Induction in Linear Logic

Fumihiko Yamaguchi¹ and Masakazu Nakanishi¹

Received April 30, 1996

Linear logic, introduced by Girard *et al.*, has a great power of expression, but no method for induction. This paper proposes a method of induction using knowledge represented by linear logical formulas. In linear logic, the number of propositions is controlled by logical operators. When a background theory and a hypothesis prove an example, the number of propositions on each side must be equivalent.

1. INTRODUCTION

Linear logic, introduced by Girard *et al.* (1995), has a great power of expression, but no method of induction. Induction is an operation of inferring hypotheses. A hypothesis and a given background theory should explain given examples. Examples are classified as positive examples and negative examples. Positive examples are propositions that some events obey, by the background theory and the hypothesis. Negative examples are propositions that some events do not follow, and that are consistent with the background theory and the hypothesis.

Some incremental inductive methods based on traditional logic employ the first given positive example as the first hypothesis. However, in linear logic, the first positive example is not always the first hypothesis. Since the number of propositions is controlled by logical operators, induction must make the accounts balance.

Inductive operations are considered to be the inverse of deductive operations, because deduction can infer a positive example from a background theory and the hypothesis. For example, in the field of inductive logic programming (ILP), some inductive operations which are inverses of one or two

2107

Department of Computer Science, Keio University, Hiyoshi 3-14-1, Yokohama, Kanagawa, Japan; e-mail: yamagu@nak.math.keio.ac.jp.

steps of resolutions are employed (Muggleton and De Raedt, 1994). In this paper, some inductive operators for positive examples are defined. Those operators translate the succedent to the antecedent.

In this paper, we define linear logic in Section 2. Section 3 defines induction. In Section 4, we explain the syntactical inductive operations. We explain the case of one positive example in Section 4, and the case of plural positive examples in Section 5.

2. LINEAR LOGIC

Linear logic is defined by formation rules and inference rules as follows.

Definition 2.1. Formula. Let \mathcal{F} be the set of all formulas in linear logic. \mathcal{F} is defined as follows.

- Proposition $P \in \mathcal{F}$.
- Propositional constants \top , 1, \bot , $0 \in \mathcal{F}$.
- If $X \in \mathcal{F}$, then $!X, ?X, X^{\perp} \in \mathcal{F}$.
- If $X, Y \in \mathcal{F}$, then $X \otimes Y, X \not\in Y, X \multimap Y, X \oplus Y, X \& Y \in \mathcal{F}$.

In this paper we deal with the first-order predicate linear logic. Thus, propositions are defined as in traditional first-order logic including 0-adic predicates.

Conventionally, upper case Latin letters A, B, \ldots denote formulas, upper case Greek letters Γ, Δ, \ldots denote finite multisets of formulas. Parentheses are used to represent the order of connections.

Definition 2.2. Proof Tree. $\Gamma \vdash \Delta$ is a sequent which consists of antecedent Γ and succedent Δ . When Γ is the multiset $\{X_1, X_2, \ldots, X_n\}$ and Δ is the multiset $\{Y_1, Y_2, \ldots, Y_m\}$, the sequent $\Gamma \vdash \Delta$ has the same meaning as $X_1 \otimes X_2 \otimes \cdots \otimes X_n \multimap Y_1 \not P Y_2 \not P \cdots \not P Y_m$.

 Γ , A denotes a multiset Γ including at least one formula A.

Linear logic is defined by the initial sequents and rules of inference of Fig. 1.

The relation \vdash is reflexive because of the initial sequent $D \vdash D$, and transitive because of the cut rule. We define X = Y iff $X \vdash Y$ and $Y \vdash X$. Thus \vdash is the partial order on \mathcal{F} . When $X \vdash Y$, X is smaller than Y. According to this order, 0 and \top are the minimum and the maximum element of \mathcal{F} , respectively.

Theorem 2.1. X & Y is the greatest lower bound of X and Y.

Proof. 1. X & Y is a lower bound of both X and Y:

$$\frac{X \vdash X}{X \And Y \vdash X} \quad \frac{Y \vdash Y}{X \And Y \vdash Y}$$

Initial Sequents	Rules of Inference
1. $D \vdash D$	$\Gamma \vdash \Delta, D$ $D, \Pi \vdash \Lambda$
2. Γ,0⊢ Δ	$\Gamma,\Pi\vdash\Delta,\Lambda$ (Cut)
3. $\Gamma \vdash \top, \Sigma$	$\Gamma \vdash \Delta, D$ $D, \Gamma \vdash \Lambda$
4. ⊥⊢	$\frac{1}{D^{\perp} \Gamma \vdash \Delta} \stackrel{(\perp \text{ left})}{=} \frac{1}{\Gamma \vdash \Delta D^{\perp}} \stackrel{(\perp \text{ right})}{=}$
5. ⊢1	
	$\frac{A, B, \Gamma \vdash \Delta}{A \otimes B, \Gamma \vdash \Delta} (\otimes \text{ left}) \qquad \frac{\Gamma \vdash \Delta, A \qquad \Pi \vdash \Lambda, B}{\Gamma, \Pi \vdash \Delta, \Lambda, A \otimes B} (\otimes \text{ right})$
	$\frac{A,\Gamma\vdash\Delta}{A\mathcal{P}B,\Gamma,\Pi\vdash\Delta,\Lambda} \stackrel{B,\Pi\vdash\Lambda}{(\mathcal{P} \text{ left})} \qquad \frac{\Gamma\vdash\Delta,A,B}{\Gamma\vdash\Delta,A\mathcal{P}B} (\mathcal{P} \text{ right})$
Α, ΓΗΔ Α& Β, ΓΗ	$\frac{1}{\Delta} (\& \text{ left}) \frac{B, \Gamma \vdash \Delta}{A \& B, \Gamma \vdash \Delta} (\& \text{ left}) \qquad \frac{\Gamma \vdash \Delta, A \qquad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \& B} (\& \text{ right})$
$\frac{A,\Gamma\vdash\Delta}{A\oplus B,}$	$\frac{B,\Gamma\vdash\Delta}{\Gamma\vdash\Delta} (\oplus \text{ left}) \qquad \frac{\Gamma\vdash\Delta,A}{\Gamma\vdash\Delta,A\oplus B} (\oplus \text{ right}) \frac{\Gamma\vdash\Delta,B}{\Gamma\vdash\Delta,A\oplus B} (\oplus \text{ right})$
-	$\frac{3,\Gamma\vdash\Delta}{A\multimap B,\Gamma,\Pi\vdash\Delta,\Lambda} (\multimap \text{ left}) \qquad \frac{A,\Gamma\vdash\Delta,B}{\Gamma\vdash\Delta,A\multimap B} (\multimap \text{ right})$
<u>ΓΗΔ</u> !A,ΓΗ	$\frac{A}{\Delta} (! \text{ increase}) \qquad \frac{!A, !A, \Gamma \vdash \Delta}{!A, \Gamma \vdash \Delta} (! \text{ decrease}) \qquad \frac{A, \Gamma \vdash \Delta}{!A, \Gamma \vdash \Delta} (! \text{ left})$
<u>יר !</u> יר	$\frac{\vdash ?\Sigma, A}{\vdash ?\Sigma, !A} (! \text{ right}) \qquad \frac{\Gamma \vdash \Delta}{1, \Gamma \vdash \Delta} (1 \text{ left}) \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \bot} (\bot \text{ right})$
$\frac{\Gamma\vdash\Delta}{\Gamma\vdash\Delta,?A}$	$\frac{\Gamma \vdash \Delta, ?A, ?A}{\Gamma \vdash \Delta, ?A} (? \text{ decrease}) \qquad \frac{A, !\Gamma \vdash ?\Sigma}{?A, !\Gamma \vdash ?\Sigma} (? \text{ left})$
$\frac{\Gamma\vdash\Delta,A}{\Gamma\vdash\Delta,?A}$	(? right) $\frac{F[t/x], \Gamma \vdash \Delta}{\forall xF, \Gamma \vdash \Delta} (\forall \text{ left}) \qquad \frac{\Gamma \vdash \Delta, F[a/x]}{\Gamma \vdash \Delta, \forall xF} (\forall \text{ right})$ a is not free in Γ and Δ
	$\frac{F[a/x], \Gamma \vdash \Delta}{\exists x F, \Gamma \vdash \Delta} (\exists \text{ left}) \qquad \frac{\Gamma \vdash \Delta, F[t/x]}{\Gamma \vdash \Delta, \exists x F} (\exists \text{ right})$ a is not free in Γ and Δ

Fig. 1. Sequent calculus of linear logic.

2. When Z is a lower bound on both X and Y, X & Y is greater than Z:

$$\frac{Z \vdash X \quad Z \vdash Y}{Z \vdash X \& Y} \quad \blacksquare$$

Similarly, the least upper bound of X and Y is given using \oplus . Both operators & and \oplus satisfy the commutative law and the associative law. Moreover, they satisfy the absorption law. Therefore \mathcal{F} is a lattice.

3. INDUCTION

Induction is a process to induce the hypothesis that explains given examples.

Definition 3.1. Induction. $X \models Y$ denotes that X semantically entails Y. Let B represent the background theory, let E_i^+ $(1 \le i \le n)$ and $E_j^ (1 \le j \le m)$ represent the positive examples and the negative examples, respectively, and let them satisfy the following conditions:

for all *i*,
$$B \nvDash E_i^+$$

for all *j*, $B \otimes E_j^- \nvDash \bot$

A formula *H* which satisfies the following conditions is called an inductive conclusion of *B*, E_1^+, \ldots, E_n^+ , and E_1^-, \ldots, E_m^- :

```
for all i, B \otimes H \models E_i^+
for all j, B \otimes H \otimes E_i^- \nvDash \bot
```

Induction is a process to induce the inductive conclusion from the background theory and examples.

From the symbol processing point of view, the binary relation \vDash can be replaced with \vdash .

In this paper we consider especially induction for positive examples.

4. ONE POSITIVE EXAMPLE

One difficulty of induction in linear logic is that the first given positive example is not always the inductive conclusion.

Theorem 4.1. Let E^+ be one positive example. H is an inductive conclusion of E^+ and B iff $H \vdash E^+ \rho B^{\perp}$.

Proof. $B \otimes H \vdash E^+$ if $H \vdash E^+ \rho B^{\perp}$. We have

$$\frac{H \vdash E^+ \not P B^\perp}{B, H \vdash E^+} \frac{\frac{B \vdash B}{B, B^\perp \vdash}}{\frac{B, H \vdash E^+}{B \otimes H \vdash E^+}}$$

 $H \vdash E^+ \rho B^{\perp}$ if $B \otimes H \vdash E^+$. We have

$$\frac{H \vdash H}{H \vdash B^{\perp}, B \otimes H} \xrightarrow{B \vdash B} B \otimes H \vdash E^{+}}{\frac{H \vdash E^{+}, B^{\perp}}{H \vdash E^{+}, B^{\perp}}} \blacksquare$$

The process to generate the formula E satisfying $B \otimes H \vdash E$ from $B \otimes H$ is deduction. Therefore, induction is regarded as the inverse of deduction.

Definition 4.1. Inductive Operator $\rightsquigarrow X \rightsquigarrow Y$ iff $Y \vdash X$.

Taking this relation as an operator, we can obtain the inductive conclusions of the background theory B and positive example E^+ by starting with $E^+ \rho B^{\perp}$.

Theorem 4.2. The relation \rightarrow is transitive.

Proof. Since \vdash is transitive, this is obvious.

For any formula $A, A \rightarrow 0$ because of the initial sequent $\Gamma, 0 \vdash \Delta$. Indeed 0 is a hypothesis for any positive example and background theory. However, this is meaningless. Another extreme is $A \rightarrow A$. Thus $E^+ \rho B^{\perp}$ is a hypothesis for the positive example E^+ and background theory B. Basically, we consider removing some redundancy in $E^+ \rho B^{\perp}$ by using operator \sim .

Knowledge such as $X \to X$ does nothing. When one X is given, this knowledge yields only the same one. Since $X \to X$ is equivalent to $X \not P X^{\perp}$, we get the following lemma.

Lemma 4.1. $X P X^+ \rightarrow 1$.

Proof. We have

$$\frac{X \vdash X}{\vdash X \land P X^{\perp}} \quad \blacksquare$$

Therefore, we can rewrite $X P X^{\perp}$ as 1 using the operator \sim . A more general form of the above rule is given as follows.

Theorem 4.3. $(V \otimes X) \land (W \otimes X^{\perp}) \rightsquigarrow V \otimes W$.

Proof. We have

$$\frac{\begin{array}{c}
W \vdash W \quad \overleftarrow{K, X^{\perp}} \\
W \vdash V \quad \overrightarrow{W \vdash X, (W \otimes X^{\perp})} \\
\hline
\frac{V \vdash V \quad W \vdash (V \otimes X), (W \otimes X^{\perp})}{V, W \vdash (V \otimes X) \ P (W \otimes X^{\perp})} \\
\hline
\overline{V \otimes W \vdash (V \otimes X) \ P (W \otimes X^{\perp})} \\
\hline
\end{array}} \bullet$$

Lemma 4.1 is the special case of Theorem 4.3 where both V and W are equal to 1.

Some knowledge may be used many times in order to prove the positive example. The following three theorems deal with the logical operators which allow weakening and contraction.

Theorem 4.4. $?A \rightarrow A$.

Proof. We have

$$\frac{A \vdash A}{A \vdash ?A} \quad \blacksquare$$

Theorem 4.5. $?A \rightarrow \bot$.

Proof. We have

$$\frac{\perp}{\perp + ?A} \quad \blacksquare$$

Theorem 4.6. $?A \rightarrow ?A P ?A$.

Proof. We have

$$\frac{?A \vdash ?A \quad ?A \vdash ?A}{?A \vdash ?A \vdash ?A, ?A}$$

The following theorem allows us to apply the operator \rightarrow to a part of a formula.

Theorem 4.7. If $A \rightarrow A'$, then $A \otimes X \rightarrow A' \otimes X$, $A \land X \rightarrow A' \land X$, $A \land X \rightarrow A' \land X$, $A \oplus X \rightarrow A' \oplus X$, $A \& X \rightarrow A' \& X, X \rightarrow A \rightarrow X \rightarrow A'$, $!A \rightarrow !A'$, and $?A \rightarrow ?A'$.

Proof. We have

<u>A'</u>	$\vdash A X \vdash X$	$A' \vdash A$	$X \vdash X$
$A', X \vdash A \otimes X$		$A' P X \vdash A, X$	
Α' ($\otimes A \vdash A \otimes X$	$A' P A \vdash A$	ΑΡΧ
$A' \vdash A$	$X \vdash X$	$A' \vdash A$	$X \vdash X$
$A' \vdash A \oplus \overline{X}$	$X \vdash A \oplus X$	$\overline{A' \& X \vdash A}$	$\overline{A} \overline{A' \& X \vdash X}$
$\overline{A'\otimes X\vdash A\oplus X}$		A' & 2	$X \vdash A \& X$
$A' \vdash A$	$X \vdash X$	$A' \vdash A$	$A' \vdash A$
$X, X \multimap A' \vdash A$		$!A' \vdash A$	$\overline{A' \vdash ?A}$
$\overline{X \multimap A' \vdash X \multimap A}$		$A' \vdash A$	$\overrightarrow{A'} + \overrightarrow{A}$

Example 4.1. Let D mean that you have just one dollar, and C mean that you have just one pack of cigarettes, and P mean that you have just one pack of fried potatoes; the background theory B is given as follows:

$$B = !(D \multimap C) \otimes !(D \multimap P)$$

This background theory means, "If you spend 1 dollar, you get one pack of cigarettes, and if you spend 1 dollar, you get one pack of fried potatoes, and this knowledge can be used as many times as required." As a positive example, "You have two packs of cigarettes and a pack of fried potatoes" is given. What kind of hypothesis can we get which explains the situation? We have

$$E^{+} = C \otimes C \otimes P$$

The condition of hypothesis H is written as follows:

$$B \otimes H \vdash E^{+}$$

= $H \vdash E^{+} \rho B^{\perp}$
= $H \vdash (C \otimes C \otimes P) \rho (!(D \multimap C) \otimes !(D \multimap P))^{\perp}$
= $H \vdash (C \otimes C \otimes P) \rho ?(D \otimes C^{\perp}) \rho ?(D \otimes P^{\perp})$

Using the inductive operations introduced above, we can rewrite this formula as follows:

$$(C \otimes C \otimes P) \land ?(D \otimes C^{\perp}) \land ?(D \otimes P^{\perp})$$

$$\rightarrow (C \otimes C \otimes P) \land ?(D \otimes C^{\perp}) \land ?(D \otimes C^{\perp}) \land ?(D \otimes P^{\perp})$$

$$\rightarrow (C \otimes C \otimes P) \land (D \otimes C^{\perp}) \land (D \otimes C^{\perp}) \land (D \otimes P^{\perp})$$

$$\rightarrow (D \otimes C \otimes P) \land (D \otimes C^{\perp}) \land (D \otimes P^{\perp})$$

$$\rightarrow (D \otimes D \otimes P) \land (D \otimes P^{\perp})$$

$$\rightarrow D \otimes D \otimes D$$

Thus, we get the hypothesis, "You have just 3 dollars."

To take the predicate logic into account, the inductive operator with unification is introduced. As in the resolution principle, a formula X(t) and its complement $X(x)^{\perp}$ are found with the substitution [t/x].

Theorem 4.8. The following holds:

$$(\exists x(V \otimes X)) \mathrel{\mathcal{P}} (W \otimes X^{\perp}[t/x]) \rightsquigarrow V[t/x] \otimes W$$

Proof. Some abbreviations are employed in following proof:

	$X[t/x] \vdash X[t/x]$
	$W \vdash W \vdash X[t/x], X^{\perp}[t/x]$
$V[t x] \vdash V[t x]$	$W \vdash X[t/x], (W \otimes X^{\perp}[t/x])$
$V[t/x], W \vdash (V$	$(\otimes X)[t/x], (W \otimes X^{\perp}[t/x])$
$V[t/x], W \vdash (V$	$X \otimes X$ [t/x], ($W \otimes X^{\perp}[t/x]$)
$V[t/x], W \vdash (\exists x)$	$x(V \otimes \overline{X})), (W \otimes X^{\perp}[t/X])$
$V[t/x] \otimes W \vdash (\exists$	$\overline{lx(V \otimes X)} \mathcal{P}(W \otimes X^{\perp}[t/x])$

Example 4.2. There is a background theory, "A bird has wings, and this knowledge can be used as many times as required":

 $B = ! \forall x (Bird(x) \multimap Haswings(x))$

When a positive instance "Tweety has wings" is given,

 E^+ = Haswings (Tweety)

we can rewrite the condition of the hypothesis as follows:

$$B \otimes H \vdash E^{+}$$

= $H \vdash E^{+} \rho B^{\perp}$
= $H \vdash \text{Haswings}(\text{Tweety}) \rho (! \forall x(\text{Bird}(x) \multimap \text{Haswings}(x)))^{\perp}$
= $H \vdash \text{Haswings}(\text{Tweety}) \rho ?(\forall x(\text{Bird}(x) \multimap \text{Haswings}(x)))^{\perp}$
= $H \vdash \text{Haswings}(\text{Tweety}) \rho ?\exists x(\text{Bird}(x) \multimap \text{Haswings}(x))^{\perp}$
= $H \vdash \text{Haswings}(\text{Tweety}) \rho ?\exists x(\text{Bird}(x) \otimes \text{Haswings}^{\perp}(x))$

We can use the inductive operation introduced above to unify this, with the substitution [Tweety/x]. Thus, we can rewrite this formula as follows.

Haswings(Tweety) \mathcal{P} ?($\exists x(Bird(x) \otimes Haswings^{\perp}(x))$)

 \rightarrow Haswings(Tweety) $\mathcal{P}(\exists x(\operatorname{Bird}(x) \otimes \operatorname{Haswings}^{\perp}(x)))$

 \rightarrow Bird(Tweety)

Thus, we get the hypothesis, "Tweety is a bird."

5. PLURAL POSITIVE EXAMPLES

When plural positive examples are given, the hypothesis must explain all of them. The next theorem leads to a hypothesis which explains two positive examples at the same time.

Theorem 5.1. When H_1 is an inductive conclusion for positive example E_1^+ and H_2 is an inductive conclusion for positive example E_2^+ , $H_1 \& H_2$ is an inductive conclusion for both E_1^+ and E_2^+ .

Proof. $H_1 \& H_2$ is the greatest lower bound of H_1 and H_2 . The relation \vdash is transitive. Therefore obvious.

The hypothesis may be too complex for using the above theorem. A complicated hypothesis can be reduced by the following rule.

Theorem 5.2. Let A denote a finite number of A's connected by \otimes or & We have $A \rightarrow A$.

Proof. $A \rightarrow !A$:

$$\frac{A \vdash A}{!A \vdash A}$$

 $(!A) \otimes \mathbf{A}' \rightarrow !A \text{ if } \mathbf{A}' \rightarrow !A:$

$$\frac{|A \vdash |A|}{|A| + |A| + |A'|} \frac{|A \vdash A'|}{|A \vdash (|A|) \otimes \mathbf{A}'}$$

 $(!A) \& \mathbf{A}' \rightarrow !A \text{ if } \mathbf{A}' \rightarrow !A:$

$$\frac{|A \vdash |A|}{|A \vdash (|A|) \& \mathbf{A}'}$$

And $|A \rightarrow |A|$; thus Theorem 4.7 leads to $A \rightarrow |A|$.

For example, when E_1^+ suggests $D \otimes D$ and E_2^+ suggests $D \otimes D \otimes D$, the inductive conclusion may be $(D \otimes D) \& (D \otimes D \otimes D)$ or simply !D.

6. CONCLUSION AND FURTHER ISSUES

We have presented a method for induction in linear logic. This method successfully deals with positive examples including predicate linear logic.

The operation which removes redundancy of the form $X \rightarrow X$ is similar to resolution. However, this operation is given as the inverse of deduction.

The inductive operation introducing ! is ambiguous. There are many choices of simplification using ! in general. This rule induces a numerical relation, such as, an even number of D's is represented by $!(D \otimes D)$, one or more D's is represented by $D \otimes !D$, etc.

The treatment of negative examples is another theme. In general, symbolic induction may be regarded as solving a simultaneous inequality on a set of formulas \mathcal{F} using partial order.

Unlike traditional logic, formulas of linear logic can represent quantity. Thus induction in linear logic can allow machine intelligence to find general knowledge expressed in terms of quantitative attributes.

ACKNOWLEDGMENT

I would like to thank Prof. Toshio Nishimura, Faculty of Science, Kanagawa University, for insightful advice and for verifying proofs in this paper.

REFERENCES

- Deville, Y., and Lau, K.-K. (1994). Logic program synthesis, Journal of Logic Programming, 19-20, 321-350.
- Filinski, A. (1992). Linear continuations, in Conference Record of the Nineteenth Annual Symposium on Principles of Programming Languages, ACM Press, pp. 27-38.
- Genesereth, M. R., and Nilsson, N. J. (1987). Logical Foundations of Artificial Intelligence, Morgan Kaufmann Publishers.
- Girard, J.-Y., Lafont, Y., and Regnier, L. (1995). Advances in Linear Logic, Cambridge University Press, Cambridge.
- Muggleton, S., and De Raedt, L. (1994). Inductive logic programming: Theory and methods, Journal of Logic Programming, **19–20**, 629–679.