

Department of Computer Science
CSC 444 (grad) topic

NATURAL DEDUCTION (Pelletier / Kalish & Montague)

1. General Remarks

Resolution caught the imagination of the “automated theorem proving” community when Alan Robinson proposed it in 1965. One result was the rather unjust neglect of “natural” deduction methods – the sorts of methods which people naturally tend to use when proving mathematical theorems (if any such activity can be called natural), and which are outlined in many books on logic (an exemplary text is Kalish & Montague’s *Logic* (Harcourt & Brace 1964)).

It turns out that natural deduction can be just as efficient as resolution-based systems, and sometimes more so. The proofs are often easier to follow – they are structured more like the mathematical proofs we are used to. As a result it’s also somewhat easier to come up with good proof strategies, i.e., we can consult our intuitions about how we do such proofs. A disadvantage is that there are more proof rules to be encoded procedurally, and the structure of proofs is nested rather than uniform as in resolution. However, the simplicity of resolution is a little deceptive, since the conversion to clause form involves multiple rules, and furthermore can lead to exponential “blow-up” in the size of the premises (e.g., consider $((\phi_1 \wedge \psi_1) \vee \dots \vee (\phi_n \wedge \psi_n))$).¹

There are two general characteristics of natural deduction: (1) it typically uses a relatively large number of inference rules (e.g., one or two dozen), including Modus Ponens, Modus Tollens, Universal Instantiation, etc.; and (2) it typically uses “backward reasoning” rules, in addition to ordinary forward reasoning rules. Backward reasoning means that we pursue a goal by transforming it into one or more subgoals, often with systematic use of *assumption making*.

For instance, the goal of proving $(\chi \Rightarrow \psi)$ can be achieved by *assuming* χ , and on that assumption (in addition to all the premises and assumptions already in force) proving ψ . Why is this legitimate? Well, one way of justifying this is via the deduction theorem, which says that if ψ is derivable given χ (plus other premises), then $(\chi \Rightarrow \psi)$ is derivable (from the premises). This is what Genesereth & Nilsson call “proving provability” – i.e., by constructing a proof of ψ , given χ , we have verified that there *is* a proof of $(\chi \Rightarrow \psi)$, using the standard forward reasoning rules (just MP in the proof system presented in G&N) – even though we haven’t actually constructed it; thus we have proved the provability of $(\chi \Rightarrow \psi)$.

Symbolically, we could write this backward rule as

¹However, the blow-up can be avoided by inventing new symbols for pieces of formulas, and adding equivalences. For example, we could rewrite the example here as $(\chi_1 \vee \dots \vee \chi_n)$, $\chi_1 \equiv (\phi_1 \wedge \psi_1)$, ..., $\chi_n \equiv (\phi_n \wedge \psi_n)$, where each χ_i is a new predication over all variables in ϕ_i and ψ_i ; this yields a polynomial-size clause set. The same trick is used in proving the NP-completeness of SAT (satisfiability).

$$\frac{\Delta \vdash (\chi \Rightarrow \psi)}{\Delta, \chi \vdash \psi}.$$

The goal of proving that $(\chi \Rightarrow \psi)$ is derivable (using forward rules only) from premises Δ is written above the line, and the “reduced” problem is shown below the line.

There is an equivalent way of looking at this rule. We might say that our real goal is not to establish “provability” relations based on some set of forward rules and axiom schemas; rather, our goal is to show that certain conclusions *logically follow* from (are *entailed* by) the premises (i.e., all models of the premises are models of the conclusion). In a sense this is a more fundamental way of viewing it, since after all, the purpose of deductive inference is to try to capture whatever follows logically from given premises. From that perspective, the above rule can be rewritten as

$$\frac{\Delta \models (\chi \Rightarrow \psi)}{\Delta, \chi \models \psi};$$

i.e., “Given the goal of showing that $(\chi \Rightarrow \psi)$ logically follows from premises Δ , you need only show that ψ logically follows from those premises, plus assumption χ ”.

2. Pelletier’s System

As a particularly natural system (and one that has been implemented and tested on many “challenge problems” in the literature), consider F.J Pelletier’s THINKER theorem prover.² It is based on the Kalish & Montague proof system, and includes over a dozen forward rules and over a dozen backward rules (and no axioms). The forward rules are Quantifier Negation (QN), Double Negation (DN), Biconditional to Conditional (BC), And Elimination (AE), Modus Ponens (MP), Modus Tollens (MT), Modus Tollendo Ponens (MTP), Existential Instantiation (EI), Universal Instantiation (UI), and Repetition (R):

$$\begin{array}{l} QN : \frac{\neg(\forall\nu)\phi}{(\exists\nu)\neg\phi}, \quad \frac{\neg(\exists\nu)\phi}{(\forall\nu)\neg\phi} \quad DN : \frac{\neg\neg\phi}{\phi} \\ BC : \frac{\chi \Leftrightarrow \psi}{\chi \Rightarrow \psi}, \quad \frac{\chi \Leftrightarrow \psi}{\psi \Rightarrow \chi} \quad AE : \frac{\chi \wedge \psi}{\chi}, \quad \frac{\chi \wedge \psi}{\psi} \\ MP : \frac{\chi, \chi \Rightarrow \psi}{\psi}, \quad MT : \frac{\neg\psi, \chi \Rightarrow \psi}{\neg\chi} \quad MTP : \frac{\neg\chi, \chi \vee \psi}{\psi}, \quad \frac{\psi, \chi \vee \neg\psi}{\chi}, \text{ etc.} \\ EI : \frac{(\exists\nu)\phi(\dots\nu\dots)}{\phi(\dots\nu'\dots)} \quad UI : \frac{(\forall\nu)\phi(\dots\nu\dots)}{\phi(\dots\tau\dots)} \quad R : \frac{\phi}{\phi} \text{ (alternatively, } \frac{\phi, \neg\phi}{\square} \text{)} \end{array}$$

where ν' is a new constant and τ is a term free for ν in ϕ .

²See F.J. Pelletier (1998) “Natural Deduction Theorem Proving in THINKER”, *Studia Logica* 60, pp. 3-43.

The problem reduction rules are as follows. The first three don't involve assumptions, and can be thought of as the "inverses" of forward rules BC, AE, and UI respectively:

$$1. \frac{\Delta \models (\chi \Leftrightarrow \psi)}{\Delta \models (\chi \Rightarrow \psi), \text{ and } \Delta \models (\psi \Rightarrow \chi)} \quad 2. \frac{\Delta \models (\chi \wedge \psi)}{\Delta \models \chi, \text{ and } \Delta \models \psi} \quad 3. \frac{\Delta \models (\forall \nu)\psi}{\Delta \models \psi}$$

The next three rules are for "assuming the antecedent" (to derive a consequent of a conditional) and "assuming the contrary" (to derive a contradiction):

$$4. \frac{\Delta \models (\chi \Rightarrow \psi)}{\Delta, \chi \models \psi} \quad 5. \frac{\Delta \models \neg\phi}{\Delta, \phi \models \square} \quad 6. \frac{\Delta \models \phi}{\Delta, \neg\phi \models \square}$$

The next four rules tell us that to derive a contradiction from a negated formula, it is sufficient to derive the unnegated formula from the negated one (and in the case of disjunction, distributing the negation):

$$7. \frac{\Delta, \neg(\chi \Rightarrow \psi) \models \square}{\Delta, \neg(\chi \Rightarrow \psi) \models (\chi \Rightarrow \psi)} \quad 8. \frac{\Delta, \neg(\chi \Leftrightarrow \psi) \models \square}{\Delta, \neg(\chi \Leftrightarrow \psi) \models (\chi \Leftrightarrow \psi)}$$

$$9. \frac{\Delta, \neg(\chi \wedge \psi) \models \square}{\Delta, \neg(\chi \wedge \psi) \models (\chi \wedge \psi)} \quad 10. \frac{\Delta, \neg(\chi \vee \psi) \models \square}{\Delta, \neg(\chi \vee \psi) \models \chi, \text{ or } \neg(\chi \vee \psi) \models \psi}$$

(Here I'm not sure why these aren't collapsed into $\frac{\Delta, \neg\phi \models \square}{\Delta, \neg\phi \models \phi}$ and $\frac{\Delta \models (\chi \vee \psi)}{\Delta, \neg\chi \models \psi, \text{ or } \Delta, \neg\psi \models \chi}$. It also seems to me it would help to have rules for directly distributing negation over conjunctions, disjunctions, and conditionals.)

The next four rules formally are directed at the goal of proving ϕ , but they are best thought of as rules for strengthening the premises, while disregarding the form of the current goal formula. The idea is that a "stalled" proof attempt may be able to proceed towards the goal again once some additional consequences of the premises have been worked out. (Actually the theorem prover is also able to switch into forward inference mode when the backward proof attempt bogs down.)

$$11. \frac{\Delta, (\chi \Rightarrow \psi) \models \phi}{\Delta, (\chi \Rightarrow \psi) \models \chi, \text{ and } \Delta, (\chi \Rightarrow \psi), \chi \models \phi} \quad 12. \frac{\Delta, (\chi \Rightarrow \psi) \models \phi}{\Delta, (\chi \Rightarrow \psi) \models \neg\psi, \text{ and } \Delta, (\chi \Rightarrow \psi), \neg\psi \models \phi}$$

$$13. \frac{\Delta, (\chi \vee \psi) \models \phi}{\Delta, (\chi \vee \psi) \models \neg\chi, \text{ and } \Delta, (\chi \vee \psi), \neg\chi \models \phi} \quad 14. \frac{\Delta, (\chi \Rightarrow \psi) \models \phi}{\Delta, (\chi \vee \psi) \models \neg\psi, \text{ and } \Delta, (\chi \vee \psi), \neg\psi \models \phi}$$

Finally, there are at least the first two of the following "termination rules" which recognize that the desired result could be proved from the premises in one forward inference step, and so obviously follows from them. Many more rules which "jump to obvious conclusions" could be added:

$$15. \frac{\Delta, \phi, \neg\phi \models \square}{\text{True}} \quad 16. \frac{\Delta, \phi \models \phi}{\text{True}}$$

$$17. \frac{\Delta, \chi, (\chi \Rightarrow \psi) \models \psi}{True} \quad 18. \frac{\Delta, \psi, \neg\psi \models \chi}{True} \quad 19. \frac{\Delta, (\chi \Leftrightarrow \psi) \models (\chi \Rightarrow \psi)}{True}, \text{ etc.}$$

One point I'm not sure about is why MP, MT and MTP aren't used in "inverted form" as backward rules, e.g., $\frac{\Delta \models \psi}{\Delta \models (\phi \Rightarrow \psi), \text{ and } \Delta \models \phi}$ (where we would select an appropriate conditional by matching the goal formula ψ to consequents of conditionals among the premises).

3. Proofs in Pelletier's System

In a proof, we write the desired conclusion ϕ as "Show ϕ ". When we apply a problem reduction rule, we write down the assumptions (if any) that it introduces and generate new "Show"s, which we indent relative to the superordinate "Show". When a subproof is complete, we "star" the corresponding "Show" (putting a "*" beside it) and "box" the corresponding subproof, by drawing a vertical line downward from the star to the end of the subproof. Formulas may also be added by forward inference rules.

An important point to notice about the problem reduction rules is that they are all "monotone in the premises", i.e., they only *add* assumptions, never *revoke* them. Consequently, in any step we can use all assumptions that have not yet been "boxed". More generally, a proof step may use any prior formulas, including starred "Show"-formulas, that can be reached by moving monotonically to less indented prior parts of the proof, or staying at the same level of indentation. Other details are best understood with the help of some examples:

Example 3.1

Premises: none; *to be proved:* $((\forall x)(Px \Rightarrow Qx) \Rightarrow ((\forall x)Px \Rightarrow (\forall x)Qx))$.

Note: we are dropping some brackets for brevity.

1.	*	Show	$((\forall x)(Px \Rightarrow Qx) \Rightarrow ((\forall x)Px \Rightarrow (\forall x)Qx))$.		
2.			$(\forall x)(Px \Rightarrow Qx)$	Assume	rule 4
3.		*	Show	$(\forall x)Px \Rightarrow (\forall x)Qx$	
4.			$(\forall x)Px$	Assume	rule 4
5.			*	Show	$(\forall x)Qx$
6.			*	Show	Qx
7.				$(Px \Rightarrow Qx)$	2, UI
8.				Px	4, UI
9.				Qx	7, 8, MP

Example 3.2

Premises: $(\exists x)(Q \Rightarrow Px)$, $(\exists x)(Px \Rightarrow Q)$

To be proved: $(\exists x)(Px \Leftrightarrow Q)$

NB: “Show \square ” is conventionally omitted.

1.	* Show $(\exists x)(Px \Leftrightarrow Q)$		
2.	$\neg(\exists x)(Px \Leftrightarrow Q)$	Assume	rule 6 (“Show \square ” omitted)
3.	$(\forall x)\neg(Px \Leftrightarrow Q)$	2, QN	
4.	$(\exists x)(Q \Rightarrow Px)$	Prem	Forward inferences are made as long as there are no applicable problem reduction rules
5.	$(Q \Rightarrow Py)$	4, EI	
6.	$(\exists x)(Px \Rightarrow Q)$	Prem	
7.	$(Pz \Rightarrow Q)$	6, EI	
8.	$\neg(Py \Leftrightarrow Q)$	3, UI	NB: y was introduced in line 5
9.	* Show $(Py \Leftrightarrow Q)$		rule 8
10.	* Show $(Py \Rightarrow Q)$		rule 1, first half
11.	Py	Assume	rule 4
12.	* Show Q		
13.	$\neg Q$	Assume	rule 6 (“Show \square ” omitted)
14.	$\neg(Pz \Leftrightarrow Q)$	3, UI	NB: z was introduced in line 5
15.	* Show $(Pz \Leftrightarrow Q)$		rule 8
16.	* Show $(Q \Rightarrow Pz)$		rule 1, second half
17.	Q	Assume	rule 4 (implicit “Show Pz ”)
18.	$\neg Q$	13, R	Alternatively, could infer: \square 17, 13
19.	$(Pz \Leftrightarrow Q)$	7, 16, CB	
20.	$(Py \Leftrightarrow Q)$	5, 10, CB	

Note that when using conjunctive problem reduction rules such as rule 1 (see lines 10, 16, 19, and 20 above) we do not introduce the second “Show” right away, but only after the first “Show” has been completed. If upon completion of the first “Show”, it turns out that the second part has already been proved, it won’t be generated as a goal again. In either case, the *forward* inference rule corresponding to the conjunctive deduction rule is used as justification of the desired conclusion from the two parts (as in 19, 20 above). (This is handled by a special SIMPLEPROOF procedure in THINKER.)