, generalisation, logic, induction algorithm, theory, context, inductive logic programming

## Contents

# 1 The reasoning process of generalisation

Consider a very common reasoning process that is based on an interesting situation: two individuals, say Tom and Sally, never show up at the same time. When you see both frequently but never at the same time, you start to suspect a dependency between the two states of affair, which could be expressed as $(\forall\alpha)\neg\mathrm{Present}(\alpha, \mathrm{Tom}) \vee \neg\mathrm{Present}(\alpha, \mathrm{Sally})$, that is, for any event $\alpha$, either Tom or Sally is absent from $\alpha$.

The theory-based logic of inductive generalisation (**TLIG**) is an attempt to formalise this kind of reasoning. It is an algorithm to construct, from a set of observations, all the rules that hold across the individuals involved in the observations. **TLIG** is based on three central ideas:

1. the correct generalisations are those that are unrefuted by our observational data;
2. generalisations can be formed from particular instances by substituting arbitrary terms in them with variables; and
3. the differences between observational situations can be used to systematically construct *rule candidates*, disjunctive claims that have the potential to hold over the whole range of our observations when generalised.

I will first define what I mean by induction in this paper, followed by a quick survey of earlier approaches to logic of induction and their relation to **TLIG**. Then, I define the language and theory operations of **TLIG**, and provide some examples of their results.

This article uses many words in a manner that either differs from common usage or is more precise than the common use of the word. Listed below are the definitions of such terms.

**event** An atomic unit of phenomenal world. Epistemologically, an event is the source of new information.

**observation** Experiencing an event.

**conceptualisation** A proposition or set of propositions that describe an event.

**situation** A set of events.

**condition** Any attribute of a situation.

**theory** A set of propositions.

**observational theory** A theory that describes some situation.

**generalised theory** A theory that is produced from another by tentative generalisation.

**context** A set of background assumptions that can be partially unspecified. The propositions of a theory are all relative to a common context, the context of the theory.

**merging of observational theories** Producing a new theory that describes a situation that is a union of the situations of the argument theories.

**particular proposition** A proposition that has no universally quantified variables in it (a ground clause). A proposition $p$ is *more particular* than another proposition $q$ if $q$ has variables in every place where $p$ has.

**generalisation** A proposition formed from a more particular one by universally quantifying some of its terms.

**rule candidate** A disjunction of propositions formed by merging two observational theories.

**rule** A generalisation of a rule candidate.

# 2 Formal systems of induction

Flach [Fla95, 16–17] explicates the two-step process of inductive reasoning that dates back to C. S. Peirce. The steps of the process are (1) construction of new hypotheses from observations, called *abduction* by Peirce; and (2) Evaluation of these new hypotheses for acceptability, plausibility etc., called *induction* by Peirce. This distinction is also present behind Popper's *Conjectures and Refutations*: conjectures are constructed scientific theories, suggested but not certain; in the process of evaluating them, some are refuted while others are not.

However, even though both steps — construction and evaluation of hypotheses — have been acknowledged in the philosophy of induction from the very beginning, most of the work in the area has been put into the formalisation of the latter. The reason is easy to see: on one hand, classical logic is too rigid to deal with hypotheses, so there is an acute problem with evaluation of hypotheses; on the other hand, the construction of new hypotheses presents an essentially open-ended problem as it does not *need* to obey any specific rules.

However, for both practical and theoretical reasons, it is important how hypotheses are constructed. Firstly, the construction of hypotheses occurs as an everyday phenomenon and its formalisation provides important insight into the operation of human mind and means to build effective systems that are able to construct new hypotheses. Secondly, some hypotheses are clearly more sensible than others while being equally plausible in the light of some observations. For instance, an agent that observes many sunny days but no cloudy ones might conclude that all days are sunny, or that all trees are sunny. Both are possible hypotheses from the same observations, but the former feels more relevant than the latter. This suggests that it is possible to give a normative description of how relevant hypotheses are obtained from examples.

In this section, I first talk about the meaning of "induction". Then I take a look at different frameworks to tentatively propose, evaluate, and refute propositions; and in the last part, I study different approaches to constucting hypotheses.

## 2.1 What is induction?

The terms *induction* and *abduction* warrant a terminological note. Both are nowadays used to refer to many phenomena, but the most widely accepted definitions are probably these: induction is reasoning that produces a general rule out of many particular instances, while abduction is "inverse deduction", reasoning from conclusions to premisses. But general propositions usually logically entail their particular instances, and premisses logically entail their conclusions. Thus, for many cases, the definitions actually coincide or at least overlap: both are types of reasoning where we take a proposition $p$ and produce from that other propositions that logically entail $p$.

However, there are two alternatives as to what it means to produce a "general rule" out of instances, as explained by [Fla95, 16–26]. Both kinds of induction produce rules that are not refuted by the observations. However, the positive qualifications differ: in *explanatory induction* the rules must imply (or "explain", hence the name) the observations, i.e. abduction; in *confirmatory induction* the rules should just be unrefuted. So every theory that explains the observations is confirmed by them, but not vice versa. The logic of confirmation was formally studied by Hempel [Hem43].

As Flach notes [Fla95, 26], the difference between the two kinds of induction is essentially in the direction of the transitivity of the relation. If a set of observations $E$ is explained by a theory $T$, it is also explained by a stronger theory $T' \to T$, whereas if $E$ confirms a theory $T$, it also confirms a weaker theory $T'$ for which $T \to T'$.

However, induction is just not a search for any theory that explains or is confirmed by our observations. The union of all our observations both explains and is confirmed by itself. Any consistent theory that implies it is "better" in the sense that it is more general. So, if we are allowed to include facts about the individuals in the observations in our theory, induction means producing a maximal theory that is consistent with our observations. Such a theory, of course, also explains our observations, since it is maximal. This is the kind of result that **TLIG** constructs.

If propositions about individuals are not allowed in our theory, explanatory induction is not always possible. It is possible in the case that the generalisations can be explained without referring to the factual treats of the individuals involved; for instance, abstract data types such as numbers and lists possess no factual treats that cannot be known from the expressions that refer to them. But real-world individuals, such as the earth, have many properties that just happen to be true or false for that particular individual. However, even in such cases, it is possible to look for a maximal theory that is confirmed by the observations.

An illustrative example is the classic "all ravens are black": $(\forall\alpha)R\alpha \to B\alpha$. It is an inductive rule, but does not *explain* the blackness of an individual $r$ unless we additionally know that $r$ is a raven, and it does not explain $r$ not being a raven, unless we know that $r$ is not black. However, it is certainly a hypothesis that is confirmed by numerous individuals.

## 2.2 Frameworks for evaluating hypotheses

Classical logic deals with *truth*: deriving irrefutably correct consequences from irrefutably correct premises. However, the requirement of irrefutably correct premises is usually too strong for information about the real world. For instance, there is no way to obtain an irrefutable rule from experience: there is always a possibility that future experiences will refute our rule.

Consequently, any system that wants to allow for refutable truths must employ a framework that formalises the process of retracting propositions when they are refuted and keeps track of which propositions should be refuted and when. This subsection provides a survey of these frameworks.

This article shows a way to systematically use a fragment of first-order predicate logic (**FOL**) to keep track of situations that have existed and form correct rules (*generalisations*) from this information. The information is used in two ways: firstly, to come up with relevant rule candidates in the first place; secondly, to drop out those rules that have counterexamples.

To deal with these tentatively true generalisations, a concept of *theory* is used. There are two kinds of theories in this framework: those that collect information about the world (observational theories), and those that are generalised from this information. In our system, when new observations are obtained, the old generalised theory is outdated and a new one is formed from our updated information.

This logic of generalisation thus is a framework to describe correct generalisations when our information about the world is incomplete (as it necessarily is) and the correct action when facing new information. What it does not take into account, however, are incorrect observations or incorrect conceptualisations of observations; for these, even more delicate means have to be developed.

### 2.2.1 Bayesianism

The Bayesian tradition is probably the most prominent among inductive frameworks. In this framework, each proposition is associated with a probability that changes dynamically as new observations are acquired. When a new hypothesis is formed, it is assigned a prior probability; as relevant information comes in, this probability may rise or sink. The latter is a soft form of "refutation".

It would be interesting to attempt to use the Bayesian approach with the hypothesis-formation techniques presented in this paper. However, the dynamic probabilities are heavily affected by the prior probabilities of the propositions. If we take our observations to be certain, the method degenerates into a flip-flop: an observation that is consistent with our hypothesis keeps the probability of the hypothesis, while a counterexample nullifies it.

I have chosen not to use probabilities in this paper to simplify the theory-based logic of inductive generalisation and to avoid having to invent an algorithm providing initial probabilities for hypotheses and observations.

### 2.2.2 Adaptive logic

Adaptive logics [Bat01] are a natural fit for inductive reasoning, but not constrained to it. An adaptive logic has a lower limit logic whose rules are safe in the sense that the conclusion unconditionally follows from the premisses, and an upper limit logic whose rules can be used to form tentative conclusions. Every proposition is associated with the propositions it depends on; if a proposition is contradicted, this information is used to track the contradiction back to its source(s). This is a kind of a constructive *reductio ad absurdum.*

The point of adaptive logics is to provide a framework that (1) matches the dynamics of human reasoning and (2) provides a way to construct maximally consistent consequence sets from rules of reasoning that would lead to inconsistent results if applied without constraints. There are many different adaptive logics that could be used as a basis for induction. These differ in choice of the lower limit logic, the upper limit logic, and the adaptive strategy that is used to retract propositions when contradictions (or "abnormalities", in the terminology of adaptive logics) occur.

It is possible to combine adaptive logics with the theory-based approach taken in this paper. The only reason for not doing so is to keep the theory-based logic of inductive generalisation as simple as possible. Theories, which are already needed for construction of disjunctive hypotheses, provide a sufficiently powerful tool for the needs of **TLIG**, so there is no need to use an adaptive logic as a backbone for the system. However, one possible direction of future research could be an adaptive logic of induction that employs some scheme of forming disjunctive generalisations that is not based on theories.

Batens [Bat05] proposes a straightforward logic of induction (**LI**), which heavily restricts the kind of generalisations we can make: they cannot contain constants and they are always of the form $(\forall xy\ldots)Pxy\ldots \rightarrow Qxy\ldots$, where $P$ and $Q$ are first-order logic clauses dependent on the variables $x$, $y$, ...

Meheus [Meh05] studies a logic of abduction $\mathbf{LA^k}$. This logic is based on adaptive logic and is specifically designed to allow for comparing already falsified generalisations (scientific theories). Adaptive logic is only used to invalidate old abductive explanations when they are proven trivial or when better explanations are available. Mutually inconsistent generalisations are handled with a modal logic: they are given a modality that prevents derivation of trivialities from them.

### 2.2.3 Theory-based approach

The theory-based logic of inductive generalisation presented in this paper relies on the concept of *theories*, as the name says. Instead of having a single set of propositions all with a similar ontological status, we have multiple sets, or theories, of propositions. The propositions in one theory are related by having the same context, or domain of discourse. Theories are related to each other in various ways that are explicitly specified.

Theories, also called contexts, situations, and microtheories, were developed in artificial intelligence research to formalise the way people use logic (see [McC87], further formalised in [McC93]), but have similarities with models and possible world semantics. Usually, whenever we have a statement whether written in a logical language or not, it is relative to some *context.* It is almost always possible to discover another context where the statement is false. So it makes sense to group statements together when they are meant to be interpreted in the same context.

The way theories are used as the framework for induction in this paper, there are two kinds of theories: observational theories that describe observations of particular situations, and generalised theories produced from these. The propositions that result from generalisations are never "retracted", since they are always sensible with respect to the data they were originally constructed from; on the other hand, a theory that might have been relevant at some stage can become outdated when new observation data are available.

## 2.3 Construction of hypotheses

Mitchell's concept of version space [Mit82, 204] provides a theoretical framework for talking about possible hypotheses. The version space contains different versions of *generalisations*, each of which matches a set of *instances*. The task is to find the generalisation(s) that match(es) all of a set of positive instances while matching none of a set of negative instances.

For induction, a generalisation is a logical formula and an instance is a conceptualisation of an observation (which is also a logical formula). The positive instances are our observations and negative instances are their negations. A logical formula "matches" an instance when it logically implies it. This allows us to reformulate the problem of finding a correct inductive generalisation thus: given a set of propositions that describe observations, produce a logical formula that implies these propositions while being consistent with them, i.e. not implying their negations. It may be noted that since the generalisation must imply our observations, it would have to be inconsistent itself in order to imply their negations. Consequently, a generalisation is simply a consistent logical formula that implies the propositions that describe our observations.

Such an inductive generalisation is usually not uniquely defined. Every (consistent) theory has as its generalisation at least the theory itself, while other generalisations are possible. Consequently, we choose the *most general generalisation* (MGG) as the most interesting of these. The MGG is a clause $c \in G$ such that $(\forall g \in G)(c \vdash g)$, where $G$ is the set of generalisations that are correct for our observations.

The generality of logical formulae forms a partial order (which is the same as the partial order of entailment). In the case of non-recursive predicates, this partial order is very elegantly described by the notion of *$\theta$-subsumption*: a clause $c_1$ $\theta$-subsumes another clause $c_2$ if there is a substitution $s$ from variables to arbitrary terms such that the disjuncts of $c_1[s]$ are an improper subset of disjuncts in $c_2$. The significance of $\theta$-subsumption is that given a clause $c$, one can algorithmically construct all (non-recursive) clauses that imply $c$ or that are implied by $c$. Plotkin [Plo71] was the first one to notice the difference between $\theta$-subsumption and implication. A thorough analysis of $\theta$-subsumption and implication is in [IA93].

Most attempts to make systems that synthesise hypotheses come from artificial intelligence. This section studies these approaches.

### 2.3.1 Top-down $\theta$-subsumption

Under $\theta$-subsumption, a clause can be made less general in two ways: (1) by substituting all occurrences of a given variable in it by some term (e.g. a variable already in use, a constant or a function expression) or (2) by adding a disjunct to it. Shapiro's Model Inference System [Sha81] induced logic programs by simply searching the version space from the most general predicate definition (one where all parameters are universally quantified) towards the specialisations in order of "simplicity". The search was complete when a definition was found that implied all the positive examples while implying none of the negative ones.

The problem was formulated to better match the formalism, Prolog, where this search took place. Clauses were restricted to Horn clauses and the solution was not required to imply the negations of negative examples, not implying them was sufficient. The system is elegant in its own way but very inefficient, because the space of possible clauses is enormous: it contains all possible predicate definitions. Especially the adding of disjuncts (called *literals* in Prolog parlance) causes the search tree fork a lot.

Top-down search in the version space can be improved by generating more efficient search techniques. Since these optimisations in search mean delaying some possibilities or not considering them at all, they present a difficult balancing problem between generating too many results and omitting some interesting results.

### 2.3.2 Inverse resolution

It is wasteful to search the whole version space for solutions, when the examples already give hints about what kind of generalisations could be relevant for the current case. By inverting rules of deduction, we can construct from a proposition $p$ those propositions that logically entail $p$. This is the approach in Muggleton's CIGOL [MB88] which inverts the deductive rule of resolution in various ways in order to obtain more general clauses as explanations of the examples.

This approach resembles the one taken in this paper, because in the theory-based logic of inductive generalisation, generalisations are constructed by moving up the $\theta$-subsumption hierarchy. Under $\theta$-subsumption, a clause can be made more general in two ways: (1) by replacing some or all occurrences of a term $t$ with a new (previously unused) universally quantified variable or (2) by taking away a disjunct from it. However, since every example is non-disjunctive, this means that in practice only (1) is possible, which is equivalent to inverting the rule of universal quantification elimination.

CIGOL is also interesting in that it creates new concepts (predicate definitions) heuristically. New concepts are created when they express the information content of the examples more succinctly than the previous predicates. Human intervention is of course required to give these concepts meaningful names (to humans, that is).

### 2.3.3 Relative least-general generalisations

Another method for constructing generalisations from the examples is combining the examples. The relative least-general generalisation (RLGG) of two propositions $p$ and $q$ is the most specific clause that, with a given background theory

$\Gamma$, implies both $p$ and $q$. The program GOLEM [MF90] finds generalisations by combining examples in this way.

RLGG's expose the problem of combinatorial explosion in another way: the propositions constructed by RLGG can be very long (even infinite with some background theories). They are also completely dependent on a sensible background theory, because the resulting proposition has to imply both $p$ and $q$ without referring to either. The theory-based logic of inductive generalisation does not aim so high: the theory may include propositions about individuals, so examples can be used as explanations of one another.

# 3 How TLIG works

When searching generalisations, it is easy and efficient to find basic generalisations of the examples — those that are obtained by universally quantifying some terms in the examples. However, these generalisations are insufficient to describe the truth conditions of a predicate, because the truth of a predicate might depend on the truth of another. So, additionally we need a means to produce disjunctions of the ground clauses that might be causally connected to enable generalisation from these disjunctions.

The idea of **TLIG** is to use the differences between observational theories to find these causal dependencies. Given two theories $\Gamma$ and $\Delta$, we construct disjunctions of propositions that are found in one theory but not the other.

The process of forming generalised theories from observations proceeds by the following steps:

1. Simple observational theories are formed from observations by conceptualising them.

2. The basic theories are merged to obtain more elaborate observational theories that describe multiple observations. These theories only contain information that is irrefutably true within the context of the observations described.

3. In the process of merging, disjunctions are formed between propositions that could have a dependency.

4. The combined theory of all observations is used to form generalisations: propositions that are more general (under $\theta$-subsumption) than those which describe the observations.

5. From these generalisations, those are filtered out that are incompatible with our observations.

## 3.1 Forming the initial theories

Well-known examples of propositions in classical first-order predicate logic (**FOL**) do not show how **FOL** can be used to describe observations for several reasons. For instance, the classic proposition

$$\text{Man}(\text{Socrates})$$

which states that Socrates is a man, cannot describe information obtained from an observation for two reasons.

Firstly, the proposition is *too broad* to be the result of an observation: it effectively claims that Socrates is a man independent of the situation, world, or domain of speech of the claim. An observation does not have enough information to rule out the possibility that Socrates will turn into a plant during night-time or when no one sees.

Secondly, the proposition makes an ungrounded assumption about the *valence* (or *arity*) of the predicate "Man". Semantically, this assumption translates into a conviction that being a man does not depend on external factors. For instance, whether Socrates is a man does not depend on who is judging. This might be unproblematic in the case of "man" (although I doubt it). However, it is not uncommon for further investigation to show that supposedly unary predicates — *attributes*, such as "good" — prove to be binary relations, and relations in general prove to have a greater valence than was supposed.

The problems can be solved by introducing the *event* of each observation as the first parameter of every predicate. An event constant uniquely identifies the observation in which the relation was observed. Adding it as a parameter is effectively equivalent to claiming that our proposition is particular to a certain observation. Because every situation we might speak of has its unique set of events, the event holds information about everything that might be relevant to the truth of our proposition.

It is quite clear that the event parameter prevents the proposition from being too broad. What is not so evident is that the event parameter also solves the problem of unknown valence. This is because every proposition that has too few parameters (for example, $\mathrm{Good}(e_6, \mathrm{Sunshine})$) is in practice universally quantified with respect to a parameter that it should not be. Thus, there is an unnoticed condition that needs to hold for the proposition to be true (for example, $\mathrm{OpinionOf}(e_6, \mathrm{Panu})$). So the proposition that really holds is actually $\mathrm{Good}(e_6, \mathrm{Sunshine}) \vee \neg\mathrm{OpinionOf}(e_6, \mathrm{Panu})$, and too low a valence is equivalent with not noticing some condition that affects the truth of a predicate.

## 3.2 Explorative logic

Explorative logic (**EL**) is the part of **TLIG** that is used in observational theories. For generalisations, full **FOL** is used. **EL** is a fragment of **FOL**: it only contains ground clauses, so it is in practice a logic over a Herbrand universe.

### 3.2.1 An overview of EL

The explorative logic of generalisation (**EL**) deals with information that can be safely concluded from experience. In practice, **EL** is a simplified version of **FOL** accompanied with the theory-merging operation, $\oplus$.

**EL** differs from normal **FOL** in the following ways:

1. **EL** does not deal with universally quantified variables, because no observation could possibly justify a universal claim.

2. Propositions are kept in negation normal form (*NNF*) and conjunction normal form (*CNF*). This means that negation symbols are only found in

front of atomic formulae and theories are sets of disjunctions of atomic formulae.

3. Constants are used for both known and unknown individuals, so there is no need for existentially quantified variables, and, consequently, for variables at all.

Theories are merged for two reasons: to ultimately gather information about experience into one theory, and to construct good candidates for disjunctive rules that hold across all theories merged.

### 3.2.2 Formal definition of EL

As the *vocabulary* of **EL** we have:

- A set of *constants*, here written with $a$, $b$, $c$, $\ldots$ including the *event constants* $e_1$, $e_2$, $\ldots$
- A set of *function constants*, here written with $f$, $g$, $h$, $\ldots$
- A set of *predicate constants*, here written with $P$, $Q$, $R$, $\ldots$
- *Logical operators*: $\neg$ (negation), $\vee$ (disjunction)
- Parentheses: "(" and ")", used to denote precedence

Expressions are defined thus:

1. A *term* is a constant or an expression $(fab\ldots)$ where $f$ is a function symbol and $a$, $b$, $\ldots$ are terms.
2. If $P$ is a predicate constant, $e$ is an event constant and $a$, $b$, $\ldots$ are terms, then $Peab\ldots$ and $\neg Peab\ldots$ are *atomic formulae*.
3. A *formula* is either an atomic formula or $(p \vee q)$ where $p$ and $q$ are formulae.
4. A proposition is a formula.

**EL** deals only with observational theories. There are no rules of deduction in **EL**. The only activity that resembles reasoning is the process of merging theories and generalising from them.

### 3.2.3 Examples

Suppose we have an event $e_1$ where we observe sunshine. We could conceptualise this observation as the proposition $Se_1$, where $S$ means sunshine and $e_1$ is the constant that names this particular event. On the other hand, there could be another event $e_2$ where we observe lack of sunshine, giving $\neg Se_2$.

A proposition need not be about an event in general. We can use constants to denote concepts within and across events. For instance, if we observe that Peter is taller than Tom in an event $e$, we can conceptualise this as $Tepj$ ($T$ means taller, $e$ means this event, $p$ means Peter, and $j$ means John). If we observe the same fact in another event $e_2$, we can conceptualise it as $Te_2pj$.

## 3.3 Operations between theories

The operations between theories form the backbone of the framework of logic of generalisation. They include $\oplus$ (*merge*) and gen (*generalise*), presented below. Merging produces new observational theories that cover the situations covered by earlier theories while being careful to retain the truth of all claims; generalisation uses the results of merging to find general rules that are undefeated as yet. The theories that are arguments of $\oplus$, as well as the result, only contain clauses in **EL**; however, results of gen contain clauses in full **FOL**.

### 3.3.1 Definition of $\oplus$

An observational theory $\Gamma$ is a set of propositions that describes a particular situation. The merging of theories, $\Gamma \oplus \Delta$,[1] produces a new theory that describes a situation that is a union of the situations of $\Gamma$ and $\Delta$. The merged theory includes relevant information that can be derived from $\Gamma$ and $\Delta$ collectively.

Let us denote the *terms* of a proposition $p$ by $\mathrm{terms}(p)$. A *event substitution* of a proposition $p$ is $p[t/e]$ where $t \in \mathrm{terms}(p)$ corresponds to an event (is in the first place of some predicate), $e$ is an event constant and $p[t/e]$ means $p$ with occurrences of $t$ replaced by $e$. An event substitution is *safe* for a theory $\Gamma$ if it does not substitute with an event constant that is already in use in $\Gamma$. The *substitution set* $\mathrm{ss}(\Gamma, e)$ is defined as:

$$\mathrm{ss}(\Gamma, e) = \{p[t/e] \mid p \in \Gamma\}$$

where $p[t/e]$ is an event substitution that is safe with respect to $\Gamma$. The set of *possibly causally relevant substitutions* $\mathrm{pcrs}(\Gamma, \Delta, e)$ filters out irrelevant formulae from the substitution set:

$$\begin{aligned}\mathrm{pcrs}(\Gamma, \Delta, e) \quad &= \quad \{p \mid p \in \mathrm{ss}(\Gamma, e), p \notin \mathrm{ss}(\Delta, e)\} \\ &\cup \quad \{p \mid p \in \mathrm{ss}(\Gamma, e), \neg p \in \mathrm{ss}(\Delta, e), \neg p \notin \mathrm{ss}(\Gamma, e)\}\end{aligned}$$

The set of *rule candidates* $\mathrm{rc}(\Gamma, \Delta)$ is given as disjunctions of causally relevant propositions in $\Gamma$ and $\Delta$ with different events in different theories replaced by a common event:

$$\mathrm{rc}(\Gamma, \Delta) = \{p \vee q \mid p \in \mathrm{pcrs}(\Gamma, \Delta, e), q \in \mathrm{pcrs}(\Delta, \Gamma, e)\}$$

where $e$ is a new event constant. This allows us to devise a definition of $\oplus$:

$$\Gamma \oplus \Delta = \Gamma \cup \Delta \cup \mathrm{rc}(\Gamma, \Delta)$$

### 3.3.2 Rationale behind $\oplus$

A simple union of propositions is not sufficient for merging theories. It disregards potentially interesting *rules* that arise from the possibility that differences between theories are causally relevant. However, although it is not sufficient to include the propositions from the argument theories in the resultant theory, it

[1]The symbol $\oplus$ is an arbitrary choice for the merge operation which does not have an allegoric well-known operation.

is safe to do so. This is why: a proposition $p \in \Gamma$ is about an event $e$ (known or unknown) that belongs to the situation described by $\Gamma$. Since the event must also belong to the situation described by $\Gamma \oplus \Delta$, $p \in \Gamma$ implies $p \in \Gamma \oplus \Delta$.

The propositions that are good candidates for building rules are disjunctions where:

1. One of the disjuncts comes from $\Gamma$ and the other from $\Delta$.

2. The disjuncts talk about at least one common individual, so they depict a dependency between two predicates.

In order to obtain more propositions that satisfy condition (2), we can substitute terms of propositions in $\Gamma$ and $\Delta$ with constants not previously occurring in $\Gamma$ or $\Delta$. This is equivalent to existence introduction and then weakening the result by disjunction introduction. For instance, if $\Gamma$ has $Ce_1$ ("$e_1$ is counterfactual") and $\Delta$ has $Fe_2$ ("$e_2$ is factual"), these can be rewritten to $Ce'$ and $Fe'$ to form a disjunction that may be later generalised to showing a causal relationship between counterfactuality and factuality.

In principle, we could generate all possible disjunctions whose disjuncts are some propositions of $\Gamma$ and $\Delta$ with constant substitutions. However, it turns out that the number of generated disjunctions can be drastically reduced without losing any causally relevant ones by two simple measures: considering which terms are *connectable*, and filtering out disjuncts that are causally irrelevant.

There are many possible criteria as to which terms are connectable, i.e. good candidates to substitute with a common new constant that makes the formulae satisfy condition (2). In this paper I am especially interested in finding causal relationships that hold across events, so the $\oplus$ operation presented here only connects event parameters. Other options include connecting all terms that occur in the same place of a predicate or function, or employing a type framework to cut down insensible connections.

For reasoning about causal relationships, I defined the concept of *substitution set*. A substitution set of a theory $\Gamma$ is the set of propositions that is obtained from $\Gamma$ by substituting mutually connectable parameters by a common constant. For any proposition $p$, the substitution set may lack it, include it, include its negation, or include both. If the substitution set has both an affirmative and negative version of a proposition, that proposition is *factual* in theory $\Gamma$ with respect to the connected parameter, i.e. its truth depends on the value of the substituted parameter.

It is possible to filter out causally irrelevant propositions by comparing the substitution sets of the theories being merged. If the status of a proposition $p$ is the same in both substitution sets (absent, affirmative, negative, or factual), it cannot be used to explain differences in other propositions and so is left out. And if the proposition is factual in one and affirmative / negative in the other, the only interesting disjunct on the factual side is the one that is the complement from the other side.

For instance, if $\Gamma$ has $He_1x$ ("$x$ is a human in event $e_1$"), and $\Delta$ has $De_2y$ ("$y$ is a dog in event $e_2$"), then $\Gamma \oplus \Delta$ should have $(He'x \vee De'y)$ ("there is either a dog $y$ or a human $x$ in some event"). But if $\Delta$ also has $He_2x$, we can safely conclude that the humanity of $x$ is irrelevant to the dogness of $y$. On another occasion, if $\Gamma$ has $He_1x$ and $\Delta$ has $De_2x$, then $\Gamma \oplus \Delta$ should have $(He'x \vee De'x)$ ("$x$ is either a human or a dog in some event.")

### 3.3.3 Generalising

When we have the information gathered by merging, it is relatively easy to devise the generalisation operation. A generalisation of a theory $\text{gen}(\Gamma)$ is simply the set of universal propositions that have examples but no counterexamples in $\Gamma$. The argument of $\text{gen}(\Gamma)$ is an observational theory in **EL**, but the result is a generalised theory in **FOL**.

A *generalising substitution* of a proposition $p$ is $(\forall\alpha)p[t/\alpha]$, where $t \in \text{terms}(p)$, $\alpha$ is an universally quantified variable previously unused in $p$ and $p[t/\alpha]$ means $p$ with occurrences of $t$ replaced by $\alpha$. The generalisations of a proposition $p$ is defined as:

$$\text{gen}(p) = \{p\} \cup \{(\forall\alpha)q[t/\alpha] \mid q \in \text{gen}(p)\}$$

where $(\forall\alpha)q[t/\alpha]$ is a generalising substitution. The generalisations of a theory $\Gamma$ are defined as:

$$\text{gen}(\Gamma) = \{q \mid q \in \text{gen}(p), p \in \Gamma, \Gamma \cup \{q\} \nvdash \bot\}$$

Then, let us look at the rationale behind gen. When we have a particular proposition making an example of a universal proposition, we can form a more general proposition simply by substituting some or all occurrences of a term $t$ with an universally quantified variable. This process gives us a new proposition that $\theta$-subsumes our original proposition; the process can be repeated until all terms in the formula are universally quantified variables. For instance, if we have the claim $Hep$ ("Panu is a human in event $e$"), the corresponding generalisations are $(\forall\alpha)H\alpha p$ ("Panu is (always) human"), $(\forall\alpha)He\alpha$ ("everything is human in event $e$") and $(\forall\alpha)(\forall\beta)H\alpha\beta$ ("everything is human").

While partial substitutions (substitutions where only some of the occurrences of a term $t$) are possible, they generate uninteresting results because they break causal relations between disjuncts. If the resulting "rule" is consistent, it can only be used to prove things that are already known. This is because merging leaves the original disjuncts in the merged theory, so they are available to directly generalise from. Thus, we disregard partial substitutions.

If a generalisation has counterexamples, it forms a contradiction when combined with $\Gamma$. This gives us a simple way to filter out generalisations that have counterexamples.

## 3.4 Examples: putting it all together

In this section, we walk through a couple of examples to see what kind of results **TLIG** yields.

### 3.4.1 Causal connections of weather

For this example, we pick the following predicates: $Se$ means sunshine in event $e$, $Ce$ means cloudy weather in event $e$ and $Re$ means rain in event $e$. An observation of a cloudy day is conceptualised as:

$$\Gamma_1 = \{\neg Se_1, Ce_1, \neg Re_1\}$$

If this is all the data we have, our generalised theory looks like this:

$$\text{gen}(\Gamma_1) = \{(\forall\alpha)\neg S\alpha, (\forall\alpha)C\alpha, (\forall\alpha)\neg R\alpha\}$$

Another similar day does not alter our generalisations. Let $\Gamma_2$ be a similar conceptualisation about another event, $e_2$. Then the merged theory contains no additional disjunctions, because there are no differences in the event-substituted clauses of $\Gamma_1$ and $\Gamma_2$:

$$\Gamma_1 \oplus \Gamma_2 = \{\neg Se_1, Ce_1, \neg Re_1, \neg Se_2, Ce_2, \neg Re_2\}$$

The generalised theory $\text{gen}(\Gamma_1 \oplus \Gamma_2) = \text{gen}(\Gamma_1) = \text{gen}(\Gamma_2)$. But suppose there is another day when it is raining in addition to being cloudy. We have:

$$\Gamma_3 = \{\neg Se_3, Ce_3, Re_3\}$$

Now $\Gamma_1 \oplus \Gamma_3$ has a new disjunction, $Re' \vee \neg Re'$, but this can be dropped out because it is a tautology. Thus $\Gamma_1 \oplus \Gamma_3$ is effectively the union of the theories and

$$\text{gen}(\Gamma_1 \oplus \Gamma_3) = \{(\forall\alpha)C\alpha, (\forall\alpha)\neg S\alpha, \neg Re_1, Re_3\}$$

because any more general proposition about rain is refuted by one or the other of the particular statements.

Then, let us have a sunny day. With this, we have:

$$\begin{aligned}
\Gamma_4 &= \{Se_4, \neg Ce_4, \neg Re_4\} \\
\Gamma_4 \oplus (\Gamma_1 \oplus \Gamma_3) &= \{Re' \vee Se', Se' \vee Ce', \neg Ce' \vee Re', \\
&\quad\ \neg Ce' \vee \neg Se', \neg Re' \vee \neg Se', \neg Re' \vee Ce'\} \\
&\cup\ \ \Gamma_1 \cup \Gamma_3 \cup \Gamma_4 \\
\text{gen}(\Gamma_4 \oplus (\Gamma_1 \oplus \Gamma_3)) &= \{(\forall\alpha)C\alpha \vee \neg R\alpha, \\
&\quad\ (\forall\alpha)\neg R\alpha \vee \neg S\alpha, \\
&\quad\ (\forall\alpha)\neg C\alpha \vee \neg S\alpha, \\
&\quad\ (\forall\alpha)C\alpha \vee S\alpha, \\
&\quad\ Re' \vee Se', \neg Ce' \vee Re'\} \\
&\cup\ \ \Gamma_1 \cup \Gamma_3 \cup \Gamma_4
\end{aligned}$$

This lets us note a couple of interesting things. First, the ground clauses of the observations are present in the generalised theory, because the general clauses no longer imply them. However, if any single singular proposition is taken out, the others imply it *together* with the generalisations if this is possible (it is not in the case of $Re_3$ which has to be taken as factual in the light of our observations). In a way, the observations work both as the explanans for each other and the explananda.

Second, there are now some leftover disjunctions ($Re' \vee Se'$, $\neg Ce' \vee Re'$) in our generalised theory that are not subsumed by any general rule. These are side products of the merging process: rule candidates whose universality cannot be proven impossible by looking at the substitution sets of different situations. However, when a generalisation is attempted from them, the generalisations are refuted by instances in the observation data.

Third, we have as broad generalisations as possible because **TLIG** constructs every (event-relative) generalisation that has not been refuted. For instance, we already have the rules that rain implies cloudy weather, rain and sunshine are mutually exclusive, and clouds are equivalent with no sunshine.

### 3.4.2 Observations of the physical height of two individuals

Another example, which demonstrates how non-unary predicates work, is a set of observations about two individuals. I have chosen a subjective conceptualisation to better match the kind of data that an independent agent might have. There are two predicates, $Texy$ which means "$x$ looks taller than $y$ in event $e$" and $Nexy$ which means "$x$ is nearer (the observer) than $y$ in event $e$", and two individuals, John ($j$) and Peter ($p$). John is actually taller than Peter, so he looks taller whenever he is nearer than Peter. With this, we may have the following observations:

$$
\begin{aligned}
\Gamma_1 &= \{Te_1jp, \neg Te_1pj, Ne_1jp, \neg Ne_1pj\} \\
\Gamma_2 &= \{Te_2pj, \neg Te_2jp, Ne_2pj, \neg Ne_2jp\} \\
\Gamma_3 &= \{Te_3jp, \neg Te_3pj, Ne_3pj, \neg Ne_3jp\}
\end{aligned}
$$

The merged theory looks like this:

$$
\begin{aligned}
\Gamma_1 \oplus (\Gamma_2 \oplus \Gamma_3) \quad = \quad & \{(Te'pj) \vee (Te'jp), (Te'jp) \vee (Ne'pj), (Te'jp) \vee (\neg Ne'jp), \\
& (\neg Te'jp) \vee (\neg Te'pj), (\neg Te'pj) \vee (Ne'pj), (\neg Te'pj) \vee (\neg Ne'jp), \\
& (Ne'jp) \vee (Te'pj), (Ne'jp) \vee (\neg Te'jp), (Ne'jp) \vee (Ne'pj), \\
& (\neg Ne'pj) \vee (Te'pj), (\neg Ne'pj) \vee (\neg Te'jp), (\neg Ne'pj) \vee (\neg Ne'jp)\} \\
\cup \quad & \Gamma_1 \cup \Gamma_2 \cup \Gamma_3
\end{aligned}
$$

The generalised theory is too long to clearly present here — 40 clauses not $\theta$-subsumed by each other. The added number of clauses is mostly due to more places in predicates. This leads to more partial generalisations, such as generalised sentences particular to Peter. There are three kinds of generalisations in the theory:

1. Generalisations with no constants (except predicate constants), that usually point to general treats of the predicates involved, such as

   $$(\forall\alpha)(\forall\beta)(\forall\gamma)\neg N\alpha\beta\gamma \vee \neg N\alpha\gamma\beta$$

   (something cannot be nearer than a thing that is nearer than it).

2. Non-disjunctive generalisations about individuals. In our theory, the only individuals that allow such generalisations are events; for example, $(\forall\alpha)Te_3j\alpha$ (John looks taller than anything else in $e_3$).

3. Disjunctive *rules* about individuals, such as $(\forall\alpha)(\forall\beta)T\alpha j\beta \vee \neg N\alpha j\beta$ (if John is nearer than something else, he always looks taller).

The third kind of generalisation is very interesting. It cannot be used to explain individual facts about our observations as such because it doesn't imply them, but it gives a clue that there is a property which the individual $j$ has and which is not universal. This can be used as a base for concept formation. To add the new concept to our theory, we can augment the rule into $(\forall\alpha)(\forall\beta)(\forall\gamma)\neg T'\gamma \vee T\alpha\gamma\beta \vee \neg N\alpha\gamma\beta$ where $T'$ is a new predicate (intuitively meaning "tall"), and add the fact that $T'j$ (John is tall).

# 4 Significance of TLIG

While **TLIG** is mostly based on earlier work especially in the field of inductive logic programming, it also presents some novel ideas. These ideas and open questions are studied in this section.

## 4.1 Theories as a framework

The concept of theory contributes an enormous practical and theoretical help to this paper. The idea is that no proposition is ever "just there": it always belongs to some theory that is relative to some context. A context is a (usually partially unspecified) set of background assumptions.

The observational theories in **TLIG** have the background assumption that they describe observations of some particular situation. The definition of $\oplus$ is heavily motivated by intuition about how rules can be formed by observing differences between situations. In the case of observational theories, the theory framework provides a way to reason about situations.

On the other hand, the generalised theories in **TLIG** have the background assumption that they are thought experiments (not unrefutably true) that hold universally true in the light of observations that we have about some particular situation. In this case, theories provide a way to tentatively propose generalisations without messing our observational data with them. These generalised theories could be combined with other observational data to make predictions.

In general I think that theories capture an important aspect of human thinking. **TLIG** can be seen as an application that demonstrates the practical and theoretical value of theories and operations between them. One possible direction of future work is a reformalisation of **TLIG** in other formalisms, such as adaptive logic.

## 4.2 The $\oplus$ and gen operations

Perhaps the most original feature of **TLIG** is the $\oplus$ operation, which constructs *rule candidates* that might be *causally relevant*. Further work may define what it actually means for a disjunction to be "causally relevant". Then it would be possible to show whether or not $\oplus$ actually produces all such rule candidates.

However, the basic intuition behind causal relevance is this: a disjunction is causally relevant when it has at least one generalisation that holds over the set of our observational data, while none of its disjuncts do. That is, a causally relevant rule candidate is a minimal disjunction whose generalisation is unrefuted.

The $\oplus$ operation presented in this paper only rewrites, or *connects*, event parameters in order to keep the combinatorial explosion of disjunctions down.

However, a more sophisticated connecting scheme could provide more interesting rules.

The generalisation operation gen forms generalisations by moving up the $\theta$-subsumption hierarchy. However, this does not produce all clauses that imply the original proposition, since in the case of recursive predicates, $\theta$-subsumption may miss some generalisations. It should be possible to extend gen to include more generalisations; inverting implication in general is studied in [Mug92].

### 4.3 Concept formation

As noticed in the height observation example, constants that are left in $\text{gen}(\Gamma)$ imply that there is something special about these individuals. It is possible to form new concepts in the following way:

1. All particular claims that have a similar form up to variable and constant renamings are gathered together.
2. These claims are replaced by one generic rule by introducing a new predicate whose truth for particular individuals implies that the rule holds for those individuals.
3. The new predicate is then used to enumerate those individuals for which the rule holds.

### 4.4 Mistakes in conceptualisation

The observational theories in **TLIG** are taken to be irrefutably true, but this actually holds only as far as the conceptualisations of our observations are correct. It would be interesting to extend **TLIG** with measures to deal with problems of incorrect conceptualisation such as accidental homonymy or synonymy.

In the field of computer ontologies, contexts (or theories) are sometimes used to combine ontologies that use constants differently; see e.g. [dLdM05]. This suggests that it would be possible to augment the $\oplus$ operation to be flexible enough to combine sensibly observational theories that use constants (names) differently.

Another direction is to make gen somewhat more resilient towards incorrect observations. For instance, it would be possible to keep track of the number of rules that a particular observation refutes. If the number of such rules is very high, that could be grounds to assume that the observation is incorrect, so the observation would be refuted instead.

## References


[Bat01] Diderik Batens. A general characterization of adaptive logics. *Logique et Analyse*, 173–175:45–68, 2001. Appeared 2003.

[Bat05] Diderik Batens. On a logic of induction. In Festa et al. [FAP05], pages 221–242.

[dLdM05] P. de Leenheer and A. de Moor. Context-driven disambiguation in ontology elicitation. In P. Shvaiko and J. Euzenat, editors, *Context* *and Ontologies: Theory, Practice and Applications*, volume WS-05-01 of *AAAI Technical Report*, pages 17 – 24. AAAI Press, 2005.

[FAP05] Roberto Festa, Atocha Aliseda, and Jeanne Peijnenburg, editors. *Confirmation, Empirical Progress, and Truth Approximation. Essays in Debate with Theo Kuipers. Vol. 1*, volume 83 of *Poznan Studies in the Philosophy of the Sciences and the Humanities*. Rodopi, Amsterdam/New York, 2005.

[Fla95] Peter Flach. *Conjectures – an inquiry concerning the logic of induction*. Institute for Language Technology and Artificial Intelligence, April 1995.

[Hem43] C.G. Hempel. A purely syntactical definition of confirmation. *Journal of Symbolic Logic*, 6(4):122–143, 1943.

[IA93] P. Idestam-Almquist. *Generalization of clauses*. PhD thesis, Stockholm University, Department of Computer and Systems Sciences, 1993.

[MB88] S.H. Muggleton and W. Buntine. Machine invention of first-order predicates by inverting resolution. In *Proceedings of the 5th International Conference on Machine Learning*, pages 339–352. Kaufmann, 1988.

[McC87] John L. McCarthy. Generality in artificial intelligence. *Communications of the Association of Computing Machinery*, 30(12):1029–1035, 1987.

[McC93] John McCarthy. Notes on formalizing context. In *International Joint Conference on Artificial Intelligence*, pages 555–562, 1993.

[Meh05] Joke Meheus. Empirical progress and ampliative adaptive logics. In Festa et al. [FAP05], pages 193–217.

[MF90] S.H. Muggleton and C. Feng. Efficient induction of logic programs. In *Proceedings of the First Conference on Algorithmic Learning Theory*, pages 368–381, Tokyo, 1990. Ohmsha.

[Mit82] Tom M. Mitchell. Generalization as search. *Artificial Intelligence*, 18(2):203–226, 1982.

[Mug92] S.H. Muggleton. Inverting implication. In *Proceedings of the Second Inductive Logic Programming Workshop*, pages 19–39, Tokyo, 1992. ICOT (Technical report TM-1182).

[Plo71] G.D. Plotkin. *Automatic Methods of Inductive Inference*. PhD thesis, Edinburgh University, 1971.

[Sha81] Ehud Y. Shapiro. The model inference system. In Patrick J. Hayes, editor, *International Joint Conference on Artificial Intelligence*, page 1064. William Kaufmann, August 1981.