[Tar56] A. Tarski. *Logic, Semantics, Metamathematics.* Oxford University Press, 1956.

Press, 1989. Also IRST-Technical Report 8902-04 and DAI Research Paper 375, University of Edinburgh.

[GS91] F. Giunchiglia and L. Serafini. Multilanguage first order theories of propositional attitudes. In *Proceedings 3rd Scandinavian Conference on Artificial Intelligence*, pages 228–240, Roskilde University, Denmark, 1991. IOS Press. Also IRST-Technical Report 9001-02, IRST, Trento, Italy.

[GS92] F. Giunchiglia and L. Serafini. Hierarchical meta-logics for beliefs and provability: how we can do without modal logics. In *Proc. Canadian Conference of Artificial Intelligence AI '92*, Vancouver, Canada, 1992. Also IRST-Technical Report 9112-03, IRST, Trento, Italy. Long version to appear in: Artificial Intelligence Journal.

[GS94] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: how we can do without modal logics). *Artificial Intelligence*, 65:1–42, 1994.

[GSGF93] F. Giunchiglia, L. Serafini, E. Giunchiglia, and M. Frixione. Non-Omniscient Belief as Context-Based Reasoning. In *Proc. of the 13th International Joint Conference on Artificial Intelligence*, pages 548–554, Chambery, France, 1993. Also IRST-Technical Report 9206-03, IRST, Trento, Italy.

[GSS92] F. Giunchiglia, L. Serafini, and A. Simpson. Hierarchical meta-logics: intuitions, proof theory and semantics. In *Proc. META-92, Workshop on Metaprogramming in Logic*, number 649 in LNCS, pages 235–249, Uppsala, Sweden, 1992. Springer Verlag. Also IRST-Technical Report 9101-05, IRST, Trento, Italy.

[GT91] F. Giunchiglia and P. Traverso. Reflective reasoning with and between a declarative metatheory and the implementation code. In *Proc. of the 12th International Joint Conference on Artificial Intelligence*, pages 111–117, Sydney, 1991. Also IRST-Technical Report 9012-03, IRST, Trento, Italy.

[GT92] F. Giunchiglia and P. Traverso. A Metatheory of a Mechanized Object Theory. Technical Report 9211-24, IRST, Trento, Italy, 1992. Submitted for publication to: Journal of Artificial Intelligence.

[Pra65] D. Prawitz. *Natural Deduction - A proof theoretical study.* Almquist and Wiksell, Stockholm, 1965.

[Sho67] J.R. Shoenfield. *Mathematical logic.* Addison Wesley, London, 1967.

[Smo85] C. Smorynski. *Self-Reference and Modal Logic.* Springer-Verlag, Berlin, 1985.

[Sol70] R. Solovay. Provability Interpretations of Modal Logic. *Israel Journal of Mathematics*, 25:287–304, 1970.

[Tar36] A. Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1:261–405, 1936. English translation in [Tar56].

Artificial Intelligence. As argued in [GS93], MR systems closely resemble the current (implementational) practice in meta-theoretic theorem proving and in the representation of propositional attitudes. In this paper we have studied the proof theoretic properties of MR systems and also their relationship with normal modal logics.

Acknowledgements

The research described in this paper owes a lot to the openess and sharing of ideas which exists in the Mechanized Reasoning groups in Trento and in Genoa. Alex Simpson and Vanni Criscuolo have provided useful feedback on many of the topics described in this paper. The discussions with Alessandro Cimatti, Enrico Giunchiglia, Paolo Traverso, have provided useful insights. We also thank Alessandro Agostini for his reading of earlier versions of this paper and Toby Walsh for his comments about style.

References

[Boo79] G. Boolos. *The Unprovability of Consistency, an essay in modal logic.* Cambridge University Press, 1979.

[BS84] R. Bull and K. Segerberg. Basic Modal Logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume III, pages 1–88. D. Reidel Publishing Company, 1984.

[Che80] B. F. Chellas. *Modal Logic – an Introduction.* Cambridge University Press, 1980.

[FH88] R. Fagin and J.Y. Halpern. Belief, awareness, and limited reasoning. *Artif. Intell.*, 34:39–76, 1988.

[Gab90] D. Gabbay. Labeled Deductive Systems. Technical Report CIS-Bericht-90-22, Universität München – Centrum für Informations und Sprachverarbeitung, 1990.

[Giu91] F. Giunchiglia. Multilanguage systems. In *Proceedings of AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning*, 1991. Also IRST-Technical Report no. 9011-17.

[Giu93] F. Giunchiglia. Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine*, XVI:345–364, 1993. Short version in Proceedings IJCAI'93 Workshop on Using Knowledge in its Context, Chambery, France, 1993, pp. 39–49. Also IRST-Technical Report 9211-20, IRST, Trento, Italy.

[GS89] F. Giunchiglia and A. Smaill. Reflection in constructive and non-constructive automated reasoning. In H. Abramson and M. H. Rogers, editors, *Proc. Workshop on Meta-Programming in Logic Programming*, pages 123–145. MIT

$\mathrm{MKS}_1 \ldots \mathrm{S}_n$	$\mathrm{MK}\mathcal{S}_1 \ldots \mathcal{S}_n$
$\langle T(``A") \supset \neg T(``\neg A"),\ i+1\rangle$	$\dfrac{\dfrac{\dfrac{\langle T(``A"),\ i+1\rangle}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle \neg T(``\neg A"),\ i+1\rangle}\ \mathcal{D}_i}{\langle T(``A") \supset \neg T(``\neg A"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$
$\langle A \supset T(``\neg T(``\neg A")"),\ i+2\rangle$	$\dfrac{\dfrac{\dfrac{\langle A,\ i+2\rangle}{\langle \neg T(``\neg A"),\ i+1\rangle}\ \mathcal{B}_i}{\langle T(``\neg T(``\neg A")"),\ i+2\rangle}\ \mathcal{R}_{up.i+1}}{\langle A \supset T(``\neg T(``\neg A")"),\ i+2\rangle}\ \supset\mathrm{I}_{i+2}$
$\langle T(``A") \supset A,\ i+1\rangle$	$\dfrac{\dfrac{\dfrac{\langle T(``A"),\ i+1\rangle}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A,\ i+1\rangle}\ \mathcal{T}_i}{\langle T(``A") \supset A,\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$
$\langle T(``A") \supset T(``T(``A")"),\ i+2\rangle$	$\dfrac{\dfrac{\dfrac{\langle T(``A"),\ i+2\rangle}{\langle T(``A"),\ i+1\rangle}\ \mathcal{S}4_{i+1}}{\langle T(``T(``A")"),\ i+2\rangle}\ \mathcal{R}_{up.i+1}}{\langle T(``A") \supset T(``T(``A")"),\ i+2\rangle}\ \supset\mathrm{I}_{i+2}$
$\langle \neg T(``A") \supset T(``\neg T(``A")"),\ i+2\rangle$	$\dfrac{\dfrac{\dfrac{\langle \neg T(``A"),\ i+2\rangle}{\langle \neg T(``A"),\ i+1\rangle}\ \mathcal{S}4_{i+1}}{\langle T(``\neg T(``A")"),\ i+2\rangle}\ \mathcal{R}_{up.i+1}}{\langle \neg T(``A") \supset T(``\neg T(``A")"),\ i+2\rangle}\ \supset\mathrm{I}_{i+2}$
$\langle T(``T(``A") \supset A") \supset T(``A"),\ i+1\rangle$	$\dfrac{\dfrac{\dfrac{\dfrac{\langle T(``T(``A") \supset A"),\ i+1\rangle}{\langle T(``A") \supset A,\ i\rangle}\ \mathcal{R}_{dn.i} \quad \langle T(``A"),\ i\rangle}{\langle A,\ i\rangle}\ \supset\mathrm{E}_i}{\langle T(``A"),\ i+1\rangle}\ \mathcal{G}_i}{\langle T(``T(``A") \supset A") \supset T(``A"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$

Q.E.D.

Corollary 7.2 *If $KS_1 \ldots S_n$ is a modal system respecting* (47) *(i.e. it is one among K, KD, KT, K4, KDB, KD4, KTB, KT4, KT5 and KG) then, $\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A^*$ if and only if there exists an index i such that $\vdash_{\mathrm{MK}\mathcal{S}_1 \ldots \mathcal{S}_n} \langle A,\ i\rangle$.*

Corollary 7.3 *For any normal modal logic $KS_1 \ldots S_n$, $\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A^{**}$ if and only if $\vdash_{\mathrm{MBK}\mathcal{S}_1 \ldots \mathcal{S}_n}$ $\langle A,\ 0\rangle$*

8 Conclusion

The work presented in this paper is part of a much larger project whose goal is to provide logical and philosophical foundations to the implementational work done in

$\mathcal{D}_i^B$ and $\mathcal{T}_i^B$ have the same restriction as $\mathcal{R}_{up.i}$. $\mathcal{G}_i^B$ is applicable only if $\langle A,\ i+1\rangle$ does not depend on any assumption with index $i+1$ which differs from $\langle B(\text{“}A\text{”}),\ i+1\rangle$.

Definition 7.7 *Let $\mathcal{S}_1, \ldots, \mathcal{S}_n$ be any bridge rules chosen among $\{\mathcal{D}, \mathcal{B}, \mathcal{T}, \mathit{4}, \mathit{5}, \mathcal{G}\}$. Then the ML system MK$\mathcal{S}_1 \ldots \mathcal{S}_n$ is obtained by adding, for each $i \in \omega$, $\mathcal{S}_{1i} \ldots \mathcal{S}_{ni}$ to the deduction machinery of MK. Furthermore, the ML system MBK$\mathcal{S}_1 \ldots \mathcal{S}_n$ is obtained by adding, for each $i \in \omega$, $\mathcal{S}_{1i}^B \ldots \mathcal{S}_{ni}^B$ to the deduction machinery of MBK.*

Theorem 7.9 *If A be an L_i-wff, then $\vdash_{\mathrm{MKS}_1 \ldots \mathrm{S}_n} \langle A,\ i\rangle$ if and only if $\vdash_{\mathrm{MK}\mathcal{S}_1 \ldots \mathcal{S}_n} \langle A,\ i\rangle$. If A is an $L(B)$-wff, then $\vdash_{\mathrm{MBKS}_1 \ldots \mathrm{S}_n} \langle A,\ i\rangle$ if and only if $\vdash_{\mathrm{MBK}\mathcal{S}_1 \ldots \mathcal{S}_n} \langle A,\ i\rangle$.*

Proof The first table below shows how each bridge rule $\mathcal{S}_i$ is a derived inference rules of MKS$_1 \ldots$S$_n$. The second shows how the each axiom schema S$_i$ is a theorem of MK$\mathcal{S}_1 \ldots \mathcal{S}_n$. Similar tables can be built for MBK.

MK$\mathcal{S}_1 \ldots \mathcal{S}_n$	MKS$_1 \ldots$S$_n$
$\dfrac{\langle A,\ i\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\ \mathcal{D}_i$	$\dfrac{\dfrac{\langle A,\ i\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{R}_{up.i} \quad \langle T(\text{“}A\text{”}) \supset \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\ \supset\mathrm{E}_{i+1}$
$\dfrac{\langle A,\ i+1\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i\rangle}\ \mathcal{B}_i$	$\dfrac{\dfrac{\langle A,\ i+2\rangle \quad \langle A \supset T(\text{“}\neg T(\text{“}\neg A\text{”})\text{”}),\ i+2\rangle}{\langle T(\text{“}\neg T(\text{“}\neg A\text{”})\text{”}),\ i+1\rangle}\ \supset\mathrm{E}_{i+2}}{\langle \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\ \mathcal{R}_{dn.i+1}$
$\dfrac{\langle A,\ i\rangle}{\langle A,\ i+1\rangle}\ \mathcal{T}_i$	$\dfrac{\dfrac{\langle A,\ i\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{R}_{up.i} \quad \langle T(\text{“}A\text{”}) \supset A,\ i+1\rangle}{\langle A,\ i+1\rangle}\ \supset\mathrm{E}_{i+1}$
$\dfrac{\langle T(\text{“}A\text{”}),\ i+1\rangle}{\langle T(\text{“}A\text{”}),\ i\rangle}\ \mathcal{S}4_i$	$\dfrac{\dfrac{\langle T(\text{“}A\text{”}),\ i+1\rangle \quad \langle T(\text{“}A\text{”}) \supset T(\text{“}T(\text{“}A\text{”})\text{”}),\ i+1\rangle}{\langle T(\text{“}T(\text{“}A\text{”})\text{”}),\ i+1\rangle}\ \supset\mathrm{E}_{i+1}}{\langle T(\text{“}A\text{”}),\ i\rangle}\ \mathcal{R}_{dn.i}$
$\dfrac{\langle \neg T(\text{“}A\text{”}),\ i+1\rangle}{\langle \neg T(\text{“}A\text{”}),\ i\rangle}\ \mathcal{S}5_i$	$\dfrac{\dfrac{\langle \neg T(\text{“}A\text{”}),\ i+1\rangle \quad \langle \neg T(\text{“}A\text{”}) \supset T(\text{“}\neg T(\text{“}A\text{”})\text{”}),\ i+1\rangle}{\langle T(\text{“}\neg T(\text{“}A\text{”})\text{”}),\ i+1\rangle}\ \supset\mathrm{E}_{i+1}}{\langle \neg T(\text{“}A\text{”}),\ i\rangle}\ \mathcal{R}_{dn.i}$
$\begin{array}{c}\langle T(\text{“}A\text{”}),\ i\rangle \\ \vdots \\ \dfrac{\langle A,\ i\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{G}_i\end{array}$	$\dfrac{\dfrac{\dfrac{\begin{array}{c}[\langle T(\text{“}A\text{”}),\ i\rangle] \\ \vdots \\ \langle A,\ i\rangle\end{array}}{\langle T(\text{“}A\text{”}) \supset A,\ i\rangle}\ \supset\mathrm{I}_i}{\langle T(\text{“}T(\text{“}A\text{”}) \supset A\text{”}),\ i+1\rangle}\ \mathcal{R}_{up.i} \quad \langle T(\text{“}T(\text{“}A\text{”}) \supset A\text{”}) \supset T(\text{“}A\text{”}),\ i+1\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \supset\mathrm{E}_{i+1}$

The restriction on $\mathcal{R}_{up.i}$ determines the restrictions on the corresponding bridge rules.

(b') *Homogeneization of the premisses and conclusion:* In an MR system the same wff at different levels has different meanings. Thus a statement involving wffs at multiple levels, cannot be directly translated into an equivalent statement about the corresponding modal logic. In order to translate

$$\langle \bullet_i(\text{“}A\text{”}),\ j\rangle \vdash_{\text{MS}} \langle A,\ i\rangle \tag{65}$$

in modal logic we have to “lift up” the conclusion to the same language of the assumption, obtaining

$$\langle \bullet_i(\text{“}A\text{”}),\ j\rangle \vdash_{\text{MS}} \langle \bullet_i(\text{“}A\text{”}),\ j\rangle \tag{66}$$

But in this process the information on the interaction between the i-theory and the j- theory stated by (65) is lost. This can be summed up by saying that, even if it is possible to prove an equivalence result between provability in MR systems and provability in modal logics the same cannot be done for derivability. The derivations with assumptions and conclusions in different languages have no counterpart in modal logiccs.

7.4 Using Bridge Rules to Obtain Normal Modal Logics

ML systems which are equivalent to some (all) modal logics can be obtained by adding bridge rules instead of axioms. Let us define the following bridge rules:

$$\frac{\langle A,\ i\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\ \mathcal{D}_i \qquad \frac{\langle A,\ i+1\rangle}{\langle \neg B(\text{“}\neg A\text{”}),\ i\rangle}\ \mathcal{D}_i^B$$

$$\frac{\langle A,\ i+1\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i\rangle}\ \mathcal{B}_i \qquad \frac{\langle A,\ i\rangle}{\langle \neg B(\text{“}\neg A\text{”}),\ i+1\rangle}\ \mathcal{B}_i^B$$

$$\frac{\langle A,\ i\rangle}{\langle A,\ i+1\rangle}\ \mathcal{T}_i \qquad \frac{\langle A,\ i+1\rangle}{\langle A,\ i\rangle}\ \mathcal{T}_i^B$$

$$\frac{\langle T(\text{“}A\text{”}),\ i+1\rangle}{\langle T(\text{“}A\text{”}),\ i\rangle}\ 4_i \qquad \frac{\langle B(\text{“}A\text{”}),\ i\rangle}{\langle B(\text{“}A\text{”}),\ i+1\rangle}\ 4_i^B$$

$$\frac{\langle \neg T(\text{“}A\text{”}),\ i+1\rangle}{\langle \neg T(\text{“}A\text{”}),\ i\rangle}\ 5_i \qquad \frac{\langle \neg B(\text{“}A\text{”}),\ i\rangle}{\langle \neg B(\text{“}A\text{”}),\ i+1\rangle}\ 5_i^B$$

$$\frac{\begin{array}{c}[\langle T(\text{“}A\text{”}),\ i\rangle]\\ \langle A,\ i\rangle\end{array}}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{G}_i \qquad \frac{\begin{array}{c}[\langle B(\text{“}A\text{”}),\ i+1\rangle]\\ \langle A,\ i+1\rangle\end{array}}{\langle B(\text{“}A\text{”}),\ i\rangle}\ \mathcal{G}_i^B$$

Restrictions: $\mathcal{D}_i$ and $\mathcal{T}_i$ have the same restriction as $\mathcal{R}_{up.i}$. 4_i and 5_i are applicable only if their consequences are L_i-wffs. $\mathcal{G}_i$ is applicable only if $\langle A,\ i\rangle$ does not depend of any assumption with index i which differs from $\langle T(\text{“}A\text{”}),\ i\rangle$.

Example 7.3 Consider the MR system MBKB, obtained by adding the axioms:

$$\langle A \supset B(\text{“}\neg B(\text{“}A\text{”})\text{”}),\, i\rangle$$

to each theory of MBK. Consider the deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle B(\text{“}\bot\text{”}),\, 1\rangle}{\langle \bot,\, 2\rangle}\,\mathcal{R}_{dn.2}}{\langle \neg\neg\bot,\, 2\rangle}\,\bot_2}{\langle B(\text{“}\neg\neg\bot\text{”}),\, 1\rangle}\,\mathcal{R}_{up.2} \qquad \dfrac{\dfrac{\dfrac{\langle \bot,\, 0\rangle}{\langle \neg\bot,\, 0\rangle}\,{\supset}\mathrm{I}_0 \quad \langle \neg\bot \supset B(\text{“}\neg B(\text{“}\neg\neg\bot\text{”})\text{”}),\, 0\rangle}{\langle B(\text{“}\neg B(\text{“}\neg\neg\bot\text{”})\text{”}),\, 0\rangle}\,{\supset}\mathrm{E}_0}{\langle \neg B(\text{“}\neg\neg\bot\text{”}),\, 1\rangle}\,\mathcal{R}_{dn.1}}{\langle \bot,\, 1\rangle}\,{\supset}\mathrm{E}_1}{\langle \neg B(\text{“}\bot\text{”}),\, 1\rangle}\,\bot_1 \tag{64}$$

(64) is the translation (via $(.)^{\mathrm{MK2MBK}}$) of deduction (57) on page 63. It is a proof of $\langle \neg B(\text{“}\bot\text{”}),\, 1\rangle$. Again the translation of $\neg B(\text{“}\bot\text{”})$ is not a theorem of KB. This is due to the fact that the 1-theory has a theory on the top of it, which is the 0-theory. On the other hand the 0-theory of MBKB is equivalent to KB because, it is the higher theory.

We end this subsection with a general remark on the equivalence between MR systems and modal logics.

Remark 7.5 The two main features of MR systems which are not shared by modal logics can be summarized as follows:

(a) MR systems allow multiple theories. As a consequence the same wff in different theories has different meanings.

(b) MR systems allow derivations that span different theories. In particular assumptions in one theory may influence deductions in other theories.

Multiple theories and derivations with assumptions and conclusions in distinct languages have an epistemological significance.

Although provability in MR systems is embeddable into provability in normal modal logics, MR systems are more expressive than modal logics. Indeed the structure $\langle I, \prec\rangle$ of the languages allows for a more fine grained formalization, which is lost when collapsing the languages into a unique amalgamated modal language. The loss of expressivity is due to two main facts directly connected with points (a) and (b) above.

(a') *Elimination of the structural information:* In a equivalence statement such as that of theorem 7.1 the existential quantification over the set I makes it impossible in S to distinguish the case in which $\langle A,\, i\rangle$ is provable in MS from the case in which $\langle A,\, j\rangle$ is provable in MS, for some $i \neq j$. Analogously in the equivalence statement of the form; “$\vdash_{\mathrm{S}} A$ if and only if forall $i \in I \vdash_{\mathrm{MS}} \langle A,\, i\rangle$”, the universal quantification over I, makes it impossible to distinguish the case in which $\langle A,\, i\rangle$ is not provable in MS from the case in which $\langle A,\, j\rangle$ is *not* provable in MS, for some $i \neq j$.

Theorem 7.7 (Axioms vs. Assumptions) *Let A be an L_i-wff and B an L_j-wff, then:*

(i) *If $i \geq j$: $\vdash_{\text{MBK}+\langle B,\, j\rangle} \langle A,\, i\rangle$ if and only if $\langle B,\, j\rangle \vdash_{\text{MK}} \langle A,\, i\rangle$;*

(ii) *If $i < j$: $\vdash_{\text{MBK}+\langle B,\, j\rangle} \langle A,\, i\rangle$ if and only if $\langle T^{i-j}(\text{“}B\text{”}),\, i\rangle \vdash_{\text{MK}} \langle A,\, i\rangle$.*

Proof The proof is analogous to that of the corresponding theorem of MBK, theorem 6.1 on page 55. Q.E.D.

Let $(.)^{**}$ and $(.)^{++}$ be two mappings from each language L_i of MBK to $L(\Box)$. They behave the same as $(.)^{*}$ and $(.)^{+}$, except that but $\Box$ is mapped into B.

Theorem 7.8 *For any normal modal logic $KS_1 \dots S_n$, $\vdash_{\text{KS}_1 \dots \text{S}_n} A^{**}$ if and only if $\vdash_{\text{MBKS}_1 \dots \text{S}_n}$ $\langle A,\, 0\rangle$.*

Proof Suppose that $\vdash_{\text{MBKS}_1 \dots \text{S}_n} \langle A,\, 0\rangle$; by compactness let Ω be a finite subset of the characteristic axioms of $\text{MBKS}_1 \dots \text{S}_n$ such that:

$$\vdash_{\text{MBK}+\Omega} \langle A,\, 0\rangle \tag{58}$$

By the MBK version of theorem 7.7

$$\Omega' \vdash_{\text{MBK}} \langle A,\, 0\rangle \tag{59}$$

where $\Omega' = \{\langle B^j(\text{“}C\text{”}),\, 0\rangle : \langle C,\, j\rangle \in \Omega\}$. Notice that each element of Ω' has index equal to 0. By theorem 7.6, there exists an i such that:

$$\Omega'^{B}_{T} \vdash_{\text{MBK}} \langle A^{B}_{T},\, i\rangle \tag{60}$$

By the equivalence between K and MK, (theorem 7.1 on page 60) we have that:

$$\left(\Omega'^{B}_{T}\right)^{+} \vdash_{\text{K}} \left(A^{B}_{T}\right)^{+} \tag{61}$$

which implies that:

$$\vdash_{\text{K}S_1 \dots \text{S}_n} \left(A^{B}_{T}\right)^{+} \tag{62}$$

and for the definition of $(.)^{++}$

$$\vdash_{\text{K}S_1 \dots \text{S}_n} A^{++} \tag{63}$$

Viceversa: Notice that the 0-theory contains the translation under $(.)^{**}$ of the axioms of $\text{KS}_1 \dots \text{S}_n$ and is closed under the derivability conditions of $\vdash_{\text{KS}_1 \dots \text{S}_n}$. Q.E.D.

Remark 7.4 Theorem 7.8 does not hold, in general, for the theories distinct from the 0-theory. Consider the following example.

Let us define the operator $(.)^{MK2MBK}$, that maps MK deductions into MBK deductions. Intuitively $(.)^{MK2MBK}$ substitutes the predicate T with B, reverses the indexes of the occurrences of a deduction. The dual operation is performed by its inverse $(.)^{MBK2MK}$.

Definition 7.5 *For any deduction Π in MK with maximum index i_0, the formula tree $(\Pi)^{MK2MBK}$ in MBK is obtained by replacing every occurrence $\langle A, i\rangle$ in Π with $\langle A^T_B, i_0 \perp i\rangle$. For any deduction Π in MBK with maximum index i_0, the formula tree $(\Pi)^{MBK2MK}$ is obtained by replacing every occurrence $\langle A, i\rangle$ in Π with $\langle A^B_T, i_0 \perp i\rangle$.*

Example 7.2 An example on how $(.)^{MK2MBK}$ works, is as follows:

$$\left(\dfrac{\dfrac{\dfrac{\dfrac{\langle T(\text{“}T(\text{“}p\text{”})\text{”}),\, 2\rangle}{\langle T(\text{“}p\text{”}),\, 1\rangle}\,\mathcal{R}_{dn.1}}{\langle p,\, 0\rangle}\,\mathcal{R}_{dn.0} \quad \dfrac{\langle T(\text{“}q\text{”}),\, 1\rangle}{\langle q,\, 0\rangle}\,\mathcal{R}_{dn.1}}{\langle p\wedge q,\, 0\rangle}\,\wedge\mathrm{I}_0}{\langle T(\text{“}p\wedge q\text{”}),\, 1\rangle}\,\mathcal{R}_{up.0}\right)^{MK2MBK} = \dfrac{\dfrac{\dfrac{\dfrac{\langle B(\text{“}B(\text{“}p\text{”})\text{”}),\, 0\rangle}{\langle B(\text{“}p\text{”}),\, 1\rangle}\,\mathcal{R}_{dn.0}}{\langle p,\, 2\rangle}\,\mathcal{R}_{dn.1} \quad \dfrac{\langle B(\text{“}q\text{”}),\, 1\rangle}{\langle q,\, 2\rangle}\,\mathcal{R}_{dn.1}}{\langle p\wedge q,\, 2\rangle}\,\wedge\mathrm{I}_2}{\langle B(\text{“}p\wedge q\text{”}),\, 1\rangle}\,\mathcal{R}_{up.1}$$

Lemma 7.3 *If Π is a deduction in MK of $\langle A, i\rangle$ depending on Γ and i_0 is the greatest index of the occurrences of Π, then $(\Pi)^{MK2MBK}$ is a deduction in MBK of $\langle A^T_B, i_0 \perp i\rangle$ depending on Γ'; where $\Gamma' = \{\langle B^T_B, i_0 \perp j\rangle : \langle B, j\rangle \in \Gamma\}$.*
If Π is a deduction in MBK of $\langle A, i\rangle$ depending on Γ and i_0 is the greatest index of the occurrences of Π, then $(\Pi)^{MBK2MK}$ is a deduction in MK of $\langle A^B_T, i_0 \perp i\rangle$ depending on Γ'; where $\Gamma' = \{\langle B^B_T, i_0 \perp j\rangle : \langle B, j\rangle \in \Gamma\}$.

Proof The proof is a straightforward induction on the complexity of Π. Q.E.D.

Theorem 7.6 (MK and MBK) *For any set $G \cup \{A\}$ of L_i-wffs, $\langle G, i\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$ if and only if $\langle G^T_B, 0\rangle \vdash_{\mathrm{MBK}} \langle A^T_B, 0\rangle$.*

Proof Let Π be a weak normal deduction of $\langle A, i\rangle$ depending on $\langle G', i\rangle \subseteq \langle G, i\rangle$. For the sublevel property (theorem 4.3 on page 31) the maximum index of Π is i. By lemma 7.3, Π^{MK2MBK} is a deduction of $\langle A^T_B, 0\rangle$ depending of $\langle {G'}^T_B, 0\rangle$. We conclude that $\langle G^T_B, 0\rangle \vdash_{\mathrm{MBK}} \langle A^T_B, 0\rangle$. For the opposite direction we proceed in the same way applying $(.)^{MBK2MK}$ instead of $(.)^{MK2MBK}$. Q.E.D.

Definition 7.6 ($\mathbf{MBKS_1 \ldots S_n}$) *For each modal system $KS_1 \ldots S_n$, $MBKS_1 \ldots S_n$ is an MR system based on MBK such that:*

$$MBKS_1 \ldots S_n = MBK + \langle S^+_1, 1\rangle + \ldots + \langle S^+_n, 1\rangle + \langle S^+_1, 2\rangle + \ldots + \langle S^+_n, 2\rangle + \ldots$$

where for each $1 \leq k \leq n$, $\langle S_k, i\rangle$ is the translation of the schema S_k, restricted to the language L_i.

In order to prove the equivalence theorem between MR systems based on MBK and modal logics, we adopt the same methodology as in the case of MK. We first state a theorem which allows us to translate axioms of $\mathrm{MBKS_1 \ldots S_n}$ into assumpions on MBK.

Remark 7.1 As property (47) does not hold for all normal modal systems. theorem 7.4 is not applicable to all normal modal systems. Let us understand the role of property (47) in the proof of theorem 7.4.

Property (47) is exploited in the very last step of the proof to delete the $i_0 \perp i$ boxes in front of the formula A. Such a step is useless, if the maximum index i_0 of the axioms in Ω is i. In other words if we do not use any axioms on the upper theory to prove a theorem at level i. $i = i_0$ and (54) becomes:

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A^+ \tag{56}$$

and, the application of (47) is not necessary.

This means that the theorem should be provable for an MR system in which there exists at least one theory which does not have any other upper theory. This is indeed the case for MBK (definition 3.4 on page 18) and its theory labeled with 0. We state this fact formally in the next subsection

Example 7.1 Consider the following proof of $\langle \neg T(\text{“}\perp\text{”}),\, i\rangle$ in the MR system MKB

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle T(\text{“}\perp\text{”}),\, i\rangle}{\langle \perp,\, i \perp 1\rangle}\,\mathcal{R}_{dn.i-1}}{\langle \neg\neg\perp,\, i \perp 1\rangle}\,\perp_{i-1}}{\langle T(\text{“}\neg\neg\perp\text{”}),\, i\rangle}\,\mathcal{R}_{up.i-1} \qquad \dfrac{\dfrac{\dfrac{\langle \perp,\, i+1\rangle}{\langle \neg\perp,\, i+1\rangle}\,\supset\mathrm{I}_{i+1} \quad \langle \neg\perp \supset T(\text{“}\neg T(\text{“}\neg\neg\perp\text{”})\text{”}),\, i+1\rangle}{\langle T(\text{“}\neg T(\text{“}\neg\neg\perp\text{”})\text{”}),\, i+1\rangle}\,\supset\mathrm{E}_{i+1}}{\langle \neg T(\text{“}\neg\neg\perp\text{”}),\, i\rangle}\,\mathcal{R}_{dn.i}}{\langle \perp,\, i\rangle}\,\supset\mathrm{E}_i}{\langle \neg T(\text{“}\perp\text{”}),\, i\rangle}\,\perp_i \tag{57}$$

The translation of $\neg T(\text{“}\perp\text{”})$ is not a theorem of KB, hence the two systems are not equivalent. Notice that in the proof (57), we use an axiom at level $i+1$; *i.e.* $\langle \neg\perp \supset T(\text{“}\neg T(\text{“}\neg\neg\perp\text{”})\text{”}),\, i+1\rangle$.

Remark 7.2 From the previous remark we can conclude that the sublevel property does not hold for the MR systems MKB, MK5, MKD5, MK45, MKD45 and MKB4. In fact in each of these systems, we have theorems at some level which are provable only by using axioms at a higher level.

Remark 7.3 The systems based on MK proposed so far, are characterized by the fact that the axioms are homogeneously added to all the levels. In other words: an axiom A is added to each i-theory such that A is an L_i-wff. These systems have corresponding modal logics. On the other hand, there are MR systems based on MK, which can be obtained by adding axioms in one or a specific subset of theories (see for instance example 6.1 at the beginning of section 6. Such systems do not have any modal counterpart. The ability to add axioms locally to theories is very important in metatheoretic theorem proving and in modelling of propositional attitudes [GS93, GSGF92].

7.3 MK vs. MBK

The goal of this subsection is to make explicit the connection between MK and MBK. We will show that each normal modal logic is equivalent to an MR system based on MBK (see remark 7.1).

Theorem 7.4 *If $KS_1 \ldots S_n$ is a modal system respecting property* (47)*, then* $\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A^*$ *if and only if there exists an index i such that* $\vdash_{\mathrm{MKS}_1 \ldots \mathrm{S}_n} \langle A,\ i\rangle$.

Proof Suppose that $\vdash_{\mathrm{MKS}_1 \ldots \mathrm{S}_n} \langle A,\ i\rangle$; by corollary 6.5 on page 55, let Ω be a finite subset of the characteristic axioms of $\mathrm{MKS}_1 \ldots \mathrm{S}_n$ such that:

$$\vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle \tag{48}$$

Let i_0 be the greatest index of the elements in Ω. By theorem 6.1 on page 55

$$\Omega' \vdash_{\mathrm{MK}} \langle A,\ i\rangle \tag{49}$$

where $\Omega' = \{\langle T^{i_0-j}(\text{“}B\text{”}),\ i_0\rangle : \langle B,\ j\rangle \in \Omega\}$. By shifting up the conclusion of (49) to i_0 we obtain:

$$\Omega' \vdash_{\mathrm{MK}} \langle T^{i_0-i}(\text{“}A\text{”}),\ i_0\rangle \tag{50}$$

By the equivalence between K and MK, (theorem 7.1 on page 60) we have that:

$$\Omega'^{+} \vdash_{\mathrm{K}} \Box^{i_0-i} A^{+} \tag{51}$$

Since any wff in Ω is the translation, via $()^*$, of an axiom of $\mathrm{KS}_1 \ldots \mathrm{S}_n$ and $()^+$ is the inverse of $(.)^*$, then for each $\langle B,\ j\rangle \in \Omega$:

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} B^{+} \tag{52}$$

and by (SC2):

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} \Box^{i_0-j} B^{+} \tag{53}$$

From (51) and (53) by repeated applications of (CUT) we have that:

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} \Box^{i_0-i} A^{+} \tag{54}$$

By (47) we conclude that:

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A^{+} \tag{55}$$

Viceversa: Suppose that $\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A$. For the same argument given at the beginning of the proof of lemma 7.2, there exists a finite set of axioms Ω and a finite sequence of applications of the $\vdash_{\mathrm{K}+\Omega}$'s derivability property which terminates in $\vdash_{\mathrm{K}+\Omega} A$. Let i be the maximum depth of the formulas involved in this sequence. Notice that the i-theory of $\mathrm{MKS}_1 \ldots \mathrm{S}_n$ contains the translation of the axioms in Ω and it is closed under the derivability conditions defining K. This implies that $\langle A^*,\ i\rangle$ is derivable in $\mathrm{MKS}_1 \ldots \mathrm{S}_n$.
Q.E.D.

Theorem 7.5 ([Che80] Exercise 5.10) *Property* (47) *does not hold for the modal systems KB, K5, KD5, K45, KD45 and KB4.*

Notation 7.3 We write $\Diamond A$ as an abbreviation of $\neg\Box\neg A$.

Definition 7.3 ($\mathbf{KS_1 \ldots S_n}$) *For any n-pla of schemas $S_1, \ldots, S_n$ in* {D, B, T, 4, 5, G}, *the modal system $KS_1 \ldots S_n$ is the minimal modal system containing K and the axioms schemas $S_1, \ldots, S_n$; where* D, B, T, 4, 5 *and* G *are the following schemas:*

D. $\Box A \supset \Diamond A$
B. $A \supset \Box\Diamond A$
T. $\Box A \supset A$
4. $\Box A \supset \Box\Box A$
5 $\Diamond A \supset \Box\Diamond A$
G $\Box(\Box A \supset A) \supset \Box A$

For exhaustive descriptions of the systems containing the first 5 schemas (D, T, B, 4 and 5) see [Che80]; for a description of KG see [Boo79, Smo85].

Definition 7.4 ($\mathbf{MKS_1 \ldots S_n}$) *For each modal system $KS_1 \ldots S_n$, $MKS_1 \ldots S_n$ is an MR system based on MK such that:*

$$MKS_1 \ldots S_n = MK + \langle S_1^+, 1\rangle + \ldots + \langle S_n^+, 1\rangle + \langle S_1^+, 2\rangle + \ldots + \langle S_n^+, 2\rangle + \ldots$$

where for each $1 \le k \le n$, $\langle S_k, i\rangle$ is the translation by $(.)^+$ of the schema S_k, restricted to the language L_i.

We want now to prove a theorem which states that any $L(\Box)$-wff A is provable in $\mathrm{KS}_1 \ldots \mathrm{S}_n$ if and only if there is some $i \in \omega$ such that $\langle A, i\rangle$ is provable in $\mathrm{MKS}_1 \ldots \mathrm{S}_n$. In order to state this theorem, let us first recall the soundness property for MR systems (proposition 3.1 on page 14). Point (i) of this proposition, instantiated to an MR system $\mathrm{MKS}_1 \ldots \mathrm{S}_n$, says that:

$$\vdash_{\mathrm{MKS}_1 \ldots \mathrm{S}_n} \langle T(\text{``}A\text{''}),\ i+1\rangle \text{ if and only if } \vdash_{\mathrm{MKS}_1 \ldots \mathrm{S}_n} \langle A,\ i\rangle$$

This means that we can successfully state the equivalence theorem for those modal systems which respect the following property:

$$\vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} \Box A \text{ if and only if } \vdash_{\mathrm{KS}_1 \ldots \mathrm{S}_n} A \tag{47}$$

Theorem 7.2 ([Che80] Exercise 4.61) *Property* (47) *holds for the modal systems K, KD, KT, K4, KDB, KD4, KTB, KT4 and KT5.*

Theorem 7.3 *Property* (47) *holds for the modal system KG.*

Proof By Solovay's completeness theorem (see [Sol70]) an $L(\Box)$-wff A is provable in KG if and only if, for every realization Φ, A^Φ is a theorem of PA.

Hence if $\vdash_{\mathrm{KG}} \Box A$ then by Solovay's completeness theorem, for each realization Φ $\vdash_{\mathrm{PA}}$ $\mathrm{Bew}(\ulcorner A^\Phi \urcorner)$, which implies that $\vdash_{\mathrm{PA}} A^\Phi$ and again by Solovay's completeness theorem $\vdash_{\mathrm{KG}} A$. Q.E.D.

(LM): see lemma 4.4 on page 28;

(CUT): Suppose that $\langle G, i\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$ and $\langle G, j\rangle, \langle A, j\rangle \vdash_{\mathrm{MK}} \langle B, j\rangle$. Let $k = \max(i, j)$. By lemma 4.6 on page 32, $\langle G, k\rangle \vdash_{\mathrm{MK}} \langle A, k\rangle$ and $\langle G, k\rangle, \langle A, k\rangle \vdash_{\mathrm{MK}} \langle B, k\rangle$. By applying (MR–CUT) (theorem 4.1 on page 28), $\langle G, k\rangle \vdash_{\mathrm{MK}} \langle B, k\rangle$.

(SB): If $\langle G, i\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$, p is a propositional letter and B is an L_j-wff, then $\langle G^p_B, i + j\rangle \vdash_{\mathrm{MK}} \langle A^p_B, i + m\rangle$.

Boolean Properties: They are exactly the classical ND inference rules.

(SC2): see proposition 3.1 on page 14.

Q.E.D.

Theorem 7.1 (Equivalence of MK and K's provability relation) *For any set $G \cup \{A\}$ of L_i-wffs, $\langle G, i\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$ if and only if $G^+ \vdash_{\mathrm{K}} A^+$. For any set $G \cup \{A\}$ of $L(\Box)$-wffs $G \vdash_{\mathrm{K}} A$, if and only if there exists a finite subset $G' \subseteq G$ and a natural number i such that, $\langle G'^*, i\rangle \vdash_{\mathrm{MK}} \langle A^*, i\rangle$.*

Proof Trivially, by lemma 7.1 and lemma 7.2. Q.E.D.

Corollary 7.1 *For any n-pla $A_1, \ldots, A_n$ of $L(\Box)$-wffs*

$$A_1 \ldots A_n \vdash_{\mathrm{K}} A \text{ if and only if } \langle A_1^*, i\rangle, \ldots, \langle A_n^*, i\rangle \vdash_{\mathrm{MK}} \langle A^*, i\rangle$$

where i is the maximum depth of $A_1, \ldots, A_n$ and A.

7.2 Normal Modal Logics

Normal modal systems can be defined by adding K (the minimal normal modal system) a set of axioms. In the following we consider as examples the fifteen systems presented in [Che80], that is KD, KT, KB, K4, K5, KDB, KD4, KD5, K45, KD45, KB4, KTB, KT4, KT5 plus the system KG presented in [Boo79].

The methodology is as follows. We provide MK with suitable families of sets of axioms which are the translation by $(.)^*$ of the modal axioms. In most of the cases we succeed *i.e.* we obtain an MR system based on MK such that: "if A is provable in a normal modal system, then $\langle A, i\rangle$ is provable in the corresponding MR system, for some index i". In some other cases (*e.g.* KB) the simple addition of the translation of the axioms, does not lead to an equivalent MR system.

We individuate a minimal condition (property (47)) under which a modal logic is translatable (by adding the translation of its axioms) in a system based on MK. However some modal logics do not satisfy this minimal condition. Two things cause this problem: the language structure of MK and/or the form of the equivalence statement. Remark 7.1 individuates this problem and subsection 7.3, describes a structure of languages and an equivalence statement by which it is possible to give a general equivalence result between modal logics and MR systems.

Lemma 7.1 (From MK to K) *For any set $G \cup \{A\}$ of L_i-wffs, if $\langle G,\, i\rangle \vdash_{\mathrm{MK}} \langle A,\, i\rangle$, then $G^+ \vdash_{\mathrm{K}} A^+$.*

Proof Let Π be a weak normal deduction of $\langle A,\, i\rangle$ from $\langle G,\, i\rangle$. By induction on Π we prove that $G^+ \vdash_{\mathrm{K}} A^+$.

[base case] If Π is $\langle A,\, i\rangle$ then $A^+ \in G^+$. By (RX) and (LM), we have that $G^+ \vdash_{\mathrm{K}} A^+$.

[step case] Suppose that Π ends with the application of an i-rule σ_i. If σ_i is one of $\wedge\mathrm{I}_i, \wedge\mathrm{E}_i, \vee\mathrm{I}_i, \vee\mathrm{E}_i, \supset\mathrm{I}_i$ and $\supset\mathrm{E}_i$ then, straightforwardly, by using the induction hypothesis and by applying the modal version of σ_i, $G^+ \vdash_{\mathrm{K}} A^+$. If σ_i is $\bot_i$ then by induction hypothesis, $G^+, \neg A^+ \vdash_{\mathrm{K}} \bot$. In $L(\Box)$, $\bot$ is not a primitive symbol, it can be seen as an abbreviation of $p \wedge \neg p$ for some propositional letter p. By applying the boolean properties ($\wedge$E), ($\neg$E), and (RAA) we obtain that $G^+ \vdash_{\mathrm{K}} A^+$. If Π ends with an application of $\mathcal{R}_{up.i}$, then Π is of the form:

$$\dfrac{\dfrac{\begin{matrix}\langle G_1,\, i\rangle\\ \Pi_1\\ \langle T(\text{“}A_1\text{”}),\, i\rangle\end{matrix}}{\langle A_1,\, i \perp 1\rangle} \quad \cdots \quad \dfrac{\begin{matrix}\langle G_n,\, i\rangle\\ \Pi_n\\ \langle T(\text{“}A_n\text{”}),\, i\rangle\end{matrix}}{\langle A_n,\, i \perp 1\rangle}}{\begin{matrix}\Pi'\\ \dfrac{\langle A,\, i \perp 1\rangle}{\langle T(\text{“}A\text{”}),\, i\rangle}\end{matrix}}$$

where Π' is a deduction of $\langle A,\, i \perp 1\rangle$ from $\langle A_1,\, i \perp 1\rangle, \ldots, \langle A_n,\, i \perp 1\rangle$ and Π_k $1 \le k \le n$ are deductions of $\langle T(\text{“}A_k\text{”}),\, i\rangle$ depending on $\langle G_k,\, i\rangle$ and $\bigcup_{1 \le k \le n}(\langle G_k,\, i\rangle) \subseteq \langle G,\, i\rangle$. By the induction hypothesis on Π_k, we obtain that $G_k^+ \vdash_{\mathrm{K}} \Box A_k^+$. By the induction hypothesis on Π', $A_1^+, \ldots, A_n^+ \vdash_{\mathrm{K}} A^+$ and by (SC2), $\Box A_1^+, \ldots, \Box A_n^+ \vdash_{\mathrm{K}} \Box A^+$. By n applications of (CUT) and (LM), we obtain that $G^+ \vdash_{\mathrm{K}} A^+$. Π cannot end with an application on a $\mathcal{R}_{dn.i}$ with premiss $\langle T(\text{“}A\text{”}),\, i+1\rangle$, as it is weak normal and no formulas with index greater than i can occur in Π. Q.E.D.

Lemma 7.2 (From K to MK) *For any set $G \cup \{A\}$ of $L(\Box)$-wffs, if $G \vdash_{\mathrm{K}} A$, then there exists a subset $G' \subseteq G$ and a natural number i such that, $\langle G'^*,\, i\rangle \vdash_{\mathrm{MK}} \langle A^*,\, i\rangle$.*

Proof $\vdash_{\mathrm{K}}$ is the minimal relation closed under the properties listed in definition 7.1. If $G \vdash_{\mathrm{K}} A$, then there exists a finite sequence $G_1 \vdash_{\mathrm{K}} A_1, \ldots, G_n \vdash_{\mathrm{K}} A_n$ such that: $G_n = G$, $A_n = A$ and every $G_k \vdash_{\mathrm{K}} A_k$, is either of the form $B \vdash_{\mathrm{K}} B$, or it is obtained by applying one of the deduction properties of $\vdash_{\mathrm{K}}$ to one or more elements appearing before in the sequence. It is therefore sufficient to show that the translations of the K derivability properties are preserved by MK. In general there exists no language L_i containing all the wffs in G^* (for example when the depths of the wffs in G are not upper bounded) and so the expression $\langle G^*,\, i\rangle$ does not have any sense. This is not a serious problem, as it is easy to verify that $\vdash_{\mathrm{K}}$ is compact, in the sense that, $G \vdash_{\mathrm{K}} A$ if and only if there exists a finite subset $G' \subseteq G$, such that $G' \vdash_{\mathrm{K}} A$. This allows us to consider only the instances of the rules of $\vdash_{\mathrm{K}}$ involving only finite sets. In the following we show that the translation of the derivability properties of $\vdash_{\mathrm{K}}$ are preserved by $\vdash_{\mathrm{MK}}$.

(RX): let $i = depth(A)$, then $\langle A^*,\, i\rangle \vdash_{\mathrm{MK}} \langle A^*,\, i\rangle$.

7.1 Modal K

Let L be a propositional language. We define the modal language $L(\Box)$ as the minimal set of wffs containing L, closed under the usual rules for the logical connectives plus the following rule: if A is in $L(\Box)$ then $\Box A$ is in $L(\Box)$. For an introduction to Modal Logic we refer the reader to [BS84, HC72, Che80]

Following [BS84], the deducibility relation $\vdash_K$ on $L(\Box)$ is characterized by the following properties:

Structural properties:

(RX) $A \vdash_K A$;
(LM) *if* $\Gamma \vdash_K A$ *and* $\Gamma \subseteq \Delta$ *then* $\Delta \vdash_K A$;
(CUT) *if* $\Gamma \vdash_K C$ *and* $C, \Gamma \vdash_K A$ *then* $\Gamma \vdash_K A$;
(SB) *if* $\Gamma \vdash_K A$ *then* $\Gamma^p_B \vdash_K A^p_B$; *where* p *is a propositional letter and* B *a* $L(\Box)$*-wff.*

Boolean properties:

($\wedge$E) *if* $\Gamma \vdash_K A \wedge B$, *then* $\Gamma \vdash_K A$ *and* $\Gamma \vdash_K B$;
($\wedge$I) *if* $\Gamma \vdash_K A$ *and* $\Gamma \vdash_K B$, *then* $\Gamma \vdash_K A \wedge B$;
($\vee$E) *if* $\Gamma \vdash_K A \vee B$ *and* $\Gamma, A \vdash_K C$ *and* $\Gamma, B \vdash_K C$, *then* $\Gamma \vdash_K C$;
($\vee$I) *if* $\Gamma \vdash_K A$ *or* $\Gamma \vdash_K B$, *then* $\Gamma \vdash_K A \vee B$;
($\supset$E) *if* $\Gamma \vdash_K A$ *and* $\Gamma \vdash_K A \supset B$, *then* $\Gamma \vdash_K B$;
($\supset$I) *if* $A, \Gamma \vdash_K B$, *then* $\Gamma \vdash_K A \supset B$;
($\neg$E) *if* $\Gamma \vdash_K A$ *and* $\Gamma \vdash_K \neg A$, *then* $\Gamma \vdash_K B$;
($\neg$I) *if* $A, \Gamma \vdash_K \neg A$, *then* $\Gamma \vdash_K \neg A$;
(RAA) *if* $\neg A, \Gamma \vdash_K A$, *then* $\Gamma \vdash_K A$;

Normal property:

(SC2) *if* $\Gamma \vdash_K A$, *then* $\Box\Gamma \vdash_K \Box A$.

Definition 7.1 (Modal K) *K is the minimal normal modal logic. Its derivability relation is characterized by the structural properties (RX), (LM), (CUT), (SB), by all the boolean properties and the normal property (SC2).*

Definition 7.2 *The mapping* $(.)^*$ *from* $L(\Box)$*-wffs to* L_i*-wffs is defined as follows:*

(i) *If* A *is a propositional constant then* $A^* = A$;
(ii) $(.)^*$ *distributes over the propositional connectives;*
(iii) $(\Box A)^* = T(\text{“}A^*\text{”})$.

Notation 7.1 $(.)^*$ is an isomorphism and has an inverse, $(.)^+$.

Notation 7.2 The depth of a modal wff (L_i-wff) A (in symbols $depth(A)$) is the greatest number of nested modal operators (T predicates) in A.

Corollary 6.8 *If i_0 is the greatest index of the wffs in $\Gamma \cup \{\langle A,\, i\rangle\}$, then*

$$\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A,\, i\rangle \text{ if and only if } \Gamma, \Omega' \vdash_{\mathrm{MK}} \langle A,\, i\rangle$$

where $\Omega' = \{\langle T^{i_0-j}(\text{“}B\text{”}),\, j\rangle : \langle B,\, j\rangle \in \Omega \text{ and } i_0 \leq j\} \cup \Omega|_{i_0}$.

Remark 6.5 Corollary 6.8 states the connection between MK and any system based on MK with the same language. Hence we can suitably reformulate any theorem about MK for any system based on it. In the following we propose some of them.

Corollary 6.9 *MK+Ω + $\langle A,\, i\rangle$ is equivalent to MK+Ω + $\langle T(\text{“}A\text{”}),\, i+1\rangle$, i.e. for any set of wffs Γ and any L_j-wff B:*

$$\Gamma \vdash_{\mathrm{MK}+\Omega\,+\,\langle A,\, i\rangle} \langle B,\, j\rangle \text{ if and only if } \Gamma \vdash_{\mathrm{MK}+\Omega\,+\,\langle T(\text{“}A\text{”}),\, i+1\rangle} \langle B,\, j\rangle$$

Corollary 6.10 *If Ω is a finite set of wffs,then there exists a natural number i and an L_i-wff A such that: $\nvdash_{\mathrm{MK}+\Omega} \langle A,\, i\rangle$.*

Remark 6.6 Corollary 6.10 says that, as long as we add finitely many axioms, MK cannot get into inconsistency. This result holds even if we add contradictory axioms. Let us suppose we have added the wff $\langle \bot,\, i\rangle$. We will be able to derive everything in the i-th theory and also in all the theories below it. However in the $i+1$-theory we can only derive $\langle T(\text{“}\bot\text{”}),\, i+1\rangle$. $\langle T(\text{“}\bot\text{”}),\, i+1\rangle$ is an atomic wff and, without any other axioms, it does not lead to inconsistency.

Remark 6.7 In corollary 6.10, the hypothesis that the set of axioms is finite can be relaxed by requiring that the maximum level of the axioms is finite.

Corollary 6.11 $\nvdash_{\mathrm{MK}+\langle \bot,\, i\rangle} \langle \bot,\, i+1\rangle$.

7 ML systems vs. Modal Logics

In this section we prove that the most common normal modal logics for instance K, K4, K45, T, S4, S5 and G are embeddable in suitable MR systems based on MK. First we give a general introduction of the modal system K and prove the equivalence result between K and MK. Then, we combine this result with theorem 6.1 and prove the equivalence between some normal modal systems and their corresponding multilanguage versions. The section ends by describing some new bridge rules which allow us to define the axiom free versions of the ML systems introduced.

(ii) *If* $i > j$: $\vdash_{\mathrm{MK}+\langle B,\ j\rangle} \langle A,\ i\rangle$ *if and only if* $\langle T^{i-j}(\text{“}B\text{”}),\ i\rangle \vdash_{\mathrm{MK}} \langle A,\ i\rangle$.

Proof If Π is a proof in MK+$\langle B,\ j\rangle$ of $\langle A,\ i\rangle$ then it is a quasi-deduction of $\langle A,\ i\rangle$ from $\langle B,\ j\rangle$ in MK. Let i_0 be the maximum index of the occurrences of Π. We show that the quasi-deduction:

$$\begin{array}{c}
\dfrac{\langle T^{i_0-j}(\text{“}B\text{”}),\ i_0\rangle}{\langle T^{i_0-j-1}(\text{“}B\text{”}),\ i_0 \perp 1\rangle}\ \mathcal{R}_{dn.i_0-1} \\
\vdots \\
\dfrac{\langle T(\text{“}B\text{”}),\ j+1\rangle}{(\langle B,\ j\rangle)}\ \mathcal{R}_{dn.j} \\
\Pi \\
\langle A,\ i\rangle
\end{array} \tag{45}$$

is a deduction of $\langle A,\ i\rangle$ from $\langle T^{i_0-j}(\text{“}B\text{”}),\ i_0\rangle$. Let τ be any thread of (45), then τ is either a thread of Π or it is of the form

$$\langle T^{i_0-j}(\text{“}B\text{”}),\ i_0\rangle \ldots \langle T(\text{“}B\text{”}),\ j+1\rangle\langle B,\ j\rangle\tau' \tag{46}$$

where τ' is a thread of Π.

In the first case τ is open or closed since Π is a deduction. In the second case we have that, since i_0 is the greatest index which appears in Π, the index of each occurrence of τ is lower than or equal to i_0; hence τ is either open or closed. By lemma 4.1 on page 25, this implies that (45) is a deduction of $\langle A,\ i\rangle$ from $\langle T^{i_0-j}(\text{“}B\text{”}),\ i_0\rangle$. Let Π' be the deduction obtained by pushing down (45) to the greatest level between i and j, by means of the “pushing operator” given in definition 4.3 on page 30. If $i \leq j$ then Π' is a deduction in MK of $\langle A,\ i\rangle$ from $\langle B,\ j\rangle$. If $i > j$, then Π' is a deduction in MK of $\langle A,\ i\rangle$ from $\langle T^{i-j}(\text{“}B\text{”}),\ i\rangle$.

The viceversa is provable by applying (MR–CUT) (corollary 6.3) to

$$\vdash_{\mathrm{MK}+\langle B,\ j\rangle} \langle B,\ j\rangle \text{ and } \langle B,\ j\rangle \vdash_{\mathrm{MK}} \langle A,\ i\rangle$$

if $i \leq j$, and to

$$\vdash_{\mathrm{MK}+\langle B,\ j\rangle} \langle T^{i-j}(\text{“}B\text{”}),\ i\rangle \text{ and } \langle T^{i-j}(\text{“}B\text{”}),\ i\rangle \vdash_{\mathrm{MK}} \langle A,\ i\rangle$$

if $i > j$. Q.E.D.

Remark 6.3 The main technical difference between assumptions and axioms in MK is a consequence of the restriction on the applicability of reflection up. Thus, if $\langle A,\ i\rangle$ is derived from $\langle B,\ i\rangle$, then $\langle T(\text{“}A\text{”}),\ i+1\rangle$ can be derived by applying reflection up only if $\langle B,\ i\rangle$ is an axiom.

Remark 6.4 Theorem 6.1 essentially says that assumptions behave as axioms for all the levels below that where they are assumed. This is why certain axioms must be lifted up. Assumptions, to behave as axioms, must be high enough in the hierarchy of levels. This seems to suggest that, in a sense, axioms are not necessary, in other words, that axioms could be considered as assumptions of a metatheory at a sufficiently high level.

Corollary 6.5 (Lemma 4.4 page 28) *$\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle$ if and only if there exists a finite set $\Gamma' \subseteq \Gamma$ and a finite set $\Omega' \subseteq \Omega$ such that $\Gamma' \vdash_{\mathrm{MK}+\Omega'} \langle A, i\rangle$.*

Corollary 6.6 (Lemma 4.7 page 34) *$\Gamma|_{i+n} \vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle$, if and only if $\Gamma|_{i+n} \vdash_{\mathrm{MK}+\Omega} \langle T^n(\text{“}A\text{”}), i+n\rangle$.*

The proofs or corollaries 6.1, 6.2, 6.3, 6.4, 6.5 and 6.6 proceed the same as those of the corresponding lemmas or theorems.

Corollary 6.7 *Let Ω, Σ be any two sets of wffs. For any L_i-wff A and L_j-wff B:*

(i) $\vdash_{\mathrm{MK}+\Omega+\langle A,\, i\rangle} \langle A, i\rangle$
(ii) *if* $\vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle$, *then* $\vdash_{\mathrm{MK}+\Omega+\Sigma} \langle A, i\rangle$
(iii) *if* $\vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle$ *and* $\Gamma \vdash_{\mathrm{MK}+\Sigma+\langle A,\, i\rangle} \langle B, j\rangle$ *then* $\Gamma \vdash_{\mathrm{MK}+\Omega+\Sigma} \langle B, j\rangle$.

The proof is analogous to that of lemma 4.1, page 28.

Remark 6.1 In corollary 6.7, the first precondition of property (iii), *i.e.* $\vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle$ requires that $\langle A, i\rangle$ is provable in MK+Ω. One can ask if (iii) holds even if we allow $\langle A, i\rangle$ to be derivable from a set of hypothesis Λ, *i.e.* if the following condition holds:

$$\text{if } \Lambda \vdash_{\mathrm{MK}+\Omega} \langle A, i\rangle \text{ and } \Gamma \vdash_{\mathrm{MK}+\Sigma+\langle A,\, i\rangle} \langle B, j\rangle \text{ then } \Lambda, \Gamma \vdash_{\mathrm{MK}+\Omega+\Sigma} \langle B, j\rangle \tag{43}$$

It is easy to see that (43) does not hold. A counter-example is:

$$\langle A, i\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle \text{ and } \vdash_{\mathrm{MK}+\langle A,\, i\rangle} \langle T(\text{“}A\text{”}), i+1\rangle \text{ but } \langle A, i\rangle \nvdash_{\mathrm{MK}} \langle T(\text{“}A\text{”}), i+1\rangle \tag{44}$$

Remark 6.2 Axioms are stronger than assumptions, in the sense that what is derivable from a set of assumptions is also derivable from the same set of axioms, but not viceversa. As a simple example notice that

$$\vdash_{\mathrm{MK}+\langle A,\, i\rangle} \langle T(\text{“}A\text{”}), i+1\rangle$$

but

$$\langle A, i\rangle \nvdash_{\mathrm{MK}} \langle T(\text{“}A\text{”}), i+1\rangle$$

This implies that a theorem for axioms corresponding to the effective assumptions theorem (lemma 4.2 on page 27), cannot be stated.

The next theorem clarifies the relationship between axioms and assumptions and gives a criterium for transforming axioms into assumptions and viceversa.

Theorem 6.1 (Axioms vs. Assumptions) *Let A be an L_i-wff and B an L_j-wff, then:*

(i) *If $i \leq j$: $\vdash_{\mathrm{MK}+\langle B,\, j\rangle} \langle A, i\rangle$ if and only if $\langle B, j\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$;*

Deductions (i) and (iii) are also deductions in MK by taking each occurrence of the axiom $\langle p,\ 0\rangle$ as an assumption of the same wff. However this is not the case with deductions (ii) and (iv).

Example 6.2 Let MS be the ML system based on MK with no characteristic symbols and the set $\{\langle p,\ i\rangle : i \geq 0\}$ of characteristic axioms. The following properties hold:

(i) $\vdash_{\mathrm{MS}} \langle p,\ i\rangle$;
(ii) $\vdash_{\mathrm{MS}} \langle T(\text{“}p\text{”}),\ i+1\rangle$;
(iii) $\langle T(\text{“}p \supset q\text{”}),\ i+1\rangle \vdash_{\mathrm{MS}} \langle q,\ i\rangle$;
(iv) $\langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}),\ i+1\rangle \vdash_{\mathrm{MS}} \langle q,\ i\rangle$.

The deductions of (i)–(iv) can be obtained by substituting i for 0 and $i+1$ for 1 in the deductions of the previous example.

While in MS, each theory is a conservative extension of the one below, this does not hold for MK+$\langle p,\ 0\rangle$. Theorem 5.5 holds for MS and does not hold for MK+$\langle p,\ 0\rangle$.

In the following we will consider only the systems based on MK where the languages are not extended, *i.e.* no new symbol is introduced.

Notation 6.1 Let MS$'$ be the ML system based on MS, such that for each $i \in I$, $L_i = L'_i$ and Ω is the set of its characteristic axioms; then we shortly write MS$'$ =MS+Ω. We also write MS+$\langle A,\ i\rangle$ instead of MS+$\{\langle A,\ i\rangle\}$.

Most of the theorems proved for MK in section 4 can be generalized to any ML system MK+Ω. We summarize them below.

Corollary 6.1 (Lemma 4.1 page 25) *A quasi-deduction Π is a deduction in MK+Ω, if and only if, for each thread $\tau = \langle A_1,\ i_1\rangle, \ldots, \langle A_n,\ i_n\rangle$ starting with an assumption,*

(i) *τ is* closed*: i.e. $\langle A_1,\ i_1\rangle$ is discharged at $\langle A_h,\ i_h\rangle$ and for each $1 \leq k < h$, $i_1 \geq i_k$, or*
(ii) *τ is* open*: i.e. $\langle A_1,\ i_1\rangle$ is undischarged and for each $1 \leq k \leq n$, $i_1 \geq i_k$.*

Corollary 6.2 (Lemma 4.2 page 27) $\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$ *if and only if* $\Gamma\lfloor_i \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$.

Corollary 6.3 (Theorem 4.1 page 28) *Let Ω be any set of wffs. Then $\vdash_{\mathrm{MK}+\Omega}$ is an ML derivability relation and satisfies the following properties:*

(MR–RX) $\langle A,\ i\rangle \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$.
(MR–LM) *if* $\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$ *then* $\Gamma, \Sigma \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$.
(MR–CUT) *if* $\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A,\ i\rangle$ *and* $\Sigma, \langle A,\ i\rangle \vdash_{\mathrm{MK}+\Omega} \langle B,\ j\rangle$*then* $\Gamma, \Sigma \vdash_{\mathrm{MK}+\Omega} \langle B,\ j\rangle$.

Corollary 6.4 (Theorem 4.2 page 28) $\Gamma, \langle A,\ i\rangle \vdash_{\mathrm{MK}+\Omega} \langle B,\ i\rangle$ *if and only if* $\Gamma \vdash_{\mathrm{MK}+\Omega} \langle A \supset B,\ i\rangle$

6 ML Systems Based on MK

Given a logical system S composed of a set of logical symbols, a set of logical axioms and a set of logical rules, a theory T in S is completely specified by a set of symbols, called the *nonlogical symbols* of T, and a set of wffs, called the *nonlogical axioms* of T. Analogously, given an ML system MS, we can define an ML system MS′ "based on" MS, by specifying for each language a set of symbols and a set of axioms. The languages and the sets of axioms of MS′ are obtained by adding the new symbols and axioms to the languages and axioms of MS.

The goal of this section is to define and study MR systems based on MK. Some of the basic properties of these systems can be given as a straightforward generalization of lemmas and theorems proved for MK. Other properties are proved by transforming the axioms of an MR system based on MK into suitable assumptions of MK (theorem 6.1) and by applying the properties of MK.

Definition 6.1 (...Based on...) *Let* $MS = \langle\{L_i\}_{i\in I}, \{\Omega_i\}_{i\in I}, \Delta\rangle$ *and* $MS' = \langle\{L'_i\}_{i\in I}, \{\Omega'_i\}_{i\in I}, \Delta'\rangle$ *two ML systems. We say that MS′ is* based on *MS, if and only if, for every* $i \in I$, $L_i \subseteq L'_i$, $\Omega_i \subseteq \Omega'_i$, *and* $\Delta = \Delta'$. *Every element in* L'_i *which is not in* L_i *and every element of* Ω'_i *which is not in* Ω_i *is respectively a* characteristic symbol *and a* characteristic axiom *of MS′ with respect to MS.*

Example 6.1 Let MK+$\langle p, 0\rangle$ be the ML system based on MK with no characteristic symbols and with $\langle p, 0\rangle$ as a characteristic axiom. The following properties hold:

(i) $\vdash_{\mathrm{MK}+\langle p,\,0\rangle} \langle p, 0\rangle$;

(ii) $\vdash_{\mathrm{MK}+\langle p,\,0\rangle} \langle T(\text{“}p\text{”}), 1\rangle$;

(iii) $\langle T(\text{“}p \supset q\text{”}), 1\rangle \vdash_{\mathrm{MS}+\langle p,\,0\rangle} \langle q, 0\rangle$;

(iv) $\langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}), 1\rangle \vdash_{\mathrm{MK}+\langle p,\,0\rangle} \langle q, 0\rangle$.

In fact:

(i)

$$\langle P, 0\rangle$$

(ii)

$$\dfrac{\langle p, 0\rangle}{\langle T(\text{“}p\text{”}), 1\rangle}\,\mathcal{R}_{up.0}$$

(iii)

$$\dfrac{\langle p, 0\rangle \qquad \dfrac{\langle T(\text{“}p \supset q\text{”}), 1\rangle}{\langle p \supset q, 0\rangle}\,\mathcal{R}_{dn.0}}{\langle q, 0\rangle}\,\supset\!\mathrm{E}_0$$

(iv)

$$\dfrac{\dfrac{\dfrac{\langle p, 0\rangle}{\langle T(\text{“}p\text{”}), 1\rangle}\,\mathcal{R}_{up.0} \qquad \langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}), 1\rangle}{\langle T(\text{“}q\text{”}), 1\rangle}\,\supset\!\mathrm{E}_1}{\langle q, 0\rangle}\,\mathcal{R}_{dn.0}$$

$$\dfrac{\dfrac{\begin{matrix}\Pi_1\\ \langle T(\text{“}D_1\text{”}),\ i+1\rangle\end{matrix}}{\langle D_1,\ i\rangle}\quad \cdots \quad \dfrac{\begin{matrix}\Pi_n\\ \langle T(\text{“}D_n\text{”}),\ i+1\rangle\end{matrix}}{\langle D_n,\ i\rangle}}{\begin{matrix}\Pi''\\ \langle D,\ i\rangle\end{matrix}}$$
$$\dfrac{\langle D,\ i\rangle}{\langle T(\text{“}D\text{”}),\ i+1\rangle}$$

Each $\langle T(\text{“}D_k\text{”}),\ i+1\rangle$, $(1 \leq k \leq n)$ is neither the consequence of an E-rule (for the subformula property) nor the consequence of an I-rule or $\bot$-rule (because Π is normal). This means that each $\langle T(\text{“}D_k\text{”}),\ i+1\rangle$ is an assumption and that it belongs to $\langle T(\text{“}G\text{”}),\ i+1\rangle$. From this we conclude that $\langle G,\ i\rangle \vdash_{\text{MK'}} \langle D,\ i\rangle$ for some $\langle D,\ i\rangle \in \langle H,\ i\rangle$.

(ii) If $\langle T(\text{“}D\text{”}),\ i+1\rangle$ is the consequence of an $\bot_{i+1}$, Π' is of the form:

$$\dfrac{\begin{matrix}\langle \neg T(\text{“}D\text{”}),\ i+1\rangle\\ \Pi''\\ \langle \bot,\ i+1\rangle\end{matrix}}{\langle T(\text{“}D\text{”}),\ i+1\rangle}$$

For the induction hypothesis on Π'', there exists an element of $\langle C,\ i\rangle \in \langle H,\ i\rangle$ such that $\langle G,\ i\rangle \vdash_{\text{MK'}} \langle C,\ i\rangle$. Q.E.D.

The following corollary is a consequence of theorem 5.6

Corollary 5.2 *For any L_i-wff A, B;*

(i) $\nvdash_{\text{MK}} \langle \neg T(\text{“}B\text{”}),\ i+1\rangle$;

(ii) $\langle T(\text{“}A\text{”}),\ i+1\rangle \nvdash_{\text{MK}} \langle \neg T(\text{“}B\text{”}),\ i+1\rangle$.

Remark 5.9 Corollary 5.2 states that for any L_i-wff B, $\neg T(\text{“}B\text{”})$ is not a theorem in $i+1$ (item (i)). Furthermore $\neg T(\text{“}B\text{”})$ is not derivable from any set of assumptions of the form $T(\text{“}A\text{”})$. Intuitively, the metatheories of MK are so weak that they cannot prove the consistency of their object theory.

We conclude this section with some interesting properties of the derivability relation of MK.

Corollary 5.3 *For any set of L_i-wffs $G \cup \{A\}$, the following are properties of $\vdash_{\text{MK}}$:*

$$\langle G,\ i+1\rangle \vdash_{\text{MK}} \langle A,\ i+1\rangle \quad \text{iff} \quad \langle G,\ i\rangle \vdash_{\text{MK}} \langle A,\ i\rangle \tag{36}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}),\ i+1\rangle \quad \text{iff} \quad \langle G,\ i\rangle \vdash_{\text{MK}} \langle A,\ i\rangle \tag{37}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}) \wedge T(\text{“}B\text{”}),\ i+1\rangle \quad \text{iff} \quad \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}),\ i+1\rangle \text{ and } \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}B\text{”}),\ i+1\rangle \tag{38}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}) \vee T(\text{“}B\text{”}),\ i+1\rangle \quad \text{iff} \quad \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}),\ i+1\rangle \text{ or } \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}B\text{”}),\ i+1\rangle \tag{39}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}) \wedge T(\text{“}B\text{”}),\ i+1\rangle \quad \text{iff} \quad \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A \wedge B\text{”}),\ i+1\rangle \tag{40}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A\text{”}) \supset T(\text{“}B\text{”}),\ i+1\rangle \quad \text{iff} \quad \langle T(\text{“}G\text{”}),\ i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}A \supset B\text{”}),\ i+1\rangle \tag{41}$$

$$\langle T(\text{“}G\text{”}),\ i+1\rangle \nvdash_{\text{MK}} \langle \neg T(\text{“}A\text{”}),\ i+1\rangle \tag{42}$$

Corollary 5.1 *For each L_i-wff A,*

$$\vdash_{\text{MK}} \langle A,\, i\rangle \text{ if and only if there exists a } j \geq i \text{ such that } \vdash_{\text{MK}} \langle A,\, j\rangle.$$

Remark 5.8 The combination of corollary 5.1 and the sublevel property (theorem 4.3) provides a very positive result from a computational point of view, since it means that the provability of an L_i-wff involves only the first i-theories of the MK's hierarchy.

Theorem 5.6 *Let $\langle G, i\rangle$ and $\langle H, i\rangle$ be two (possible empty) sets of L_i-wffs. $\langle T(\text{``}G\text{''}), i+1\rangle, \langle \neg T(\text{``}H\text{''}),\, i+1\rangle \vdash_{\text{MK}} \langle \bot,\, i+1\rangle$ if and only if there exists a $C \in H$ such that $\langle G,\, i\rangle \vdash_{\text{MK}} \langle C,\, i\rangle$.*

Proof The "if" direction is trivial. For the "only if" direction, we split the proof in two cases.

[case 1] $\langle H,\, i\rangle = \emptyset$. We prove that $\langle T(\text{``}G\text{''}),\, i+1\rangle \nvdash_{\text{MK}} \langle \bot,\, i+1\rangle$. The proof is by contradiction. Let $\Pi$ be a strong normal deduction of $\langle \bot,\, i+1\rangle$ from $\langle T(\text{``}G\text{''}),\, i\rangle$. The bottom formula of Π *i.e.*, $\langle \bot,\, i+1\rangle$ (i) cannot be the consequence of the $\bot_{i+1}$, (see the restriction in definition of remark 4.1 page 20); (ii) does not occur immediately below an assumption of the form $\langle \neg A,\, i\rangle$, (iii) cannot be the consequence an I-rule, (*iv*) cannot be the consequence of an E-rule, its premiss would not be a subformula of any element of $\langle T(\text{``}G\text{''}),\, i+1\rangle \cup \{\langle \bot,\, i+1\rangle\}$. This proves that there does not exist any deduction of $\langle \bot,\, i+1\rangle$ from $\langle T(\text{``}G\text{''}),\, i+1\rangle$.

[case 2] $\langle H, i\rangle \neq \emptyset$. Let Π be a strong normal deduction of $\langle \bot,\, i+1\rangle$ from $\langle T(\text{``}G\text{''}),\, i\rangle \cup \langle \neg T(\text{``}H\text{''}),\, i+1\rangle$. We show by induction on Π that, if Π a strong normal deduction of $\langle \bot,\, i+1\rangle$ from $\langle T(\text{``}G\text{''}),\, i+1\rangle \cup \langle \neg T(\text{``}H\text{''}),\, i+1\rangle$, then there exists a $\langle C,\, i\rangle \in \langle H,\, i\rangle$ such that $\langle G,\, i\rangle \vdash_{\text{MK}} \langle C,\, i\rangle$.

Let $\langle \neg T(\text{``}D\text{''}),\, i+1\rangle \in \langle \neg T(\text{``}H\text{''}),\, i+1\rangle$ be an undischarged assumption of Π. Such an assumption must exist, as $\langle T(\text{``}G\text{''}),\, i+1\rangle \nvdash_{\text{MK}'} \langle \bot,\, i+1\rangle$ (by case 1). For the subformula property, $\langle \neg T(\text{``}D\text{''}),\, i+1\rangle$ is not a premiss of an I-rule, it is not a premiss of a $\bot_{i+1}$ and it is not the minor premiss of an E-rule. This means that $\langle \neg T(\text{``}D\text{''}),\, i+1\rangle$ must be the major premiss of an $\supset\text{E}_{i+1}$ and that Π is of the form:

$$\cfrac{\langle \neg T(\text{``}D\text{''}),\, i+1\rangle \quad \begin{matrix}\Pi' \\ \langle T(\text{``}D\text{''}),\, i+1\rangle\end{matrix}}{\begin{matrix}\langle \bot,\, i+1\rangle \\ \vdots \\ \langle \bot,\, i+1\rangle\end{matrix}}$$

$\langle T(\text{``}D\text{''}),\, i+1\rangle$ is not the consequence of an E-rule (by subformula property). Therefore it can be (i) the consequence of an I-rule or (ii) the consequence of an $\bot_{i+1}$.

(i) If $\langle T(\text{``}D\text{''}),\, i+1\rangle$ is the consequence of an I-rule, it is the consequence of a $\mathcal{R}_{up.i}$. Π' is thus of the form:

Since Π_0 is a strong normal proof, the subformula property holds for Π_0. Furthermore the occurrences of Π_1 which are not in Π_0 are subformulas of $\langle A_1 \supset T^{i_1-i_2}(``A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(``A_n \supset T^{i_n-i}(``A")")"), i_1\rangle$. Let Π be the deduction of $\langle A, i\rangle$ from $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle$ obtained by removing all the maximum formulas of Π_1. Π is in normal form. Notice that the reduction steps do not introduce any occurrence which is not a subformula of some occurrences of Π_1, except for the occurrences of $\langle \bot, j\rangle$. This implies that any occurrence of Π is a subformula of $\langle A_1 \supset T^{i_1-i_2}(``A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(``A_n \supset T^{i_n-i}(``A")")"), i_1\rangle$. Q.E.D.

Remark 5.6 Deduction (iii) of proposition 4.2 has the weak subformula property.

5.4 Consequences of the Subformula Property

Theorem 5.5 (Shifting down deductions) *For any n-pla of wffs* $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle$, $\langle A, i\rangle$,

$\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle \vdash_{\text{MK}} \langle A, i\rangle$ *if and only if* $\langle A_1, i_1+1\rangle, \ldots, \langle A_n, i_n+1\rangle \vdash_{\text{MK}} \langle A, i+1\rangle$

Proof The only if direction is given by lemma 4.6 on page 32. Suppose that $\langle A_1, i_1 + 1\rangle, \ldots, \langle A_n, i_n + 1\rangle \vdash_{\text{MK}} \langle A, i + 1\rangle$. We require, without loss of generality, that $i_1 \geq i_2 \geq \ldots \geq i_n \geq i$. By theorem 5.4, let Π be a normal deduction such that each of its occurrences is a subformula of

$$\langle A_1 \supset T^{i_1-i_2}(``A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(``A_n \supset T^{i_n-i}(``A")")"), i_1 + 1\rangle \tag{35}$$

Since each A_k are L_{i_k}-wff and A is an L_i-wff, (35) is an L_{i_1}-wff. For the weak subformula property each occurrence $\langle B, j\rangle$ of Π is (i) a subformula of (35) or (ii) an assumption discharged by the application of $\bot_j$ or (iii) $\langle \bot, j\rangle$.

(i) Since $\langle B, j\rangle$ is a subformula of (34) which is an L_{i_1}-wff, then B is an L_{j-1}-wff.

(ii) if $\langle B, j\rangle$ is discharged by a $\bot_j$-rule then it is of the form $\langle \neg C, j\rangle$. $\langle C, j\rangle$ is a subformula of (34). C is an L_{j-1}-wff. Hence B (*i.e.* $\neg C$) is an L_{j-1}-wff.

(iii) B is $\bot$. If $j > 0$ then B is an L_{j-1}-wff. If $i = 0$ then $\bot$ is not a subformula of (34), which means that it is not the end formula of Π. Let $\langle p, 0\rangle$ the formula which occurs immediately below $\langle \bot, 0\rangle$. $\langle p, 0\rangle$ cannot be a subformula of (34) which contradicts the fact that Π is normal. This means that $j > 0$.

Let Π' be the tree obtained by substituting each occurrence $\langle B, j\rangle$ of Π with $\langle B, j - 1\rangle$. Analogously to what done in the proof of lemma 4.6 we can prove that Π' is a deduction of $\langle A, i\rangle$ from $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle$ Q.E.D.

Remark 5.7 Theorem 5.5 confirms what we have anticipated in section 4 (remark 4.11 page 33). Theorem 5.5 implies that each $i + 1$-theory is indeed a conservative extension of the i-theory.

have any interest as $\langle G, i\rangle$ are not effective assumptions, hence $\vdash_{\text{MK}} \langle A, j\rangle$). For the subformula property (theorem 5.3) there exists a deduction Π, whose occurrences which are not discharged by a $\perp_k$, are subformulas either of an element of $\langle G, i\rangle$ or of $\langle T^{i-j}(\text{“}A\text{”}), i\rangle$. This means that the following deduction:

$$\dfrac{\begin{array}{c}\langle G, i\rangle \\ \Pi \\ \langle T^{i-j}(\text{“}A\text{”}), i\rangle\end{array}}{\begin{array}{c}\langle T^{i-j-1}(\text{“}A\text{”}), i \perp 1\rangle \\ \vdots \\ \dfrac{\langle T(\text{“}A\text{”}), j+1\rangle}{\langle A, j\rangle}\ \mathcal{R}_{dn.j}\end{array}}\ \mathcal{R}_{dn.i-1}$$

is a deduction of $\langle A, j\rangle$ from $\langle G, i\rangle$, which satisfies conjecture 5.1.

On the other hand conjecture 5.1 does not hold for deduction (iii) of proposition 4.2 on page 23.

Despite of the counter-example (iii) of proposition 4.2, we can prove a weaker result by transforming any deduction between different levels into an equivalent deduction which starts and terminates at the same level, and by applying the subformula property.

Theorem 5.4 (Weak subformula property) *Let $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle\langle A, i\rangle$ be an (n+1)-tupla of wffs with $i_1 \geq i_2 \geq \ldots \geq i_n \geq i$, such that $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle \vdash_{\text{MK}} \langle A, i\rangle$. Then there exists a normal deduction Π such that: every occurrence in Π is a subformula of $\langle A_1 \supset T^{i_1-i_2}(\text{“}A_2 \supset \ldots T^{i_{n\perp1}-i_n}(\text{“}A_n \supset T^{i_n-i}(\text{“}A\text{”})\text{”})\text{”}), i_1\rangle$, except for the assumptions discharged by an application of the $\perp$-rule, and the occurrences of the form $\langle\perp, j\rangle$ occurring immediately below an assumption.*

Proof If $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle \vdash_{\text{MK}} \langle A, i\rangle$, then by deduction theorem (theorem 4.2 page 28) and lemma 4.7 page 34 we have that:

$$\vdash_{\text{MK}} \langle A_1 \supset T^{i_1-i_2}(\text{“}A_2 \supset \ldots \supset T^{i_{n\perp1}-i_n}(\text{“}A_n \supset T^{i_n-i}(\text{“}A\text{”})\text{”})\text{”}), i_1\rangle \tag{34}$$

Let Π_0 be a strong normal proof of (34). Let Π_1 the (possibly not normal) deduction built as follows:

$$\dfrac{\langle A_n, i_n\rangle \qquad \begin{array}{c} \dfrac{\dfrac{\langle A_1, i_1\rangle \quad \begin{array}{c}\Pi_0 \\ \langle A_1 \supset T^{i_1-i_2}(\text{“}A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(\text{“}A_n \supset T^{i_n-i}(\text{“}A\text{”})\text{”})\text{”}), i_1+1\rangle\end{array}}{\langle T^{i_1-i_2}(\text{“}A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(\text{“}A_n \supset T^{i_n-i}(\text{“}A\text{”})\text{”})\text{”}), i_1\rangle}\ \supset\!\text{E}_{i_1}}{\langle T^{i_1-i_2-1}(\text{“}A_2 \supset \ldots \supset T^{i_{n-1}-i_n}(\text{“}A_n \supset T^{i_n-i}(\text{“}A\text{”})\text{”})\text{”}), i_1 \perp 1\rangle}\ \mathcal{R}_{dn.i_1-1} \\ \vdots \\ \langle A_n \supset T^{i_n-i}(\text{“}A\text{”}), i_n\rangle \end{array}}{\begin{array}{c}\langle T^{i_n-i}(\text{“}A\text{”}), i_n\rangle \\ \vdots \\ \dfrac{\langle T(\text{“}A\text{”}), i+1\rangle}{\langle A, i\rangle}\ \mathcal{R}_{dn.i}\end{array}}\ \supset\!\text{E}_{i_n}$$

the assumptions discharged by an application of the $\perp$-rule, and the occurrences of the form $\langle \perp,\ j\rangle$ occurring immediately below an assumption.

Proof For any branch in a normal deduction we define its *order* as follows. The main branches have order 0. A branch that ends with a formula occurrence that is a minimum premiss of an $\supset$E-rule, whose maximum premiss belongs to a branch with order n, has order $n+1$.

Let Π be a normal deduction of $\langle A,\ i\rangle$ from $\langle G,\ i\rangle$. Let $\beta = \langle A_1,\ i_1\rangle, \ldots, \langle A_n,\ i_n\rangle$ be a branch with order n and $\langle A_h,\ i_h\rangle$ its minimum formula. Each occurrence in the I-part of β is a subformula of the last formula of β, $\langle A_n,\ i_n\rangle$. $\langle A_n,\ i_n\rangle$ is the minor premiss of an $\supset$E-rule; hence it is a subformula of the major premiss that belongs to a branch with order $n \perp 1$. We conclude that each occurrence of the I-part of a branch of order n is a subformula of an occurrence in a branch of order $n \perp 1$.

Each occurrence in the E-part is a subformula of $\langle A_1,\ i_1\rangle$ and

(i) if $\langle A_1,\ i_1\rangle$ is undischarged, then it belongs to $\langle G,\ i\rangle$;

(ii) if $\langle A_1,\ i_1\rangle$ is discharged by an application of an $\supset\mathrm{I}_{i_1}$, then it is a subformula of a formula occurring in the I-part of β or in a branch with order less than n;

(iii) if $\langle A_1,\ i_1\rangle$ is discharged by an application $\perp_{i_1}$-rule then it is the major premiss of an $\supset$E-rule and the minimum formula $\langle A_h,\ i_h\rangle$ is the consequence (which is equal to $\langle \perp,\ i_h\rangle$). This means that no other occurrences are in the E-part.

Q.E.D.

The previous theorem states the subformula property for deductions that start and end at the same level. As stressed before, the subformula property doesn't hold in the general case. Nevertheless we think that the following weaker results holds.

Conjecture 5.1 *Let $\Gamma \cup \{\langle A,\ i\rangle\}$ be a finite set of wffs and i_0 its greatest index. If $\Gamma \vdash_{\mathrm{MK}} \langle A,\ i\rangle$, then there exists a normal deduction Π such that: every occurrence in Π is a subformula of $\langle T^{i_0-j}(\text{“}C\text{”}),\ i_0\rangle$, for some $\langle C,\ j\rangle \in \Gamma \cup \{\langle A,\ i\rangle\}$, except for the assumptions discharged by an application of the $\perp$-rule, and the occurrences of the form $\langle \perp,\ j\rangle$ occurring immediately below an assumption.*

Example 5.5 There are several examples in favour of this conjecture.

(a) conjecture 5.1 holds for deduction (32), namely

$$\cfrac{\cfrac{\langle \perp,\ i+1\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}$$

(b) Conjecture 5.1 holds for deductions which start at a level j and terminate at level i, with $i \neq j$. Namely if $\langle G,\ i\rangle \vdash_{\mathrm{MK}} \langle A,\ j\rangle$ and $i \neq j$. If $i \geq j$ then $\langle G,\ i\rangle \vdash_{\mathrm{MK}} \langle T^{i-j}(\text{“}A\text{”}),\ i\rangle$ by lemma 4.7 (the case with $i < j$ does not wffs of the form $\langle \neg T(\text{“}A\text{”}),\ i\rangle$ on which the $\langle \bot,\ i\rangle$ that occurs immediately above $\langle T(\text{“}A\text{”}),\ i\rangle$ depends. Take N as the maximum of the $dd(\langle \neg T(\text{“}A\text{”}),\ i\rangle)$ associated to the elements of this set.

(iv) $dd(\langle \neg T(\text{“}A\text{”}),\ i\rangle) = 0$ iff it is not the major premiss of an $\supset \mathrm{E}_i$.

(v) $dd(\langle \neg T(\text{“}A\text{”}),\ i\rangle) = dd(\langle T(\text{“}A\text{”}),\ i\rangle)$, where $\langle T(\text{“}A\text{”}),\ i\rangle$ is the minor premiss of the $\supset \mathrm{E}_i$.

Let $MAX(\Pi) = IE(\Pi) \cup FT(\Pi)$ be the set of maximum formulas in Π. Let $dd(\Pi)$ be the maximum dependency degree of the elements of $MAX(\Pi)$ which are of the form $\langle T(\text{“}A\text{”}),\ i\rangle$ or $\langle \neg T(\text{“}A\text{”}),\ i\rangle$. Let $m(\Pi)$ be the maximum complexity of the wffs in $MAX(\Pi)$. Let $n(\Pi)$ be the number of occurrences of complexity equal to $m(\Pi)$ in $MAX(\Pi)$. For any two deductions Π_1, Π_2 we say that, $\Pi_1 \succ_{MAX} \Pi_2$ if and only if (i) or (ii) or (iii):

(i) $dd(\Pi_1) > dd(\Pi_2)$;

(ii) $dd(\Pi_1) = dd(\Pi_2)$ and $m(\Pi_1) > m(\Pi_2)$;

(iii) $dd(\Pi_1) = dd(\Pi_2)$ and $m(\Pi_1) = m(\Pi_2)$ and $n(\Pi_1) > n(\Pi_2)$.

Note that, if n is the dependency degree of the maximum formula removed by the $\frac{=}{T}$-reduction, then the maximum formulas introduced have dependency degree less than n and are of the form $\langle \neg T(\text{“}A\text{”}),\ i\rangle$. Furthermore the dependency degree of the maximum formulas introduced by a $\supset$-reduction (in order to remove $\langle \neg T(\text{“}A\text{”}),\ i\rangle$) is less than or equal to the dependency degree of $\langle \neg T(\text{“}A\text{”}),\ i\rangle$. This ensures that, if Π_1 and Π_2 are the deductions respectively at the beginning and at the end of the iteration, then $\Pi_1 \succ_{MAX} \Pi_2$. Therefore the termination of NORMALIZER is guaranteed. Q.E.D.

Lemma 5.5 (Branch Shape in Strong Normal Deductions) *Let Π be a strong normal deduction in MK′ and $\beta = \langle A_1,\ i_1\rangle, \ldots, \langle A_n,\ i_n\rangle$ a branch of Π. Then there is a formula occurrence $\langle A_h,\ i_h\rangle$ in β, called the* minimum formula *of β, that separates two (possible empty) parts of β, called the* E-part *and the* I-part *of β, with the following properties:*

(i) *each $\langle A_k,\ i_k\rangle$ in the E-part (i.e., $k < h$) is a premiss of an E-rule;*

(ii) *$\langle A_h,\ i_h\rangle$, provided that $h \neq n$ is the premiss of an I-rule or of a $\bot$-rule;*

(iii) *each $\langle A_k,\ i_k\rangle$, in the I-part, except the last one, (i.e., $h < k < n$) is the premiss of an I-rule.*

Proof Since Π is strong normal, then it is normal. By lemma 5.3 let $\langle A_h,\ i_h\rangle$ be the minimum formula of a branch β. Since no occurrence that is a consequence of $\bot$-rule and a premiss of $\mathcal{R}_{dn.}$ occurs in Π, the E-part of β is composed of one subE-part which is the E-part itself. Items (i), (ii) and (iii) of lemma 5.5 correspond at items (iii), (v) and (vi) of lemma 5.3 respectively. Q.E.D.

Theorem 5.3 (Subformula Property) *Every formula occurrence in a normal deduction of $\langle A,\ i\rangle$ from $\langle G,\ i\rangle$ is a subformula of an element of $\langle G,\ i\rangle \cup \{\langle A,\ i\rangle\}$, except for*

```
procedure: NORMALIZER
  begin
    Π ← Π_0;
    while (IE(Π) ∪ FT(Π) ≠ ∅)
      begin
        while (IE(Π) ≠ ∅)
          Π ← IE-reduction(Π);
        while (FT(Π) ≠ ∅)
          Π ← =/T-reduction(Π);
      end;
    return(Π);
  end.
```

IE-reduction(Π) is the result of the application of one of the reduction steps in the proof of lemma 5.1 to Π.

$\frac{=}{T}$-reduction(Π) is the result of the application of $\frac{=}{T}$-*reduction* to Π.

Figure 9: Algorithm for normalizing deductions

is strictly lower than $m(\Pi)$, and removes the occurrence of $FT(\Pi)$ of complexity $m(\Pi)$, that is $\langle T(\text{“}B\text{”}),\ j+1\rangle$. Q.E.D.

We finally compose lemma 5.1 and lemma 5.4 to state a sufficient condition for the existence of a normal form.

Theorem 5.2 (Existence of a Strong Normal Form) *If* $\langle G,\ i\rangle \vdash_{\text{MK}} \langle A,\ i\rangle$, *then there exists a strong normal deduction of* $\langle A,\ i\rangle$ *from* $\langle G,\ i\rangle$.

Proof By lemma 5.2, let Π_0 be a deduction in which the consequences of $\bot$-rules are atomic. We apply the algorithm NORMALIZER described in figure 9 to Π_0. If NORMALIZER terminates, then it returns a strong normal deduction of $\langle A,\ i\rangle$ from $\langle G,\ i\rangle$.

We need to prove the termination of NORMALIZER. We know from lemma 5.1 and lemma 5.4 that the two internal loops terminate. We show that the external loop terminates by defining a well founded relation on deductions.

The *dependency degree* of an occurrence of the form $\langle T(\text{“}A\text{”}), i\rangle$ or $\langle \neg T(\text{“}A\text{”}), i\rangle$, denoted by $dd(\langle T(\text{“}A\text{”}),\ i\rangle)$ and $dd(\langle \neg T(\text{“}A\text{”}),\ i\rangle)$ respectively, is defined as follows:

(i) $dd(\langle T(\text{“}A\text{”}),\ i\rangle) = 0$ iff it is not the consequence of an $\bot_i$-rule;

(ii) $dd(\langle T(\text{“}A\text{”}),\ i\rangle) = 0$ iff the $\langle \bot,\ i\rangle$ that occurs immediately above $\langle T(\text{“}A\text{”}),\ i\rangle$ does not depend on $\langle \neg T(\text{“}A\text{”}),\ i\rangle$;

(iii) $dd(\langle T(\text{“}A\text{”}), i\rangle) = 1 + N$ iff the $\langle \bot, i\rangle$ that occurs immediately above $\langle T(\text{“}A\text{”}), i\rangle$ depends on $\langle \neg T(\text{“}A\text{”}),\ i\rangle$. N is computed as follows. Consider the set of

Remark 5.5 Deductions that start and end at the same level form a very general class. This class contains the proofs of all the formulas which are the translations of the theorems of the modal logic K (see section 7).

Lemma 5.4 *If* $\langle G,\, i\rangle \vdash_{\mathrm{MK'}} \langle A,\, i\rangle$*, then there exists a deduction with no occurrences that are consequences of a* $\bot$*-rule and premisses of a* $\mathcal{R}_{dn.}$*-rule.*

Proof Let Π be a weak normal deduction of $\langle A,\, i\rangle$ from $\langle G,\, i\rangle$ (see theorem 4.3 page 31). Remember that the indexes of the wffs occurring in a weak normal deduction are lower than the maximum index among those of the assumptions and the conclusion. We define $FT(\Pi)$ as the set of occurrences of Π that are consequences of a $\bot$-rule and premisses of $\mathcal{R}_{dn.}$. For any occurrence $\langle T(\text{“}B\text{”}),\, j+1\rangle$ of $FT(\Pi)$, i is greater than j (for the sublevel property). This implies that there exists at least an application of $\mathcal{R}_{up.j}$ occurring below the application of $\mathcal{R}_{dn.j}$ to $\langle T(\text{“}B\text{”}),\, j+1\rangle$. Let $\langle T(\text{“}C\text{”}),\, j+1\rangle$ be the consequence of the first application of a $\mathcal{R}_{up.j}$ in a thread from $\langle T(\text{“}B\text{”}),\, j+1\rangle$ to the end formula $\langle A,\, i\rangle$. Π is of the form shown in the left side of picture (33). We can define a $\frac{-}{T}$-reduction which transforms Π in the deduction Π', shown in the right side of the same picture. Notice that in Π' the awkward occurrence $\langle T(\text{“}B\text{”}),\, j+1\rangle$ has been removed.

$\frac{-}{T}$-reduction

$$\begin{array}{c} \langle \neg T(\text{“}B\text{”}),\, j+1\rangle \\ \Pi_1 \\ \langle \bot,\, j+1\rangle \\ \hline \langle T(\text{“}B\text{”}),\, j+1\rangle \\ \hline \langle B,\, j\rangle \\ \Pi_2 \\ \langle C,\, j\rangle \\ \hline \langle T(\text{“}C\text{”}),\, j+1\rangle \\ \Pi_3 \end{array} \qquad\qquad \begin{array}{c} \begin{array}{cc} \begin{array}{c} \langle T(\text{“}B\text{”}),\, j+1\rangle \\ \hline \langle B,\, j\rangle \\ \Pi_2 \\ \langle C,\, j\rangle \\ \hline \langle T(\text{“}C\text{”}),\, j+1\rangle \end{array} & \langle \neg T(\text{“}C\text{”}),\, j+1\rangle \end{array} \\ \hline \langle \bot,\, j+1\rangle \\ \hline (\langle \neg T(\text{“}B\text{”}),\, j+1\rangle) \\ \Pi_1 \\ \langle \bot,\, j+1\rangle \\ \hline \langle T(\text{“}C\text{”}),\, j+1\rangle \\ \Pi_3 \end{array} \tag{33}$$

As usual, let us define a well founded ordering on deductions and let us show that the $\frac{-}{T}$-reduction generates deductions which are lower in the ordering. Let $m(\Pi)$ be the maximum complexity of the elements of $FT(\Pi)$. Let $n(\Pi)$ be the number of occurrences of complexity equal to $m(\Pi)$ in $FT(\Pi)$. For any two deductions Π_1, Π_2 we say that, $\Pi_1 \succ_{FT} \Pi_2$ if and only if:

(i) $m(\Pi_1) > m(\Pi_2)$ or

(ii) $m(\Pi_1) = m(\Pi_2)$ and $n(\Pi_1) > n(\Pi_2)$.

Let Π' be obtained by applying the $\frac{-}{T}$-reduction to Π. Suppose that $\langle T(\text{“}B\text{”}),\, j+1\rangle$ is the lowest most complex occurrence of $FT(\Pi)$. Then $\Pi \succ_{FT} \Pi'$. Indeed the $\frac{-}{T}$-reduction duplicates only the elements of $FT(\Pi)$ occurring below $\langle T(\text{“}B\text{”}), j+1\rangle$, whose complexity (*a*) Give a *weaker notion of subformula*. The modified notion should be such that, normal deductions (in particular (32)) possess it. In order to make (32) have the subformula property, we have to add to definition 5.6 the following fact: "$\langle T(\text{"}A\text{"}),\ i+1\rangle$ is a subformula of $\langle A,\ i\rangle$". We reject this as a wff would have infinite subformulas. It does not satisfy the intuition that a formula is constructed from its subformulas.

(*b*) Give a *stronger notion of normal deduction*. In this case we give up the fact that every deduction (*e.g.*, (32)) is reducible in normal form. We will follow this direction and give a sufficient condition for the existence of a normal form.

(*c*) *Modify the rules* of MK in order to obtain an equivalent system with the subformula property. Counter-example (32) gives us an insight for a new (admissible) rule that could be added to MK, *i.e.*, for any $j \leq i$:

$$\frac{\langle \bot,\ i\rangle}{\langle A,\ j\rangle}\ \bot_{MK}$$

This possibility is not explored in this paper.

Definition 5.7 (Weak Maximum Formula) *An occurrence $\langle A,\ i\rangle$ in a deduction is a* weak maximum formula *if and only if it satisfies one of the two conditions:*

(i) *$\langle A,\ i\rangle$ is the consequence of an I-rule and the major premiss of an E-rule;*

(ii) *$\langle A,\ i\rangle$ is the consequence of a $\bot$-rule and the major premiss of an E-rule.*

Definition 5.8 (Strong Normal Form) *A deduction is in* strong normal form*, (or it is* strongly normal*) if and only if it does not contain any weak maximum formula.*

Remark 5.2 Notice that a maximum formula (definition 5.2) is a weak maximum formula, and, consequently, that a strong normal deduction is a normal deduction.

Remark 5.3 A deduction in the i-theory, *i.e.* a deduction that contains only applications of i-rules, is in strong normal form if and only if it is in normal form. The strong normal form differs from the normal form only as far as the applications of the bridge rules are concerned.

Remark 5.4 Not all the deductions of MK can be reduced to a strong normal form. For example deduction (32) is not reducible to such a form. In general if $\langle \bot,\ i+1\rangle \vdash_{\text{MK}} \langle A,\ i\rangle$ and $\langle \bot,\ i+1\rangle$ is a "necessary" hypothesis (*i.e.* $\nvdash_{\text{MK}} \langle A,\ i\rangle$), then there exists no strong normal deduction of $\langle A, i\rangle$ from $\langle \bot, i+1\rangle$. Indeed any deduction of $\langle A, i\rangle$ form $\langle \bot, i+1\rangle$ contains at least one occurrence of the form $\langle T(\text{"}B\text{"}),\ i+1\rangle$ that is the premiss of a $\mathcal{R}_{dn.}$ (the only rule that enables one to change level). The problem is that $\langle T(\text{"}B\text{"}),\ i+1\rangle$ is neither a subformula of $\langle \bot,\ i+1\rangle$ nor of $\langle A,\ i\rangle$.

A sufficient condition for the existence of a strong normal form is that conclusion and undischarged assumptions belong to the same level. This does not mean that we consider deductions within a single level, it means only that these deductions start and end at the same level.

Let us consider item (vi). Let $\langle A,\ i\rangle$ be any occurrence in the I-part. It is not the premiss of an E-rule. If it is the premiss of a $\bot$-rule, then it cannot be the consequence of an I-rule and it cannot be the consequence of a $\bot$-rule (see restriction on the $\bot$-rule in remark 4.1). Therefore it is the consequence of an E-rule, which means that it is the minimum formula $\langle A_h,\ i_h\rangle$. This contradicts the fact that it occurs in the I-part. We can conclude that $\langle A,\ i\rangle$ is either the last occurrence of the I-part or the premiss of an I-rule. Q.E.D.

5.3 The Subformula Property

In the single language case, a wff A is said to be a subformula of another wff B if and only if A occurs in B. The formal definition of subformula is given recursively on the wff structure. An intuitive interpretation of the subformula relation is that A is a subformula of B, if and only if, following the inductive definition of wffs, before constructing B you have to construct A. In definition 5.6 we extend this intuitive idea to the formulas of MK.

Definition 5.6 (Subformula)

(i) $\langle A,\ i\rangle$ *is a* subformula *of* $\langle A,\ i\rangle$*;*

(ii) $\langle A, i\rangle$ *is a* subformula *of* $\langle B\supset C, i\rangle$, $\langle B\wedge C, i\rangle$, $\langle B\vee C, i\rangle$ *if* $\langle A, i\rangle$ *is* subformula *of* $\langle B,\ i\rangle$ *or* $\langle C,\ i\rangle$*;*

(iii) $\langle A,\ i\rangle$ *is a* subformula *of* $\langle T(\text{“}B\text{”}),\ j+1\rangle$ *if* $\langle A,\ i\rangle$ *is* subformula *of* $\langle B,\ j\rangle$*;*

(iv) *nothing else is a* subformula*.*

Remark 5.1 In the classical natural deduction system C′ [Pra65] the *subformula property* holds under very general hypotheses. This means that all the occurrences of a normal deduction are subformulas of the assumptions or of the conclusion except for the assumptions discharged by the application of a $\bot_c$. The presence of multiple levels makes the subformula property more complex to state. Indeed, while in classical ND, lemma 5.2 page 38 ensures that the consequence of a $\bot$-rule is not a premiss of an E-rule, in MK this is not always the case. Indeed atomic formulas of the form $T(\text{“}A\text{”})$ can be the premiss of an E-rule (*i.e.*, $\mathcal{R}_{dn.}$). For example consider the following deduction:

$$\cfrac{\cfrac{\langle \bot,\ i+1\rangle}{\langle T(\text{“}A\text{”}),\ i+1\rangle}}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i} \tag{32}$$

(32) does not possess the subformula property, as $\langle T(\text{“}A\text{”}),\ i+1\rangle$ is not a subformula of either $\langle\bot,\ i+1\rangle$ or $\langle A,\ i\rangle$ [1].

In order to prove a subformula property we have three choices (at least):

[1] This counter-example is due to Alex Simpson.

(v) *the minimum formula* $\langle A_h,\, i_h\rangle$ *provided that* $k \neq n$ *is the premiss of a* $\bot$*-rule or of an I-rule;*

(vi) *every occurrence in the I-part, but the last, is the premiss of an I-rule.*

Figure 8 shows the form of a branch in a normal deduction. An example is given below.

Example 5.4 Consider the deduction:

$$\dfrac{\dfrac{\langle C,\, i\rangle \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle A,\, i+2\rangle \quad \langle \neg A,\, i+2\rangle}{\langle \bot,\, i+2\rangle}\supset\mathrm{E}_{i+2}}{\langle T(\text{“}B \wedge T(\text{“}C \supset D\text{”})\text{”}),\, i+2\rangle}\bot_{i+2}}{\langle B \wedge T(\text{“}C \supset D\text{”}),\, i+1\rangle}\mathcal{R}_{dn.i+1}}{\langle T(\text{“}C \supset D\text{”}),\, i+1\rangle}\wedge\mathrm{E}_{i+1}}{\langle C \supset D,\, i\rangle}\mathcal{R}_{dn.i}}{\langle D,\, i\rangle}\supset\mathrm{E}_i}{\langle T(\text{“}D\text{”}),\, i+1\rangle}\mathcal{R}_{up.i} \tag{31}$$

the parts of the main branch of deduction (31) are shown in the following picture:

$$\begin{array}{rcl}
\text{subE}_0\text{-part} & = & \begin{cases} \langle \neg A,\, i+2\rangle \\ \langle \bot,\, i+2\rangle \end{cases} \\
\text{subE}_1\text{-part} & = & \begin{cases} \langle T(\text{“}B \wedge T(\text{“}C \supset D\text{”})\text{”}),\, i+2\rangle \\ \langle B \wedge T(\text{“}C \supset D\text{”}),\, i+1\rangle \\ \langle T(\text{“}C \supset D\text{”}),\, i+1\rangle \\ \langle C \supset D,\, i\rangle \end{cases} \\
\text{minimum formula} & = & \quad \langle D,\, i\rangle \\
\text{I-part} & = & \{\ \langle T(\text{“}D\text{”}),\, i+1\rangle
\end{array}$$

Notice that a subE-part can contain formulas of different layers.

Proof (lemma 5.3) In a branch β the applications of E-rules and $\bot$-rules precede all the applications of I-rules. Indeed if this were not the case, as the consequence of an I-rule cannot be the premiss of an $\bot$-rule, then there should be an occurrence that is the consequence of an I-rule and major premiss of an E-rule, which contradicts the fact that Π is normal.

Let $\langle A_h,\, i_h\rangle$ be the first occurrence in β that is the premiss of an I-rule, if it exists, otherwise let $\langle A_h,\, i_h\rangle$ be $\langle A_n,\, i_n\rangle$. Let $\langle A_h,\, i_h\rangle$ be the *minimum formula*. Let the E-part of β be the subpart of β which starts with $\langle A_1,\, i_1\rangle$ and ends with $\langle A_{h-1},\, i_{h-1}\rangle$. Let the I-part of β be the subpart of β which starts with $\langle A_{h+1},\, i_{h+1}\rangle$ and ends with $\langle A_n,\, i_n\rangle$.

Let $\langle A_{k_1},\, i_{k_1}\rangle, \ldots, \langle A_{k_m},\, i_{k_m}\rangle$ be all the occurrences of the E-part of β which are consequences of a $\bot$-rule and premisses of a $\mathcal{R}_{dn.}$-rule (see figure 8). Let the subE-parts be the subparts of the E-part separated by these occurrences. (Each subE-part, but the first one, begins with an $\langle A_{k_j},\, i_{k_j}\rangle$).

Items (i), (ii) and (iv) are verified by the choice of the subE-parts.

Let us consider item (iii). Let $\langle A,\, i\rangle$ be any occurrence of a subE-part. It is not the premiss of an I-rule. If it is the premiss of a $\bot$-rule, then the consequence is either the premiss of an I-rule and $\langle A,\, i\rangle$ is the minimum formula, or the premiss of an $\mathcal{R}_{dn.}$ and $\langle A,\, i\rangle$ is the last occurrence of the subE-part.

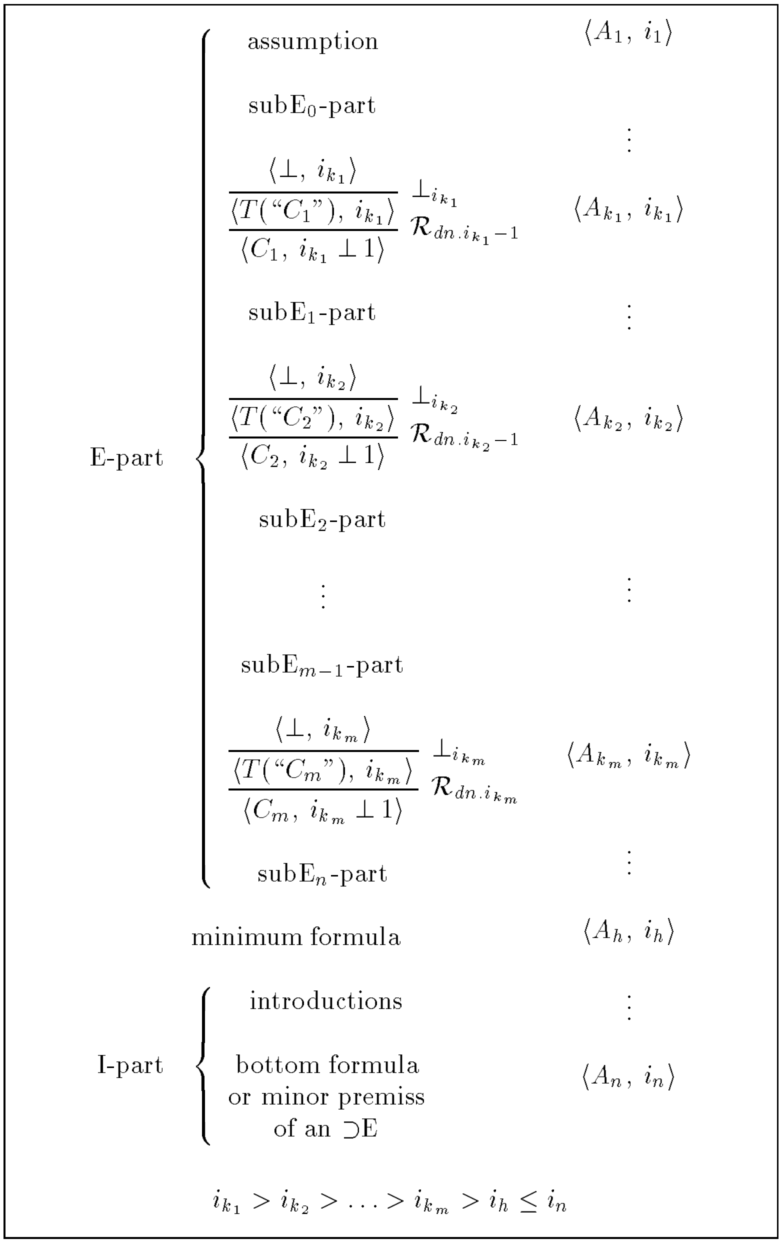

Figure 8: Form of a branch $\beta = \langle A_1,\, i_1\rangle, \ldots, \langle A_n,\, i_n\rangle$ in a normal deduction

.

(i) *The E-part is either empty or composed of* $m+1$ *not empty subE-parts, for some natural number* $m \geq 0$*;*

(ii) *each subE-part, but the first, begins with a premiss of a* $\mathcal{R}_{dn.}$*;*

(iii) *each occurrence of the subE-parts, but the last, is a major premiss of an E-rule;*

(iv) *each subE-part, but the last, ends with a premiss of a* $\bot$*-rule;*

5.2 The Form of Normal Deductions

In this section we study the form of normal deductions in MK′. We then prove some important results like for instance, the holding of the subformula property.

Definition 5.4 (Branch) *Let $\beta = \langle A_1,\, i_1\rangle, \ldots, \langle A_n,\, i_n\rangle$ be the initial part of a thread τ in a deduction Π. We say that β is a* branch *of Π if and only if it satisfies one of the following two conditions:*

- (i) *$\langle A_n,\, i_n\rangle$ is the first formula occurrence of τ, that is the minor premiss of an application of an $\supset E_{i_n}$;*
- (ii) *$\langle A_n,\, i_n\rangle$ is the last formula occurrence of τ and no minor premiss of an application of an $\supset E$ occurs in β.*

A branch which satisfies condition (ii) is called the main branch*.*

Example 5.3 The branches of the deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle T(\text{“}A \supset C\text{”}),\, i+1\rangle}{\langle A \supset C,\, i\rangle}\,\mathcal{R}_{dn.i} \quad \dfrac{\langle A \wedge B,\, i\rangle}{\langle A,\, i\rangle}\,\wedge\mathrm{E}_i}{\langle C,\, i\rangle}\,\supset\mathrm{E}_i \qquad \langle \neg C,\, i\rangle}{\langle \bot,\, i\rangle}\,\supset\mathrm{E}_i}{\langle \neg(A \wedge B),\, i\rangle}\,\bot_i \qquad \langle \neg C,\, i\rangle}{\langle \neg(A \wedge B) \wedge \neg C,\, i\rangle}\,\wedge\mathrm{I}_i}{\langle \neg C \supset \neg(A \wedge B) \wedge \neg C,\, i\rangle}\,\supset\mathrm{I}_i}{\langle T(\text{“}\neg C \supset \neg(A \wedge B) \wedge \neg C\text{”}),\, i+1\rangle}\,\mathcal{R}_{up.i} \tag{30}$$

are:

$$\begin{array}{llll}
\beta_1 = & \langle \neg C,\, i\rangle & \beta_2 = & \langle \neg C,\, i\rangle \\
 & \langle \bot,\, i\rangle & & \langle \neg(A \wedge B) \wedge \neg C,\, i\rangle \\
 & \langle \neg(A \wedge B),\, i\rangle & & \langle \neg C \supset \neg(A \wedge B) \wedge \neg C,\, i\rangle \\
 & \langle \neg(A \wedge B) \wedge \neg C,\, i\rangle & & \langle T(\text{“}\neg C \supset \neg(A \wedge B) \wedge \neg C\text{”}),\, i+1\rangle \\
 & \langle \neg C \supset \neg(A \wedge B) \wedge \neg C,\, i\rangle & & \\
 & \langle T(\text{“}\neg C \supset \neg(A \wedge B) \wedge \neg C\text{”}),\, i+1\rangle & & \\
 & & & \\
\beta_3 = & \langle T(\text{“}A \supset C\text{”}),\, i+1\rangle & \beta_4 = & \langle A \wedge B,\, i\rangle \\
 & \langle A \supset C,\, i\rangle & & \langle A,\, i\rangle \\
 & \langle C,\, i\rangle & &
\end{array}$$

β_1 and β_2 are the main branches of (30).

Definition 5.5 (Part and subpart of a thread) *A part (or subpart) τ of a thread τ' is a sequence of occurrences of τ' such that there exist two (possible empty) parts τ_1 and τ_2 such that $\tau' = \tau_1\tau\tau_2$.*

Lemma 5.3 (Branch shape) *Let Π be a normal deduction and $\beta = \langle A_1,\, i_1\rangle, \langle A_2,\, i_2\rangle, \ldots, \langle A_n,\, i_n\rangle$ a branch of Π. Then there exists an occurrence $\langle A_h,\, i_h\rangle$ in β, called* minimum formula of *β, which separates two possible empty parts of β, called the E-part $(\langle A_1,\, i_1\rangle, \ldots, \langle A_{h-1},\, i_{h-1}\rangle)$ and the I-part $(\langle A_{h+1},\, i_{h+1}\rangle, \ldots, \langle A_n,\, i_n\rangle)$ of β, with the following properties:*

$\underset{\wedge}{=}$-reduction

$$
\begin{array}{c}\langle\neg(C\wedge D),\, j\rangle\\ \Pi_1\\ \dfrac{\langle\bot,\, j\rangle}{\langle C\wedge D,\, j\rangle}\\ \Pi_2\end{array}
\qquad
\begin{array}{c}
\dfrac{
\dfrac{\begin{array}{c}\dfrac{\dfrac{\langle\neg C,\, j\rangle\quad \dfrac{\langle C\wedge D,\, j\rangle}{\langle C,\, j\rangle}\wedge\mathrm{E}_j}{\langle\bot,\, j\rangle}\supset\mathrm{E}_j}{(\langle\neg(C\wedge D),\, j\rangle)}\supset\mathrm{I}_j\\ \Pi_1\\ \langle\bot,\, j\rangle\end{array}}{\langle C,\, i\rangle}\bot_i
\quad
\dfrac{\begin{array}{c}\dfrac{\dfrac{\langle\neg D,\, j\rangle\quad \dfrac{\langle C\wedge D,\, j\rangle}{\langle D,\, j\rangle}\wedge\mathrm{E}_j}{\langle\bot,\, j\rangle}\supset\mathrm{E}_j}{(\langle\neg(C\wedge D),\, j\rangle)}\supset\mathrm{I}_j\\ \Pi_1\\ \langle\bot,\, j\rangle\end{array}}{\langle D,\, i\rangle}\bot_i
}{\{\langle C\wedge D,\, j\rangle\}}\wedge\mathrm{I}_j\\
\Pi_2
\end{array}
\tag{28}
$$

$\underset{\supset}{=}$-reduction

$$
\begin{array}{c}\langle\neg(C\supset D),\, j\rangle\\ \Pi_1\\ \dfrac{\langle\bot,\, j\rangle}{\langle C\supset D,\, j\rangle}\\ \Pi_2\end{array}
\qquad
\begin{array}{c}
\dfrac{\dfrac{\dfrac{\langle C,\, j\rangle\quad\langle C\supset D,\, j\rangle}{\langle D,\, j\rangle}\qquad\langle\neg D,\, j\rangle}{\langle\bot,\, j\rangle}}{(\langle\neg(C\supset D),\, j\rangle)}\\
\Pi_1\\
\dfrac{\dfrac{\langle\bot,\, j\rangle}{\langle D,\, j\rangle}}{\langle C\supset D,\, j\rangle}\\
\Pi_2
\end{array}
\tag{29}
$$

Let $m(\Pi)$ be the maximum complexity of the elements of $FN(\Pi)$. Let $n(\Pi)$ be the number of occurrences of maximum complexity in $FN(\Pi)$. For any two deductions Π_1, Π_2 we say that, $\Pi_1 \succ_{FN} \Pi_2$ if and only if

(i) $m(\Pi_1) > m(\Pi_2)$ or

(ii) $m(\Pi_1) = m(\Pi_2)$ and $n(\Pi_1) > n(\Pi_2)$.

We prove that the application of the reduction steps (28) and (29) to a deduction Π, returns a deduction Π' such that $\Pi \succ_{FN} \Pi'$. The proof can be carried out analogously to the proof of lemma 5.1. Q.E.D.

Theorem 5.1 (Existence of a Normal Deduction) *If $\Gamma \vdash_{\mathrm{MK}'} \langle A,\, i\rangle$, then there exists a normal deduction of $\langle A,\, i\rangle$ from Γ in MK′.*

Proof By lemma 5.2, let Π be a deduction in which all the consequences of a $\bot$-rule are atomic. By applying lemma 5.1 to Π, we obtain a deduction Π' in which any consequence of an I-rule is not the major premiss of an E-rule. Since the reduction steps defined in the proof of lemma 5.1 do not introduce non atomic consequences of $\bot$-rules, we can conclude that Π' is normal. Q.E.D.

(iii) the complexity of $T(\text{“}A\text{”})$, is 1 plus the complexity of A.

Let $m(\Pi)$ be the maximum complexity of the elements of $IE(\Pi)$. Let $n(\Pi)$ be the number of occurrences of maximum complexity in $IE(\Pi)$. For any two deductions Π_1, Π_2 we say that $\Pi_1 \succ_{IE} \Pi_2$ if and only if

(i) $m(\Pi_1) > m(\Pi_2)$ or
(ii) $m(\Pi_1) = m(\Pi_2)$ and $n(\Pi_1) > n(\Pi_2)$.

First of all, notice that the reduction steps (24)–(27) involve the subdeduction that occurs above the maximum formulas. This justifies the following. Among the elements of $IE(\Pi)$ with the same highest complexity, choose the $\langle B, j\rangle$ such that no elements of $IE(\Pi)$ with the same complexity occur above it in Π. Let Π' be the deduction obtained by applying the relative reduction step to Π. If $\langle B, j\rangle$ is the only occurrence of $IE(\Pi)$ with the maximum complexity, then $m(\Pi) > m(\Pi')$. In fact the reduction steps do not introduce any occurrence with complexity greater than or equal to $m(\Pi)$. This implies that $\Pi \succ_{IE} \Pi'$. If there are other occurrences in $IE(\Pi)$ with the same complexity as $\langle B, j\rangle$, then they occur below $\langle B, j\rangle$, and are not duplicated by any application of a reduction step. The reduction steps do not introduce any new occurrence with the same complexity as B and remove $\langle B, j\rangle$. Therefore $m(\Pi) = m(\Pi')$ and $n(\Pi) > n(\Pi')$, which implies that $\Pi \succ_{IE} \Pi'$. Q.E.D.

By lemma 5.1, any deduction in MK can be reduced in a form where each consequence of an I-rule is not the major premiss of an E-rule. This is not yet the normal form of definition 5.3, as lemma 5.1 does not consider those occurrences that are consequences of the $\bot$-rule and major premisses of an E-rule. The $\bot$ rule is very peculiar. Its antisymmetrical behavior, has devastating effects for the existence of a normal form and we need a special treatment for it. First, let us consider the ML system MK′ obtained from MK by excluding the rule for introduction and elimination of the connective $\vee$. If we define $A \vee B$ as $\neg A \supset B$, then MK′ is equivalent to MK. The following lemma introduces a stronger form for deductions in MK′, where the consequences of the $\bot$-rule are atomic. This implies that they cannot be major a premiss of an E-rule different from $\mathcal{R}_{dn.}$.

Lemma 5.2 *If* $\Gamma \vdash_{\mathrm{MK}'} \langle A, i\rangle$ *then there exists a deduction in which the consequences of the* $\bot$*-rule are atomic.*

Proof Let Π be a deduction of $\langle A, i\rangle$ from Γ. We define $FN(\Pi)$ as the set of the non atomic consequences of a $\bot$-rule in Π. For any element of $FN(\Pi)$ of the form $C \wedge D$ or $C \supset D$, Π is of the form showed in the left side of picture (28) and (29), respectively. By applying the respective reduction steps we obtain a new deduction Π', showed in the right side of each picture, in which the consequence of the $\bot$-rule has been simplified.

$\wedge$*-reduction:*

$$\begin{array}{c} \dfrac{\dfrac{\begin{array}{cc}\Pi_1 & \Pi_2\\ \langle C,\ j\rangle & \langle D,\ j\rangle\end{array}}{\langle C\wedge D,\ j\rangle}\wedge\mathrm{I}_j}{\langle C,\ j\rangle^{\bullet}}\wedge\mathrm{E}_j \\ \Pi_3 \end{array} \qquad\qquad \begin{array}{c}\Pi_1\\ \{\langle C,\ j\rangle\}\\ \Pi_3\end{array} \tag{24}$$

Π' is a deduction of $\langle A,\ i\rangle$ from Γ. Indeed $\langle C,\ j\rangle$ in Π' depends on a subset of the assumptions of $\langle C,\ j\rangle^{\bullet}$ in Π.

$\vee$ *-reduction:*

$$\begin{array}{c} \dfrac{\dfrac{\begin{array}{c}\Pi_1\\ \langle C,\ j\rangle\end{array}}{\langle C\vee D,\ j\rangle}\vee\mathrm{I}_j \quad \begin{array}{c}[\langle C,\ j\rangle]\\ \Pi_2\\ \langle E,\ j\rangle\end{array} \quad \begin{array}{c}[\langle D,\ j\rangle]\\ \Pi_3\\ \langle E,\ j\rangle\end{array}}{\langle E,\ j\rangle^{\bullet}}\vee\mathrm{E}_j \\ \Pi_4 \end{array} \qquad\qquad \begin{array}{c}\Pi_1\\ (\langle C,\ j\rangle)\\ \Pi_2\\ \{\langle E,\ j\rangle\}\\ \Pi_4\end{array} \tag{25}$$

Π' is a deduction of $\langle A,\ i\rangle$ from Γ. Indeed by lemma 4.3, $\langle E,\ j\rangle$ in Π' depends on a subset of the assumptions of $\langle E,\ j\rangle^{\bullet}$ in Π. The symmetrical versions of the $\wedge$-reduction and of the $\vee$-reduction, (when the consequence of $\wedge\mathrm{E}_i$ is $\langle D,\ i\rangle$ and the premiss of the $\vee\mathrm{I}_i$ is $\langle D,\ i\rangle$ respectively) are defined analogously.

$\supset$ *-reduction:*

$$\begin{array}{c} \dfrac{\dfrac{\begin{array}{c}[\langle C,\ j\rangle]\\ \Pi_1\\ \langle D,\ j\rangle\end{array}}{\langle C\supset D,\ j\rangle}\supset\mathrm{I}_j \quad \begin{array}{c}\Pi_2\\ \langle C,\ j\rangle\end{array}}{\langle D,\ j\rangle^{\bullet}}\supset\mathrm{E}_j \\ \Pi_3 \end{array} \qquad\qquad \begin{array}{c}\Pi_2\\ (\langle C,\ j\rangle)\\ \Pi_1\\ \{\langle D,\ j\rangle\}\\ \Pi_3\end{array} \tag{26}$$

Π' is a deduction of $\langle A,\ i\rangle$ from Γ. Indeed by lemma 4.3 the occurrence $\langle D,\ j\rangle$ in Π', depends on a subset of the assumptions of $\langle D,\ j\rangle^{\bullet}$ in Π.

T*-reduction:*

$$\begin{array}{c} \dfrac{\dfrac{\begin{array}{c}\Pi_1\\ \langle C,\ j\rangle\end{array}}{\langle T(\text{“}C\text{”}),\ j+1\rangle}\mathcal{R}_{up.j}}{\langle C,\ j\rangle^{\bullet}}\mathcal{R}_{dn.j} \\ \Pi_2 \end{array} \qquad\qquad \begin{array}{c}\Pi_1\\ \{\langle C,\ j\rangle\}\\ \Pi_2\end{array} \tag{27}$$

Π' is a deduction of $\langle A,\ i\rangle$ from Γ.

We have to show that repeated applications of the above reduction steps converge to a deduction without occurrences which are consequences of an I-rule and maximum premisses of an E-rule. We do this by defining a well founded ordering between deductions and by showing that the application of a reduction step always generates a deduction at a lower level. Let us define the complexity of a formula as follows:

(i) the complexity of $p\in P$ and $\bot$ is 0;

(ii) the complexity of $A\wedge B$, $A\vee B$ and $A\supset B$, is 1 plus the maximum complexity of A and B;

Definition 5.2 (Maximum Formula) *An occurrence $\langle A, i\rangle$ in a deduction is a* maximum formula *if and only if it satisfies one of the two conditions:*

(i) *$\langle A, i\rangle$ is the consequence of an I-rule and the major premiss of an E-rule, or*

(ii) *$\langle A, i\rangle$ is the consequence of a $\bot$-rule and the major premiss of an E-rule other than $\mathcal{R}_{dn.}$.*

Definition 5.3 (Normal Form) *A deduction is in* normal form *(or is* normal*) if and only if it does not contain any maximum formula.*

Example 5.1 The following deductions are not in normal form.

$$\dfrac{\dfrac{\langle A, i\rangle \quad \langle B, i\rangle}{\langle A\wedge B, i\rangle}\wedge\mathrm{I}_i}{\langle A, i\rangle}\wedge\mathrm{E}_i \qquad \dfrac{\dfrac{\dfrac{\langle A, i\rangle}{\langle A\supset A, i\rangle}\supset\mathrm{I}_i}{\langle T(\text{“}A\supset A\text{”}), i+1\rangle}\mathcal{R}_{up.i}}{\langle A\supset A, i\rangle}\mathcal{R}_{dn.i} \qquad \dfrac{\dfrac{\langle \bot, i\rangle}{\langle A\wedge B, i\rangle}\bot_i}{\langle A, i\rangle}\wedge\mathrm{E}_i$$

(22.a) (22.b) (22.c)

As it can be seen from example 5.1, maximum formulas are associated with useless steps. These steps can be removed.

Example 5.2 (example 5.1 continued) Deductions (22.a–c) can be reduced in normal form by removing their maximum formulas. The result is the following.

$$\langle A, i\rangle \qquad \dfrac{\langle A, i\rangle}{\langle A\supset A, i\rangle}\supset\mathrm{I}_i \qquad \dfrac{\langle \bot, i\rangle}{\langle A, i\rangle}\bot_i$$

(23.a) (23.b) (23.c)

Lemma 5.1 describes how to remove the maximum formulas satisfying point (i) of definition 5.2. Lemma 5.2 decribes how to remove those satisfying point (ii).

Lemma 5.1 *If $\Gamma \vdash_{\mathrm{MK}} \langle A, i\rangle$, then there exists a deduction of $\langle A, i\rangle$ from Γ with no formula occurrence which is both the consequence of an I-rule and the major premiss of an E-rule.*

Proof Let Π be a deduction of $\langle A, i\rangle$ from Γ. We define $IE(\Pi)$ as the set of occurrences of Π, that are both the consequence of an I-rule and the major premiss of an E-rule. For any element of $IE(\Pi)$, of the form $C\wedge D$, $C\vee D$, $C\supset D$ or $T(\text{“}C\text{”})$, Π is of the form showed in the left side of pictures (24), (25), (26) and (27), respectively. By applying the respective reduction steps we obtain a new deduction Π', showed in the right side of each picture, in which the maximum formula has been removed.

Lemma 4.8 (Shifting assumptions) *If* $\Gamma, \langle A,\, i\rangle \vdash_{\text{MK}} \langle B,\, j\rangle$*, then* $\Gamma, \langle T(\text{“}A\text{”}),\, i+1\rangle \vdash_{\text{MK}} \langle B,\, j\rangle$.

Proof by applying (MR–CUT) to $\langle T(\text{“}A\text{”}),\, i+1\rangle \vdash_{\text{MK}} \langle A,\, i\rangle$ and $\Gamma, \langle A,\, i\rangle \vdash_{\text{MK}} \langle B,\, j\rangle$. Q.E.D.

Remark 4.12 The converse of lemma 4.8 doesn't hold. A counter example is:

$$\langle T(\text{“}\bot\text{”}),\, i+1\rangle \vdash_{\text{MK}} \langle T(\text{“}\bot\text{”}),\, i+1\rangle \tag{18}$$

but:

$$\langle \bot,\, i\rangle \nvdash_{\text{MK}} \langle T(\text{“}\bot\text{”}),\, i+1\rangle \tag{19}$$

This counter example is based on the fact that shifting down the assumption $\langle T(\text{“}\bot\text{”}), i+1\rangle$ of (18) causes it to be no longer effective in (19). The cases in which an assumption becomes uneffective are not the only counter examples for the converse of lemma 4.8. Consider for instance:

$$\langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}),\, i+1\rangle, \langle T(\text{“}p\text{”}),\, i+1\rangle \vdash_{\text{MK}} \langle q,\, i\rangle \tag{20}$$

By shifting the assumptions $\langle T(\text{“}p\text{”}),\, i+1\rangle$ down to $\langle p,\, i\rangle$ we have that:

$$\langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}),\, i+1\rangle, \langle p,\, i\rangle \nvdash_{\text{MK}} \langle q,\, i\rangle \tag{21}$$

Fact (21) is not provable at this stage of the paper, we need the strong normal form result. An intuitive argument is that in order to exploit the implication $\langle T(\text{“}p\text{”}) \supset T(\text{“}q\text{”}), i+1\rangle$ of (21) we have to derive $\langle T(\text{“}p\text{”}),\, i+1\rangle$ from $\langle p,\, i\rangle$, but this is impossible for the restriction of $\mathcal{R}_{up.i}$.

Remark 4.13 The shifting between layers could be easily avoided by making L_{i+1} not contain L_i. For instance, in MPK lemma 4.8 and lemma 4.7 hold but not lemma 4.6.

5 Normal Form for Deductions in MK

5.1 Definition and Existence of Normal Deductions

Definition 5.1 *In an application* $\langle\langle A,\, i\rangle, \langle A \supset B,\, i\rangle, \langle B,\, i\rangle\rangle$ *of an* $\supset E_i$*,* $\langle A \supset B,\, i\rangle$ *and* $\langle A,\, i\rangle$ *are called* major premiss *and* minor premiss*, respectively. In an application* $\langle\langle A \vee B,\, i\rangle, \langle C,\, i\rangle, \langle C,\, i\rangle, \langle C,\, i\rangle\rangle$ *of an* $\vee E_i$*,* $\langle A \vee B,\, i\rangle$ *and the* $\langle C,\, i\rangle$*'s are called* major premiss *and* minor premisses*, respectively. In every other rule application each premiss is a major premiss.*

By I-rule we mean one of $\supset$I, $\wedge$I, $\vee$I and $\mathcal{R}_{up.}$; by E-rule we mean one of $\supset$E, $\wedge$E, $\vee$E and $\mathcal{R}_{dn.}$. $\mathcal{R}_{up.}$ and $\mathcal{R}_{dn.}$ are seen as the introduction of T and the elimination of T, respectively. Furthermore we say that an occurrence is a premiss or the consequence of a rule to mean that it is a premiss or the consequence of an application of that rule.

simply by decreasing the indexes occurring in a deduction, as in the shifting up case. Consider for example the deduction below:

$$\dfrac{\dfrac{\langle q \wedge r,\, 1\rangle}{\langle q,\, 1\rangle}\wedge\mathrm{E}_1 \quad \dfrac{\langle T(\text{“}p\text{”}),\, 1\rangle}{\langle T(\text{“}p\text{”}) \supset T(\text{“}p\text{”}),\, 1\rangle}\supset\mathrm{I}_1}{\dfrac{\langle q \wedge (T(\text{“}p\text{”}) \supset T(\text{“}p\text{”})),\, 1\rangle}{\langle q,\, 1\rangle}\wedge\mathrm{E}_1}\wedge\mathrm{I}_1 \tag{14}$$

(14) is a deduction of $\langle q,\, 1\rangle$ from $\langle q \wedge r,\, 1\rangle$. By decreasing of one level the indexes of its occurrences we obtain:

$$\dfrac{\dfrac{\langle q \wedge r,\, 0\rangle}{\langle q,\, 0\rangle}\wedge\mathrm{E}_0 \quad \dfrac{\langle T(\text{“}p\text{”}),\, 0\rangle}{\langle T(\text{“}p\text{”}) \supset T(\text{“}p\text{”}),\, 0\rangle}\supset\mathrm{I}_0}{\dfrac{\langle q \wedge (T(\text{“}p\text{”}) \supset T(\text{“}p\text{”})),\, 0\rangle}{\langle q,\, 0\rangle}\wedge\mathrm{E}_0}\wedge\mathrm{I}_0 \tag{15}$$

which is not a deduction as, for example, $T(\text{“}p\text{”})$ is not an L_0-wff.

Notice that deduction (14) contains useless steps involving wffs whose indexes cannot be decreased. These occurrences can be removed, and the result is the equivalent deduction:

$$\dfrac{\langle q \wedge r,\, 1\rangle}{\langle q,\, 1\rangle} \tag{16}$$

Deduction (16) can be shifted down to level 0, as its occurrences are all L_0-wffs. It is possible to give a general process for removing useless steps as those in (14). For this purpose, in section 5 we define a strong normal form for deductions. The strong normal form allows, among other things, to shift a deduction down to the lowest level, defined on the basis of the maximal depth of the assumptions and of the end formula.

Lemma 4.7 (Shifting conclusion) $\Gamma\lfloor_{i+n} \vdash_{\mathrm{MK}} \langle A, i\rangle$, *if and only if* $\Gamma\lfloor_{i+n} \vdash_{\mathrm{MK}} \langle T^n(\text{“}A\text{”}), i+n\rangle$.

Proof Let Π be a deduction of $\langle A,\, i\rangle$ from $\Gamma\lfloor_{i+n}$. By n applications of reflection up we obtain the following quasi-deduction:

$$\begin{array}{c} \Gamma\lfloor_{i+n} \\ \Pi \\ \dfrac{\langle A,\, i\rangle}{\langle T(\text{“}A\text{”}),\, i+1\rangle}\,\mathcal{R}_{up.i} \\ \vdots \\ \dfrac{\langle T^{n-1}(\text{“}A\text{”}),\, i+n\perp 1\rangle}{\langle T^{n}(\text{“}A\text{”}),\, i+n\rangle}\,\mathcal{R}_{up.i+n-1} \end{array} \tag{17}$$

(17) is a deduction of $\langle T^n(\text{“}A\text{”}),\, i+n\rangle$ from $\Gamma\lfloor_{i+n}$. Each $\mathcal{R}_{up.i+k}$, $(0 \le k < n)$ is indeed applicable, as the indexes of all undischarged assumptions are greater than or equal to $i+n$. The viceversa is proved by applying (MR–CUT) (theorem 4.1) to $\langle T^n(\text{“}A\text{”}),\, i+n\rangle \vdash_{\mathrm{MK}} \langle A,\, i\rangle$ and $\Gamma \vdash_{\mathrm{MK}} \langle T^n(\text{“}A\text{”}),\, i+n\rangle$. Q.E.D.

Proof We prove by induction that, by substituting each occurrence $\langle B, j\rangle$ with $\langle B, j+1\rangle$, in a deduction Π of $\langle A, i\rangle$ from Γ, we obtain a formula tree Π' that is a deduction of $\langle A, i+1\rangle$ from $\{\langle B, j+1\rangle : \langle B, j\rangle \in \Gamma\}$.

[**Base case**] If Π is the only assumption $\langle A, i\rangle$ then $\langle A, i\rangle \in \Gamma$. This implies that Π', which is $\langle A, i+1\rangle$, is a deduction of $\langle A, i\rangle$ from $\{\langle B, j+1\rangle : \langle B, j\rangle \in \Gamma\}$.

[**Step case**] Let Π be of the form:

$$\frac{\begin{matrix}\Gamma_1 \\ \Pi_1 \\ \langle A_1, i_1\rangle\end{matrix} \quad \dots \quad \begin{matrix}\Gamma_n \\ \Pi_n \\ \langle A_n, i_n\rangle\end{matrix}}{\langle A, i\rangle}\ \sigma_j$$

Notice that in MK:

(a) if $\langle\langle A_1, i_1\rangle, \dots, \langle A_n, i_n\rangle, \langle A, i\rangle\rangle \in \rho(\sigma_j)$ (*i.e.* we have an application of the inference rule σ_j), then $\langle\langle A_1, i_1+1\rangle, \dots, \langle A_n, i_n+1\rangle, \langle A, i+1\rangle\rangle \in \rho(\sigma_{j+1})$ (*i.e.* it is an application of σ_{j+1});

(b) for each $1 \leq k \leq n$, if $\langle B, l\rangle \in d_k(\sigma_j)(\alpha)$ (*i.e.* $\langle B, l\rangle$ is discharged by the application α of σ_j), then $\langle B, l+1\rangle \in d_k(\rho_{j+1})(\alpha')$ (*i.e.* $\langle B, l+1\rangle$ is discharged by the application α' of σ_{j+1});

(c) if $\langle\langle\Gamma_1, \langle A_1, i_1\rangle\rangle, \dots, \langle\Gamma_n, \langle A_n, i_n\rangle\rangle\rangle \notin rest(\sigma_i)$, (*i.e.* σ is applicable), then $\langle\langle\Gamma'_1, \langle A_n, i_n+1\rangle\rangle, \dots, \langle\Gamma'_n, \langle A_n, i_n+1\rangle\rangle\rangle \notin rest(\sigma_{i+1})$, (where for each $1 \leq k \leq n$, $\Gamma'_k = \{\langle B, j+1\rangle : \langle B, j\rangle \in \Gamma_k\}$) (*i.e.* σ_{j+1} is applicable).

From the induction hypothesis the transformation of each deduction Π_k ($1 \leq i \leq n$), denoted by Π'_k, is a deduction of $\langle A_k, i_k+1\rangle$ from $\Gamma'_k = \{\langle C, j+1\rangle : \langle C, j\rangle \in \Gamma_k\}$. From facts ($a$), ($b$), and ($c$), we derive that Π', defined as:

$$\frac{\begin{matrix}\Gamma'_1 \\ \Pi'_1 \\ \langle A_1, i_1+1\rangle\end{matrix} \quad \dots \quad \begin{matrix}\Gamma'_n \\ \Pi'_n \\ \langle A_n, i_n+1\rangle\end{matrix}}{\langle A, i+1\rangle}\ \sigma_{j+1}$$

is a deduction of $\langle A, i+1\rangle$ from $\{\langle B, j+1\rangle : \langle B, j\rangle \in \Gamma\}$. Q.E.D.

Remark 4.10 Lemma 4.6 states that the derivability relation $\vdash_{\text{MK}}$ is monotonic with respect to the hierarchy of theories. As it can be seen from the proof of the lemma, this is due to the fact that each language L_i is a subset of L_{i+1} and that the axioms and the deductive machinery at a level are as strong as those at any previous level. If $n = 0$ (there are no undischarged assumptions) lemma 4.6 states that each theory at level i is an *extension* (as defined in [Sho67] page 41) of the theory at the previous level.

Remark 4.11 Any theory in MK is a *conservative extension* (as defined in [Sho67] page 41) of the theory below. This amounts to proving the converse of lemma 4.6. This result is stated by theorem 5.5 on page 50. At this stage of the paper, we don't have the necessary machinery to prove it. Indeed the converse of lemma 4.6 is not provable

Proof Let Π be a deduction of $\langle A,\ i\rangle$ depending on $\Gamma' \subseteq \Gamma$ and i_0 the maximum index of $\Gamma' \cup \{\langle A,\ i\rangle\}$. By lemma 4.5, $\overline{\Pi}^{i_0}$ is a deduction of $\overline{\langle A,\ i\rangle}^{i_0}$ from $\overline{\Gamma'}^{i_0}$. Since i_0 is the greatest index of the elements of $\Gamma' \cup \{\langle A,\ i\rangle\}$, we have $\overline{\Gamma'}^{i_0} = \Gamma'$ and $\overline{\langle A,\ i\rangle}^{i_0} = \langle A,\ i\rangle$. Furthermore $\overline{\Pi}^{i_0}$ does not contain any overflowing formula, *i.e.* it is a weak normal deduction of $\langle A,\ i\rangle$ from Γ. Q.E.D.

Corollary 4.1 (Consistency) *For any* $i \geq 0$, $\nvdash_{\mathrm{MK}} \langle \bot,\ i\rangle$.

Proof By contradiction: let Π be a proof of $\langle \bot,\ i\rangle$. By lemma 4.5, $\overline{\Pi}^{0}$ is a deduction of $\langle \bot,\ 0\rangle$. Since $\overline{\Pi}^{0}$ does not contain any occurrence with index greater than 0, no applications of $\mathcal{R}_{up.i}$ and $\mathcal{R}_{dn.i}$ are performed in $\overline{\Pi}^{0}$. This means that Classical Propositional Logic is inconsistent, which is false. Q.E.D.

Corollary 4.2 *For any* $i \geq 0$, $\langle \bot,\ i\rangle \nvdash_{\mathrm{MK}} \langle \bot,\ i+1\rangle$.

Proof By contradiction if $\langle \bot,\ i\rangle \nvdash_{\mathrm{MK}} \langle \bot,\ i+1\rangle$, then since $\langle \bot,\ i\rangle$ is not an effective assumption, $\vdash_{\mathrm{MK}} \langle \bot,\ i+1\rangle$. But this contradicts corollary 4.1. Q.E.D.

Corollary 4.3 *For each propositional constant* p *and for any* $i \geq 0$ $\nvdash_{\mathrm{MK}} \langle p,\ i\rangle$.

Proof The same as the proof of corollary 4.1. The contradiction derives from the fact that a propositional constant p is not provable in Classical Propositional Logic. Q.E.D.

Notation 4.4 For any set G of L_i-wffs, $\langle G,\ i\rangle$ denotes the set $\{\langle A,\ i\rangle : A \in G\}$ and $T(\text{"}G\text{"})$ denotes the set $\{T(\text{"}A\text{"}) : A \in G\}$.

Corollary 4.4 (Non generating metatheory) $\langle T(\text{"}G\text{"}),\ i+1\rangle \vdash_{\mathrm{MK}} \langle T(\text{"}A\text{"}),\ i+1\rangle$ *if and only if* $\langle G,\ i\rangle \vdash_{\mathrm{MK}} \langle A,\ i\rangle$

Proof The if direction is item (ii) of proposition 3.1. For the only if direction, let Π be a deduction of $\langle T(\text{"}A\text{"}),\ i+1\rangle$ from $\langle T(\text{"}G\text{"}),\ i+1\rangle$. Then $\overline{\Pi}^{i}$ is a deduction of $\overline{\langle T(\text{"}A\text{"}),\ i+1\rangle}^{i} = \langle A,\ i\rangle$ from $\overline{\langle T(\text{"}G\text{"}),\ i+1\rangle}^{i} = \langle G,\ i\rangle$. Q.E.D.

4.4 Moving Deductions Across Levels

In MK, the language of the theory at a certain level contains the language of all the theories of the previous levels. Furthermore all the theories have the same set of axioms and the same deductive machinery. This allows deductions to be "shifted up" in the hierarchy.

Lemma 4.6 (Shifting up deductions) *If* $\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle \vdash_{\mathrm{MK}} \langle A, i\rangle$, *then* $\langle A_1, i_1+1\rangle, \ldots, \langle A_n,\ i_n+1\rangle \vdash_{\mathrm{MK}} \langle A,\ i+1\rangle$.

where $\Gamma = \mathbf{d}_\iota^\alpha(\Gamma_1, \ldots, \Gamma_n)$.

By induction hypothesis, each $\overline{\Pi_k}^{i_0}$ is a deduction of $\overline{\langle A_k,\ i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$. Since all levels have the same i-rules, if α is an application of σ_i then $\overline{\alpha}^{i_0}$ is an application either of σ_{i_0} (if $i > i_0$) or of σ_i (if $i \leq i_0$); furthermore if an assumption $\langle B,\ i\rangle$ of Π_k is discharged by α, then $\overline{\langle B,\ i\rangle}^{i_0} \in \overline{\Gamma_k}^{i_0}$ is discharged by the application $\overline{\alpha}^{i_0}$. This entails that $\overline{\Pi}^{i_0}$ is a deduction of $\overline{\langle A,\ i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

If Π ends with an application of a $\mathcal{R}_{up.i}$ *i.e.*, it is of the form:

$$\dfrac{\begin{array}{c}\Gamma\\ \Pi'\\ \langle A,\ i\rangle\end{array}}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{R}_{up.i}$$

then by induction hypothesis $\overline{\Pi'}^{i_0}$ is a deduction of $\overline{\langle A,\ i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

If $i < i_0$, then $\overline{\Pi}^{i_0}$ is

$$\dfrac{\begin{array}{c}\overline{\Gamma}^{i_0}\\ \overline{\Pi'}^{i_0}\\ \langle A,\ i\rangle\end{array}}{\langle T(\text{“}A\text{”}),\ i+1\rangle}\ \mathcal{R}_{up.i}$$

Since $\mathcal{R}_{up.i}$ is applicable in Π, the indexes of the elements of $\overline{\Gamma}^{i_0}$ are greater than i and $i < i_0$; hence $\mathcal{R}_{up.i}$ is also applicable in $\overline{\Pi}^{i_0}$. This implies that $\overline{\Pi}^{i_0}$ is a deduction of $\overline{\langle T(\text{“}A\text{”}),\ i+1\rangle}^{i_0} = \langle T(\text{“}A\text{”}),\ i+1\rangle$ from $\overline{\Gamma}^{i_0}$.

If $i \geq i_0$, then $\overline{\Pi}^{i_0}$ is $\overline{\Pi'}^{i_0}$, that is a deduction of $\overline{\langle T(\text{“}A\text{”}),\ i+1\rangle}^{i_0} = \overline{\langle A,\ i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

If Π ends with an application of a $\mathcal{R}_{dn.i}$, then it is of the form:

$$\dfrac{\begin{array}{c}\Gamma\\ \Pi'\\ \langle T(\text{“}A\text{”}),\ i+1\rangle\end{array}}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}$$

By induction hypothesis $\overline{\Pi'}^{i_0}$ is a deduction of $\overline{\langle T(\text{“}A\text{”}),\ i+1\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

If $i < i_0$ then $\overline{\Pi}^{i_0}$ is:

$$\dfrac{\begin{array}{c}\overline{\Gamma}^{i_0}\\ \overline{\Pi'}^{i_0}\\ \langle T(\text{“}A\text{”}),\ i+1\rangle\end{array}}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}$$

that is a deduction of $\overline{\langle A,\ i\rangle}^{i_0} = \langle A,\ i\rangle$ from $\overline{\Gamma}^{i_0}$.

If $i \geq i_0$, then $\overline{\Pi}^{i_0}$ is $\overline{\Pi'}^{i_0}$, that is a deduction of $\overline{\langle A,\ i\rangle}^{i_0} = \overline{\langle T(\text{“}A\text{”}),\ i+1\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

Q.E.D.

Theorem 4.3 (Existence of a weak normal deduction) *If* $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$, *then there exists a weak normal deduction in MK of* $\langle A,\ i\rangle$ *from* Γ.

Remark 4.9 If A is an L_i-wff then $(A)^{(-n)}$ is an L_{i-n}-wff. The operator $(.)^{(-1)}$ preserves provability of its arguments but not unprovability. For instance the formula $\neg T(\text{“}A\text{”}) \supset T(\text{“}\neg A\text{”})$ is not provable in MK, however $(\neg T(\text{“}A\text{”}) \supset T(\text{“}\neg A\text{”}))^{(-1)} = \neg A \supset \neg A$ is.

The operator $(.)^{(-n)}$ allows us to define a transformation which removes all the overflowing formulas from deductions. This transformation takes a proof Π and a natural number i_0, and returns the deduction $\overline{\Pi}^{i_0}$ in which all occurrences with index greater than i_0 have been pushed down to level i_0.

Definition 4.4 *For any natural number i_0, $\overline{\langle A,\, i\rangle}^{i_0}$ is $\langle A^{(i_0-i)},\, i_0\rangle$ if $i > i_0$ and $\langle A,\, i\rangle$ otherwise. If Γ is a set of wffs, $\overline{\Gamma}^{i_0} = \{\overline{\langle A,\, i\rangle}^{i_0} : \langle A,\, i\rangle \in \Gamma\}$. If Π is a deduction, then $\overline{\Pi}^{i_0}$ is the formula tree obtained by substituting every occurrence $\langle A,\, i\rangle$ in Π with $\overline{\langle A,\, i\rangle}^{i_0}$ and by removing all the consequences of $\mathcal{R}_{up.i}$ and $\mathcal{R}_{dn.i+1}$, with $i \geq i_0$.*

Example 4.4 Let Π be the deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{\langle p,\, 0\rangle}{\langle p \supset p,\, 0\rangle}\supset\mathrm{I}_0}{\langle T(\text{“}p \supset p\text{”}),\, 1\rangle}\mathcal{R}_{up.0} \qquad \langle T(\text{“}p \supset p\text{”}) \supset T(\text{“}q\text{”}),\, 1\rangle}{\langle T(\text{“}q\text{”}),\, 1\rangle}\supset\mathrm{E}_1}{\langle q,\, 0\rangle}\mathcal{R}_{dn.0}$$

of $\langle q,\, 0\rangle$ from $\langle T(\text{“}p \supset p\text{”}) \supset T(\text{“}q\text{”}),\, 1\rangle$. $\overline{\Pi}^0$ is the deduction:

$$\dfrac{\dfrac{\langle p,\, 0\rangle}{\langle p \supset p,\, 0\rangle}\supset\mathrm{I}_0 \qquad \langle (p \supset p) \supset q,\, 0\rangle}{\langle q,\, 0\rangle}\supset\mathrm{E}_0$$

of $\overline{\langle q,\, 0\rangle}^0 = \langle q,\, 0\rangle$ from $\overline{\langle T(\text{“}p \supset p\text{”}) \supset T(\text{“}q\text{”}),\, 1\rangle}^0 = \langle (p \supset p) \supset q,\, 0\rangle$.

Lemma 4.5 *Let Π be a deduction of $\langle A,\, i\rangle$ from Γ. Let i_0 be any natural number. Then $\overline{\Pi}^{i_0}$ is a deduction of $\overline{\langle A,\, i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.*

Proof We prove by induction that if Π is a deduction of $\langle A,\, i\rangle$ depending on Γ then $\overline{\Pi}^{i_0}$ is a deduction of $\overline{\langle A,\, i\rangle}^{i_0}$ from $\overline{\Gamma}^{i_0}$.

[base case] If Π is $\langle A,\, i\rangle$ then $\overline{\langle A,\, i\rangle}^{i_0}$ is a deduction of $\overline{\langle A,\, i\rangle}^{i_0}$ from $\overline{\langle A,\, i\rangle}^{i_0}$.

[step case] If Π ends with an application $\alpha = \langle\langle A_1,\, i\rangle, \ldots, \langle A_n,\, i\rangle, \langle A,\, i\rangle\rangle$ of an i-rule, σ_i, then it is of the form:

$$\dfrac{\begin{matrix}\Gamma_1 \\ \Pi_1 \\ \langle A_1,\, i\rangle\end{matrix} \quad \cdots \quad \begin{matrix}\Gamma_n \\ \Pi_n \\ \langle A_n,\, i\rangle\end{matrix}}{\langle A,\, i\rangle}\sigma_i$$

4.3 The Sublevel Property

In MK the meta-theories do not contain any specific axioms about their object theories. It seems therefore that the only way to derive, in a meta-theory, a statement about its object theory is by reflecting up from the object theory itself. If this is the case, then, in order to prove a theorem inside a theory, it becomes useless to pass through its meta-theory. In the following we formally prove the general fact that a theorem at the level i can be proved without going to the levels above i.

Definition 4.2 (Overflowing Formula, Weak Normal Form) *If Π is a deduction of $\langle A, i\rangle$ from Γ and i_0 is the greatest index of the wffs in $\Gamma \cup \{\langle A, i\rangle\}$, then an occurrence $\langle B, j\rangle$ in Π is an* overflowing formula *of Π, if and only if $j > i_0$. Π is a* weak normal deduction *(is in* weak normal form*) if and only if it does not contain any overflowing formula.*

Example 4.2 In the deduction:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle p,\ 0\rangle}{\langle p \supset p,\ 0\rangle}\ {\supset}\mathrm{I}_0}{\langle T(\text{“}p \supset p\text{”}),\ 1\rangle}\ \mathcal{R}_{up.0}}{\langle T(\text{“}T(\text{“}p \supset p\text{”})\text{”}),\ 2\rangle^{\bullet}}\ \mathcal{R}_{up.1}}{\langle T(\text{“}p \supset p\text{”}),\ 1\rangle}\ \mathcal{R}_{dn.1} \qquad \langle T(\text{“}p \supset p\text{”}) \supset T(\text{“}q\text{”}),\ 1\rangle}{\langle T(\text{“}q\text{”}),\ 1\rangle}\ {\supset}\mathrm{E}_1}{\langle q,\ 0\rangle}\ \mathcal{R}_{dn.0}$$

the occurrence flagged with the bullet is an overflowing formula.

Let us define an operator $(.)^{(-1)}$ on wffs. Intuitively $(.)^{(-1)}$ transforms an L_{i+1}-wff into an L_i-wff by removing (if there exist) the “most external” occurrences of the predicate T.

Definition 4.3

(i) $\bot^{(-1)} = \bot$;

(ii) $p^{(-1)} = p$, *if p is a propositional constant;*

(iii) $(.)^{(-1)}$ *distributes over connectives;*

(iv) $T(\text{“}A\text{”})^{(-1)} = A$.

Furthermore $A^{(0)} = A$, and for each natural number n, $(A)^{(-n)} = ((A)^{(-(n-1))})^{(-1)}$.

Example 4.3 Some examples of applications of $(.)^{(-n)}$:

$$\begin{array}{rcl}
p^{(-1)} & = & p \\
(T(\text{“}p \wedge q\text{”}) \vee r)^{(-1)} & = & (p \wedge q) \vee r \\
(p \wedge T(\text{“}q\text{”}) \wedge T(\text{“}T(\text{“}r\text{”})\text{”}))^{(-1)} & = & p \wedge q \wedge T(\text{“}r\text{”}) \\
(p \wedge T(\text{“}q\text{”}) \wedge T(\text{“}T(\text{“}r\text{”})\text{”}))^{(-2)} & = & p \wedge q \wedge r
\end{array}$$

Theorem 4.1 $\vdash_{\text{MK}}$ *is an ML-derivability relation.*

Proof We prove that $\vdash_{\text{MK}}$ satisfies all the properties of an ML-derivability relation (definition 2.11, page 11).

(MR–RX) $\langle A, i\rangle \vdash_{\text{MK}} \langle A, i\rangle$.
(MR–LM) if $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$ then $\Gamma, \Sigma \vdash_{\text{MK}} \langle A, i\rangle$.
(MR–CUT) if $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$ and $\Sigma, \langle A, i\rangle \vdash_{\text{MK}} \langle B, j\rangle$then $\Gamma, \Sigma \vdash_{\text{MK}} \langle B, j\rangle$.

(MR–RX): By definition 2.7, for each L_i-wff A, $\langle A, i\rangle$ is a deduction of $\langle A, i\rangle$ depending on $\langle A, i\rangle$, *i.e.* $\langle A, i\rangle \vdash_{\text{MK}} \langle A, i\rangle$.

(MR–LM): By definition 2.7, if $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$ then there exists a deduction of $\langle A, i\rangle$ depending on Γ' with, $\Gamma' \subseteq \Gamma \subseteq \Gamma \cup \Sigma$. This implies that $\Gamma, \Sigma \vdash_{\text{MK}} \langle A, i\rangle$.

(MR–CUT): We have two cases: $i < j$ and $i \geq j$. If $i < j$, by lemma 4.2 $\langle A, i\rangle$ is not an effective assumption and we have that $\Sigma \vdash_{\text{MK}} \langle B, j\rangle$, and by monotonicity, that $\Gamma, \Sigma \vdash_{\text{MK}} \langle B, j\rangle$. If $i \geq j$, let Π_1 be a deduction of $\langle A, i\rangle$ depending on $\Gamma' \subseteq \Gamma$ and Π_2 a deduction of $\langle B, j\rangle$ depending on $\Sigma', \langle A, i\rangle$, with $\Sigma' \subseteq \Sigma$. By lemma 4.3:

$$\begin{array}{c}\Pi \\ (\langle A, i\rangle) \\ \Pi'\end{array}$$

is a deduction of $\langle B, j\rangle$ from $\Gamma' \cup \Sigma'$. By (MR–LM) we have $\Gamma, \Sigma \vdash_{\text{MK}} \langle B, j\rangle$. Q.E.D.

Remark 4.7 (MR–CUT) is stronger than (ML–CUT) (see definition 2.11 on page 11). In (ML–CUT) the indexes of the "cut" formula ($\langle A, i\rangle$) and that of the conclusion are required to be equal, while (MR–CUT) holds even if the two indexes do not coincide.

Theorem 4.2 (Deduction theorem) $\Gamma, \langle A, i\rangle \vdash_{\text{MK}} \langle B, i\rangle$ *if and only if* $\Gamma \vdash_{\text{MK}} \langle A \supset B, i\rangle$

Proof If Π is a deduction of $\langle B, i\rangle$ from $\Gamma, \langle A, i\rangle$, then

$$\dfrac{\begin{array}{c}\Gamma, \langle A, i\rangle \\ \Pi \\ \langle B, i\rangle\end{array}}{\langle A \supset B, i\rangle} \supset\! \mathrm{I}_i$$

is a deduction of $\langle A \supset B, i\rangle$ from Γ. Viceversa, since $\langle A, i\rangle, \langle A \supset B, i\rangle \vdash_{\text{MK}} \langle B, i\rangle$, if $\Gamma \vdash_{\text{MK}} \langle A \supset B, i\rangle$, by (MR–CUT) we have $\Gamma, \langle A, i\rangle \vdash_{\text{MK}} \langle B, i\rangle$. Q.E.D.

Remark 4.8 When the assumption A and the conclusion B do not belong to the same language i the deduction theorem does not hold, as there exists no language inside which to express the implication of A and B.

Lemma 4.4 (Compactness) $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$ *if and only if there exists a finite set* $\Gamma' \subseteq \Gamma$ *such that* $\Gamma' \vdash_{\text{MK}} \langle A, i\rangle$.

Proof By definition 2.7, if $\Gamma \vdash_{\text{MK}} \langle A, i\rangle$ then there exists deduction of $\langle A, i\rangle$ depending on $\Gamma' \subseteq \Gamma$. From definition 2.6 it is straightforward to prove that Γ' is finite. The other direction is trivial. Q.E.D.

Lemma 4.2 (Effective assumptions) $\Gamma \vdash_{\text{MK}} \langle A,\, i\rangle$ *if and only if* $\Gamma\lfloor_i \vdash_{\text{MK}} \langle A,\, i\rangle$.

Proof Let $\tau = \langle A_1,\, i_1\rangle \ldots \langle A_n,\, i_n\rangle$ be any thread from an undischarged assumption $\langle A_1,\, i_1\rangle$ to the bottom formula $\langle A_n,\, i_n\rangle = \langle A,\, i\rangle$. Since Π is a deduction τ is open, which implies that $i_1 \geq i$, hence $\langle A_1,\, i_1\rangle \in \Gamma\lfloor_i$. Q.E.D.

Remark 4.6 The intuitive interpretation of lemma 4.2 is that an assumption at one level doesn't have any effects on the upper levels. We call $\Gamma\lfloor_i$ the subset of *effective assumptions* at level i of Γ.

Notation 4.3 If Π is a deduction of $\langle A,\, i\rangle$ and Π' is any deduction then,

$$\begin{array}{c}\Pi\\ (\langle A,\, i\rangle)\\ \Pi'\end{array} \tag{12}$$

is Π' itself if Π' does not contain any undischarged assumptions of the form $\langle A,\, i\rangle$, or is the quasi-deduction obtained by writing Π, without its bottom formula $\langle A,\, i\rangle$, above *all* the undischarged assumptions of Π' of the from $\langle A,\, i\rangle$. Furthermore if $\langle A,\, i\rangle$ is any top formula of Π' then

$$\begin{array}{c}\Pi\\ \{\langle A,\, i\rangle\}\\ \Pi'\end{array} \tag{13}$$

represents the quasi-deduction obtained by writing Π, without its bottom formula $\langle A, i\rangle$, above *the top formula* $\langle A,\, i\rangle$.

Lemma 4.3 (Composing deductions) *If* Π *is a deduction of* $\langle A,\, i\rangle$ *depending on* Σ *and* Π' *is a deduction of* $\langle B,\, j\rangle$ *depending on* Γ*, then* $\Pi/(\langle A,\, i\rangle)/\Pi'$ *is a deduction of* $\langle B,\, j\rangle$ *from* $(\Gamma \perp \{\langle A,\, i\rangle\}) \cup \Sigma$.

Proof $\Pi/(\langle A,\, i\rangle)/\Pi'$ is obviously a quasi-deduction. We show that it is a deduction by showing that each of its threads is either open or closed.

If a thread τ of $\Pi/(\langle A, i\rangle)/\Pi'$ does not contain the occurrence $\langle A, i\rangle$, then it is necessarily a thread of Π'. Since Π' is a deduction, τ is open or closed.

A thread that contains the occurrence $\langle A,\, i\rangle$ is of the form $\tau\langle A,\, i\rangle\tau'$. $\tau\langle A,\, i\rangle$ is a thread of Π and it is either open or closed. $\langle A,\, i\rangle\tau'$ is a thread of Π' and it is open since $\langle A,\, i\rangle$ is undischarged. If τ is closed then $\tau\langle A,\, i\rangle\tau'$ is also closed. If $\tau\langle A,\, i\rangle$ is open then the index of each occurrence of τ is lower than or equal to that of the first formula, *e.g.* the assumption $\langle A_1,\, i_1\rangle$ of $\tau\langle A,\, i\rangle$. Since $\langle A,\, i\rangle$ is undischarged in Π, $\langle A,\, i\rangle\tau$ is open and the index of each occurrence of $\langle A,\, i\rangle\tau$ is lower than or equal to i. We conclude that the index of each occurrence of $\tau\langle A,\, i\rangle\tau'$ is lower than or equal to i_1. If $\langle A_1,\, i_1\rangle$ is discharged [not discharged] in $\langle A,\, i\rangle\tau$, then we have that $\tau\langle A,\, i\rangle\tau'$ is closed [open]. This ensures that $\Pi/(\langle A,\, i\rangle)/\Pi'$ is a deduction.

The undischarged assumptions of $\Pi/(\langle A,\, i\rangle)/\Pi'$ are contained in the set of undischarged assumptions of Π' except for $\langle A,\, i\rangle$, union the undischarged assumptions of Π. This implies that $\Pi/(\langle A,\, i\rangle)/\Pi'$ is a deduction of $\langle B,\, i\rangle$ from $(\Gamma \perp \{\langle A,\, i\rangle\}) \cup \Sigma$. Q.E.D.

Remark 4.5 In order to prove lemma 4.1, it is important to notice that, if a thread τ of a quasi-deduction Π is closed, then any thread $\tau\tau'$ of a quasi-deduction Π' containing Π as sub-tree is also closed. Furthermore if τ is an open thread of Π and Π' is the deduction

$$\frac{\cdots \quad \Pi \quad \cdots}{\langle A,\, i\rangle}\ \iota$$

where ι discharges the first formula of τ, then the thread $\tau\langle A, i\rangle$ of Π' is closed. Viceversa, if $\tau\langle A,\, i\rangle$ is open then τ is open, and if $\tau\langle A,\, i\rangle$ is closed then τ is either open or closed.

Proof (of lemma 4.1) Let Π be a deduction. We prove by induction on Π that each of its threads $\tau = \langle A_1,\, i_1\rangle \ldots \langle A_n,\, i_n\rangle$ is either closed or open.

Let us suppose that $\langle A_1,\, i_1\rangle$ is discharged at some formula occurrence $\langle A_h,\, i_h\rangle$ of τ, with $1 \leq h \leq n$. Let Π' be the sub-deduction of Π with top formula $\langle A_{h-1},\, i_{h-1}\rangle$ and τ' the initial part of τ contained in Π', *i.e.* the thread of Π' such that $\tau = \tau'\langle A_h, i_h\rangle \ldots \langle A_n, i_n\rangle$. Since Π' is a deduction, by induction τ' is open. Hence, by remark 4.5, τ is closed.

Let us suppose that $\langle A_1,\, i_1\rangle$ is undischarged in Π. By contradiction let $\langle B,\, i_1+1\rangle$ be the first occurrence of τ with index equal to i_i+1. $\langle B,\, i_1+1\rangle$ is necessarily the consequence of an application of an $\mathcal{R}_{up.i_1}$. But this is impossible for the restriction of $\mathcal{R}_{up.i_1}$. This means that for each $1 \leq k \leq n$, $i_1 \geq i_k$, *i.e.* τ is open.

Viceversa, let Π be a quasi-deduction whose threads are either closed or open. We prove by induction that Π is a deduction.

[**base case**] If Π is $\langle A,\, i\rangle$, then it is a deduction.

[**step case**] Let Π end with an application of the rule ι, *i.e.* it is of the form:

$$\frac{\Pi_1 \quad \cdots \quad \Pi_n}{\langle A,\, i\rangle}\ \iota$$

In order to prove that Π is a deduction we show that (a) each Π_k is a deduction and that (b) ι is applicable.

Let $\tau = \langle A_1, i_1\rangle \ldots \langle A_n, i_n\rangle$ be a thread of Π_k. (a) $\tau\langle A, i\rangle$ is a thread of Π. By hypothesis it is closed or open. If $\tau\langle A,\, i\rangle$ is closed, then by remark 4.5, τ is either open or closed. If $\tau\langle A,\, i\rangle$ is open then τ is also open. Hence for the induction hypothesis, each Π_k is a deduction.

If ι is different from $\mathcal{R}_{up.}$, then, since it does not have any restriction, Π is a deduction. If ι is $\mathcal{R}_{up.i-1}$ then $\langle A,\, i\rangle$ is of the form $\langle T(\text{“}B\text{”}),\, i\rangle$. Let τ be the thread from an undischarged assumption $\langle A_1,\, i_1\rangle$ to $\langle T(\text{“}B\text{”}),\, i\rangle$. τ is open by hypothesis and this implies that $i_1 > i \perp 1$, which ensures the applicability of $\mathcal{R}_{up.i-1}$. Q.E.D.

Given a set of wffs Γ, for each level i we can extract three subsets of Γ: the subset of wffs with index less then or equal to i, written $\Gamma\lceil^i$, the subset of wffs with index greater than or equal to i, written $\Gamma\lfloor_i$, and the set of wffs with index equal to i, which is the intersection of the previous two sets.

Notation 4.2 Let Γ be a set of wffs. Then $\Gamma\lfloor_i = \{\langle B, j\rangle \in \Gamma : j \geq i\}$ and $\Gamma\lceil^i = \{\langle B, j\rangle \in \Gamma : j \leq i\}$. Notice that for each i, $\Gamma = \Gamma\lfloor_i \cup \Gamma\lceil^i$ and $\Gamma\lfloor_i \cap \Gamma\lceil^i = (\Gamma\lfloor_i)\lceil^i = (\Gamma\lceil^i)\lfloor_i = \{\langle B, i\rangle \in \Gamma\}$.

Definition 4.1 *A* quasi-deduction *is defined recursively as follows:*

(i) $\langle A,\ i\rangle$ *is a quasi-deduction;*

(ii) *if* $\Pi_1, \ldots, \Pi_n$ *are quasi-deductions whose end formulas are* $\langle A_1,\ i_1\rangle, \ldots, \langle A,\ n\rangle$, *then*

$$\frac{\Pi_1 \quad \ldots \quad \Pi_n}{\langle A,\ i\rangle}\ \iota$$

is a quasi-deduction, if and only if $\langle\langle A_1,\ i_1\rangle, \ldots, \langle A,\ n\rangle, \langle A,\ i\rangle\rangle \in \rho(\iota)$.

Example 4.1 Consider the following formula trees:

$$\dfrac{\dfrac{\dfrac{\langle T(\text{“}B\text{”}), i+1\rangle}{\langle B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A\wedge B,\ i\rangle}\ \wedge\mathrm{I}_i}{\langle T(\text{“}A\wedge B\text{”}), i+1\rangle}\ \mathcal{R}_{up.i} \qquad \dfrac{\dfrac{\langle A,\ i\rangle \quad \dfrac{\langle T(\text{“}B\text{”}), i+1\rangle}{\langle B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A\wedge B,\ i\rangle}\ \wedge\mathrm{I}_i}{\langle T(\text{“}A\wedge B\text{”}), i+1\rangle}\ \mathcal{R}_{up.i} \qquad \dfrac{\dfrac{\dfrac{\langle T(\text{“}A\text{”}), i+1\rangle}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i} \quad \dfrac{\langle T(\text{“}B\text{”}), i+1\rangle}{\langle B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A\wedge B,\ i\rangle}\ \wedge\mathrm{I}_i}{\langle T(\text{“}A\wedge B\text{”}), i+1\rangle}\ \mathcal{R}_{up.i}$$

(11.a) (11.b) (11.c)

The formula tree (11.a) is not a quasi-deduction, since $\langle\langle B,\ i\rangle, \langle A \wedge B,\ i\rangle\rangle$ is not an application of an $\wedge\mathrm{I}_i$. The formula tree (11.b) is a quasi-deduction but not a deduction, as the fact that $\langle A\wedge B, i\rangle$ depends on $\langle A,\ i\rangle$ forbids the application of $\mathcal{R}_{up.i}$. The formula tree (11.c) is a deduction.

Let us define threads and undischarged assumptions inside quasi-deductions the same as for deductions. The following lemma states a necessary and sufficient condition for a quasi-deduction to be a deduction.

Lemma 4.1 *A quasi-deduction* Π *is a deduction in MK, if and only if, for each thread* $\tau = \langle A_1,\ i_1\rangle, \ldots, \langle A_n,\ i_n\rangle$,

(i) τ *is* closed*: i.e.* $\langle A_1,\ i_1\rangle$ *is discharged at* $\langle A_h,\ i_h\rangle$ *and for each* $1 \leq k < h$, $i_1 \geq i_k$, *or*

(ii) τ *is* open*: i.e.* $\langle A_1,\ i_1\rangle$ *is undischarged and for each* $1 \leq k \leq n$, $i_1 \geq i_k$.

Remark 4.3 Intuitively lemma 4.1 states that, starting from an assumption and following a thread, the indexes of the formulas do not go above that of the assumption, until the assumption has been discharged. The proof is based on the fact that $\mathcal{R}_{up.}$ is the only rule which increases indexes and on the fact that the restriction on the applicability of $\mathcal{R}_{up.}$ prevents the conclusion to be at a higher level than that of any undischarged assumption.

Remark 4.4 In general a quasi-deduction contains threads which are neither open nor closed. For example the thread $\langle A,\ i\rangle\langle A \wedge B,\ i\rangle\langle T(\text{“}A \wedge B\text{”}),\ i+1\rangle$ of deduction (11.b) in the previous example, is neither open nor closed.

Notation 4.1 For any thread $\tau_1 = \langle A_1,\ i_1\rangle \ldots \langle A_n,\ i_n\rangle$, $\tau_2 = \langle B_1,\ j_1\rangle \ldots \langle B_m,\ j_m\rangle$, the notation $\tau_1\tau_2$ stands for the thread $\langle A_1,\ i_1\rangle \ldots \langle A_n,\ i_n\rangle\langle B_1,\ j_1\rangle \ldots \langle B_m,\ j_m\rangle$.

(ii)

$$\dfrac{\dfrac{\dfrac{\langle T(\text{“}A\text{”})\vee T(\text{“}B\text{”}), i+1\rangle \quad \dfrac{\dfrac{\dfrac{\langle T(\text{“}A\text{”}), i+1\rangle}{\langle A, i\rangle}\,\mathcal{R}_{dn.i}}{\langle A\vee B, i\rangle}\,\vee\mathrm{I}_i}{\langle T(\text{“}A\vee B\text{”}), i+1\rangle}\,\mathcal{R}_{up.i} \quad \dfrac{\dfrac{\dfrac{\langle T(\text{“}B\text{”}), i+1\rangle}{\langle B, i\rangle}\,\mathcal{R}_{dn.i}}{\langle A\vee B, i\rangle}\,\vee\mathrm{I}_i}{\langle T(\text{“}A\vee B\text{”}), i+1\rangle}\,\mathcal{R}_{up.i}}{\langle T(\text{“}A\vee B\text{”}), i+1\rangle}\,\vee\mathrm{E}_{i+1}}{\langle A\vee B, i\rangle}\,\mathcal{R}_{dn.i} \quad \dfrac{\langle A, i\rangle \quad \langle A\supset C, i\rangle}{\langle C, i\rangle}\,\supset\mathrm{E}_i \quad \dfrac{\langle B, i\rangle \quad \langle B\supset C, i\rangle}{\langle C, i\rangle}\,\supset\mathrm{E}_i}{\langle C, i\rangle}\,\vee\mathrm{E}_i$$

(iii)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle\neg T(\text{“}A\text{”}), i+1\rangle \quad \langle\neg T(\text{“}A\text{”})\supset T(\text{“}B\text{”}), i+1\rangle}{\langle T(\text{“}B\text{”}), i+1\rangle}\,\supset\mathrm{E}_{i+1}}{\langle B, i\rangle}\,\mathcal{R}_{dn.i}}{\langle\neg A\supset B, i\rangle}\,\supset\mathrm{I}_i}{\langle T(\text{“}\neg A\supset B\text{”}), i+1\rangle}\,\mathcal{R}_{up.i} \quad \langle\neg T(\text{“}\neg A\supset B\text{”}), i+1\rangle}{\langle\bot, i+1\rangle}\,\supset\mathrm{E}_{i+1}}{\langle T(\text{“}A\text{”}), i+1\rangle}\,\bot_{i+1}}{\langle A, i\rangle}\,\mathcal{R}_{dn.I} \quad \langle\neg A, i\rangle}{\langle\bot, i\rangle}\,\supset\mathrm{E}_1}{\langle B, i\rangle}\,\bot_i}{\langle\neg A\supset B, i\rangle}\,\supset\mathrm{I}_i}{\langle T(\text{“}\neg A\supset B\text{”}), i+1\rangle}\,\mathcal{R}_{up.i} \quad \langle\neg T(\text{“}\neg A\supset B\text{”}), i+1\rangle}{\langle\bot, i+1\rangle}\,\supset\mathrm{E}_{i+1}}{\langle T(\text{“}\neg A\supset B\text{”}), i+1\rangle}\,\bot_{i+1}}{\langle\neg A\supset B, i\rangle}\,\mathcal{R}_{dn.i} \quad \langle\neg A, i\rangle}{\langle B, i\rangle}\,\supset\mathrm{E}_i$$

Q.E.D.

4.2 Some Basic Proof Theory

We are now ready to develop some elementary proof theory. In order to do so, we need to introduce a number of operations for composing, decomposing and, in general, handling deductions. In general, the result of any transformation on a deduction is a formula tree. We give a necessary and sufficient condition for this formula tree to be a deduction.

Notice that, for a formula tree to be a deduction the following two conditions must be satisfied:

- each node (which is not a leaf) of the formula tree must be the consequence of an application of a rule in MK, whose premises are the nodes occurring immediately above that node;
- all the rule applications must be allowed.

Formula trees that satisfy the first condition are called *quasi-deductions*. Similarly to deductions, quasi-deductions are generated starting from assumptions and axioms and by applying inference rules. However restrictions on the applicability of inference rules are not considered.

(vi)

$$\dfrac{\dfrac{\dfrac{\langle \neg T(\text{“}\bot\text{”}),\ i+1\rangle \qquad \dfrac{\dfrac{\dfrac{\langle T(\text{“}A\text{”}),\ i+1\rangle}{\langle A,\ i\rangle}\mathcal{R}_{dn.i} \quad \dfrac{\langle T(\text{“}\neg A\text{”}),\ i+1\rangle}{\langle \neg A,\ i\rangle}\mathcal{R}_{dn.i}}{\langle \bot,\ i\rangle}\supset\mathrm{E}_i}{\langle T(\text{“}\bot\text{”}),\ i+1\rangle}\mathcal{R}_{up.i}}{\dfrac{\langle \bot,\ i+1\rangle}{\langle \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\bot_{i+1}}\supset\mathrm{E}_{i+1}}{\langle T(\text{“}A\text{”}) \supset \neg T(\text{“}\neg A\text{”}),\ i+1\rangle}\supset\mathrm{I}_{i+1}}{\langle \neg T(\text{“}\bot\text{”}) \supset (T(\text{“}A\text{”}) \supset \neg T(\text{“}\neg A\text{”})),\ i+1\rangle}\supset\mathrm{I}_{i+1}$$

Q.E.D.

Remark 4.2 The intuition behind the above proofs is that the meta-theorems are proved by “running” the deductive machinery of the object theory. Notice that the meta-theory does not have any specific axioms about the object theory.

Exercise 4.1 *Let A, B be L_i-wffs. Then the following are theorems of MK:*

(i) $\langle T(\text{“}A\text{”}) \supset T(\text{“}B \supset A\text{”}),\ i+1\rangle$
(ii) $\langle T(\text{“}\neg A\text{”}) \supset T(\text{“}A \supset B\text{”}),\ i+1\rangle$
(iii) $\langle T(\text{“}A \supset B\text{”}) \supset (\neg T(\text{“}\neg A\text{”}) \supset \neg T(\text{“}\neg B\text{”})),\ i+1\rangle$
(iv) $\langle (T(\text{“}A\text{”}) \wedge \neg T(\text{“}B\text{”})) \supset \neg T(\text{“}\neg(A \wedge \neg B)\text{”}),\ i+1\rangle$
(v) $\langle T(\text{“}A \vee B\text{”}) \supset (\neg T(\text{“}\neg A\text{”}) \vee T(\text{“}B\text{”})),\ i+1\rangle$
(vi) $\langle \neg T(\text{“}\neg(A \supset B)\text{”}) \vee T(\text{“}B \supset A\text{”}),\ i+1\rangle$
(vii) $\langle \neg T(\text{“}\neg(A \supset B)\text{”}) \equiv (T(\text{“}A\text{”}) \supset \neg T(\text{“}\neg B\text{”})),\ i+1\rangle$
(viii) $\langle (\neg T(\text{“}\neg A\text{”}) \supset T(\text{“}B\text{”})) \supset T(\text{“}A \supset B\text{”}),\ i+1\rangle$
(ix) $\langle (\neg T(\text{“}\neg A\text{”}) \supset T(\text{“}B\text{”})) \supset (\neg T(\text{“}\neg A\text{”}) \supset \neg T(\text{“}\neg B\text{”})),\ i+1\rangle$
(x) $\langle T(\text{“}A \supset B\text{”}) \supset ((\neg T(\text{“}B\text{”}) \supset T(\text{“}A\text{”})) \supset T(\text{“}B\text{”})),\ i+1\rangle$

So far we have proved properties about the MK provability relation. Let us now consider some properties of the MK derivability relation.

Proposition 4.2 *Let A, B and C be L_i-wffs. Then:*

(i) $\langle T(\text{“}A\text{”}) \vee T(\text{“}B\text{”}),\ i+1\rangle, \langle T(\text{“}A \supset C\text{”}),\ i+1\rangle, \langle T(\text{“}B \supset C\text{”}),\ i+1\rangle \vdash_{\mathrm{MK}} \langle C,\ i\rangle$
(ii) $\langle T(\text{“}A\text{”}) \vee T(\text{“}B\text{”}),\ i+1\rangle, \langle A \supset C,\ i\rangle, \langle B \supset C,\ i\rangle \vdash_{\mathrm{MK}} \langle C,\ i\rangle$
(iii) $\langle \neg T(\text{“}A\text{”}) \supset T(\text{“}B\text{”}),\ i+1\rangle, \langle \neg A,\ i\rangle \vdash_{\mathrm{MK}} \langle B,\ i\rangle$

Proof
(i)

$$\dfrac{\dfrac{\langle T(\text{“}A\text{”}) \vee T(\text{“}B\text{”}),\ i+1\rangle \quad \dfrac{\dfrac{\dfrac{\langle T(\text{“}A\text{”}),\ i+1\rangle}{\langle A,\ i\rangle}\mathcal{R}_{dn.i} \quad \dfrac{\langle T(\text{“}A \supset C\text{”}),\ i+1\rangle}{\langle A \supset C,\ i\rangle}\mathcal{R}_{dn.i}}{\langle C,\ i\rangle}\supset\mathrm{E}_i}{\langle T(\text{“}C\text{”}),\ i+1\rangle}\mathcal{R}_{up.i} \quad \dfrac{\dfrac{\dfrac{\langle T(\text{“}B\text{”}),\ i+1\rangle}{\langle B,\ i\rangle}\mathcal{R}_{dn.i} \quad \dfrac{\langle T(\text{“}B \supset C\text{”}),\ i+1\rangle}{\langle B \supset C,\ i\rangle}\mathcal{R}_{dn.i}}{\langle C,\ i\rangle}\supset\mathrm{E}_i}{\langle T(\text{“}C\text{”}),\ i+1\rangle}\mathcal{R}_{up.i}}{\dfrac{\langle T(\text{“}C\text{”}),\ i+1\rangle}{\langle C,\ i\rangle}\mathcal{R}_{dn.i}}\vee\mathrm{E}_{i+1}$$

(ii)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle T(``A \wedge B"),\ i+1\rangle}{\langle A \wedge B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A,\ i\rangle}\ \wedge\mathrm{E}_i}{\langle T(``A"),\ i+1\rangle}\ \mathcal{R}_{up.i} \qquad \dfrac{\dfrac{\dfrac{\langle T(``A \wedge B"),\ i+1\rangle}{\langle A \wedge B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle B,\ i\rangle}\ \wedge\mathrm{E}_i}{\langle T(``B"),\ i+1\rangle}\ \mathcal{R}_{up.i}}{\langle T(``A") \wedge T(``B"),\ i+1\rangle}\ \wedge\mathrm{I}_{i+1}}{\langle T(``A \wedge B") \supset T(``A") \wedge T(``B"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle T(``A") \wedge T(``B"),\ i+1\rangle}{\langle T(``A"),\ i+1\rangle}\ \wedge\mathrm{E}_{i+1}}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i} \qquad \dfrac{\dfrac{\langle T(``A") \wedge T(``B"),\ i+1\rangle}{\langle T(``B"),\ i+1\rangle}\ \wedge\mathrm{E}_{i+1}}{\langle B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A \wedge B,\ i\rangle}\ \wedge\mathrm{I}_i}{\langle T(``A \wedge B"),\ i+1\rangle}\ \mathcal{R}_{up.i}}{\langle T(``A") \wedge T(``B") \supset T(``A \wedge B"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$$

(iii)

$$\dfrac{\dfrac{\langle T(``A") \vee T(``B"),\ i+1\rangle \qquad \dfrac{\dfrac{\dfrac{\langle T(``A"),\ i+1\rangle}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A \vee B,\ i\rangle}\ \vee\mathrm{I}_i}{\langle T(``A \vee B"),\ i+1\rangle}\ \mathcal{R}_{up.i} \qquad \dfrac{\dfrac{\dfrac{\langle T(``B"),\ i+1\rangle}{\langle B,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A \vee B,\ i\rangle}\ \vee\mathrm{I}_i}{\langle T(``A \vee B"),\ i+1\rangle}\ \mathcal{R}_{up.i}}{\langle T(``A \vee B"),\ i+1\rangle}\ \vee\mathrm{E}_{i+1}}{\langle T(``A") \vee T(``B") \supset T(``A \vee B"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$$

(iv)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\langle T(``A \vee B"),\ i+1\rangle}{\langle A \vee B,\ i\rangle}\ \mathcal{R}_{dn.i} \qquad \dfrac{\langle A,\ i\rangle \quad \dfrac{\dfrac{\langle T(``A \supset C") \wedge T(``B \supset C"),\ i+1\rangle}{\langle T(``A \supset C"),\ i+1\rangle}\ \wedge\mathrm{E}_{i+1}}{\langle A \supset C,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle C,\ i\rangle}\ \supset\mathrm{E}_i \qquad \dfrac{\langle B,\ i\rangle \quad \dfrac{\dfrac{\langle T(``A \supset C") \wedge T(``B \supset C"),\ i+1\rangle}{\langle T(``A \supset C"),\ i+1\rangle}\ \wedge\mathrm{E}_{i+1}}{\langle B \supset C,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle C,\ i\rangle}\ \supset\mathrm{E}_i}{\langle C,\ i\rangle}\ \vee\mathrm{E}_i}{\langle T(``C"),\ i+1\rangle}\ \mathcal{R}_{up.i}}{\langle T(``A \supset C") \wedge T(``B \supset C") \supset T(``C"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}}{\langle T(``A \vee B") \supset (T(``A \supset C") \wedge T(``B \supset C") \supset T(``C")),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$$

(v)

$$\dfrac{\dfrac{\dfrac{\dfrac{\langle T(``\bot"),\ i+1\rangle}{\langle \bot,\ i\rangle}\ \mathcal{R}_{dn.i}}{\langle A,\ i\rangle}\ \bot_i}{\langle T(``A"),\ i+1\rangle}\ \mathcal{R}_{up.i}}{\langle T(``\bot") \supset T(``A"),\ i+1\rangle}\ \supset\mathrm{I}_{i+1}$$

$\mathcal{R}_{dn.}$:

$$\frac{\begin{array}{c}[\langle A,\ i\rangle]\\ \langle B,\ i\rangle\end{array}}{\langle A \supset B,\ i\rangle}\supset\mathrm{I}_i \qquad \frac{\langle A,\ i\rangle \quad \langle A \supset B,\ i\rangle}{\langle B,\ i\rangle}\supset\mathrm{E}_i$$

$$\frac{\langle A,\ i\rangle \quad \langle B,\ i\rangle}{\langle A \wedge B,\ i\rangle}\wedge\mathrm{I}_i \qquad \frac{\langle A \wedge B,\ i\rangle}{\langle A,\ i\rangle}\ \frac{\langle A \wedge B,\ i\rangle}{\langle B,\ i\rangle}\wedge\mathrm{E}_i$$

$$\frac{\langle A,\ i\rangle}{\langle A \vee B,\ i\rangle}\ \frac{\langle B,\ i\rangle}{\langle A \vee B,\ i\rangle}\vee\mathrm{I}_i \quad \frac{\langle A \vee B,\ i\rangle \quad \begin{array}{c}[\langle A,\ i\rangle]\\ \langle C,\ i\rangle\end{array} \quad \begin{array}{c}[\langle B,\ i\rangle]\\ \langle C,\ i\rangle\end{array}}{\langle C,\ i\rangle}\vee\mathrm{E}_i$$

$$\frac{\begin{array}{c}[\langle \neg A,\ i\rangle]\\ \langle \bot,\ i\rangle\end{array}}{\langle A,\ i\rangle}\bot_i$$

$$\frac{\langle A,\ i\rangle}{\langle T(``A"),\ i+1\rangle}\mathcal{R}_{up.i} \qquad \frac{\langle T(``A"),\ i+1\rangle}{\langle A,\ i\rangle}\mathcal{R}_{dn.i}$$

Restrictions: $\mathcal{R}_{up.i}$ can be applied only if the index of every undischarged assumption $\langle A,\ i\rangle$ depends on, is strictly greater than i. $\bot_i$ can be applied if A is not of the form $\bot$.

4.1 Some Basic Properties

Proposition 4.1 *Let A, B and C be any L_i-wffs. Then* (i)–(vi) *are theorems of MK.*

(i) $\langle T(``A \supset B") \supset (T(``A") \supset T(``B")),\ i+1\rangle$
(ii) $\langle T(``A \wedge B") \equiv T(``A") \wedge T(``B"),\ i+1\rangle$
(iii) $\langle T(``A") \vee T(``B") \supset T(``A \vee B"),\ i+1\rangle$
(iv) $\langle T(``A \vee B") \supset (T(``A \supset C") \wedge T(``B \supset C") \supset T(``C")),\ i+1\rangle$
(v) $\langle T(``\bot") \supset T(``A"),\ i+1\rangle$
(vi) $\langle \neg T(``\bot") \supset (T(``A") \supset \neg T(``\neg A")),\ i+1\rangle$

Proof

(i)

$$\frac{\frac{\frac{\frac{\langle T(``A \supset B"),\ i+1\rangle}{\langle A \supset B,\ i\rangle}\mathcal{R}_{dn.i} \quad \frac{\langle T(``A"),\ i+1\rangle}{\langle A,\ i\rangle}\mathcal{R}_{dn.i}}{\langle B,\ i\rangle}\supset\mathrm{E}_i}{\langle T(``B"),\ i+1\rangle}\mathcal{R}_{up.i}}{\frac{\langle T(``A") \supset T(``B"),\ i+1\rangle}{\langle T(``A \supset B") \supset (T(``A") \supset T(``B")),\ i+1\rangle}\supset\mathrm{I}_{i+1}}\supset\mathrm{I}_{i+1}$$

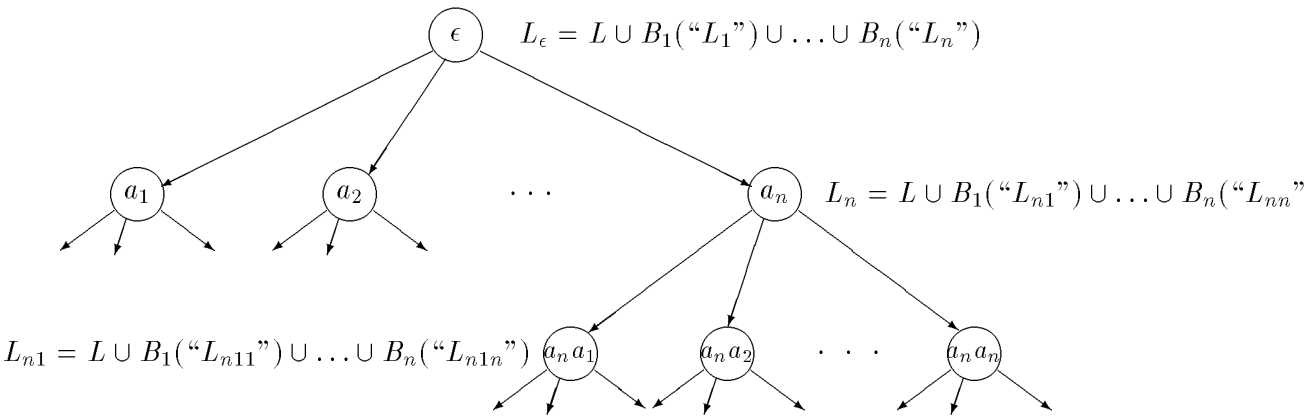

Figure 7: Structure and languages of MBK(n)

4 The MR system MK

In this section we concentrate on MK. The results presented in this section can be suitably generalized with a relatively minor effort to any MR system.

The following remark makes explicit the definition of MK given in the previous section.

Remark 4.1 Let P be a set of propositional constants, MK $= \langle \{L_i\}_{i\in\omega}, \{\Omega_i\}_{i\in\omega}, \Delta\rangle$, is such that:

(i) each language L_i is the smallest set closed under (a)–(f):

- (a) $\bot$ is an atomic L_i-wff;
- (b) if $p \in P$ then p is an atomic L_i-wff;
- (c) if A is an atomic L_i-wff then it is an L_i-wff;
- (d) if A, B are L_i-wffs then $A \supset B, A \wedge B, A \vee B$ are L_i-wffs;
- (e) if A is an L_i-wff then "A" is an L_{i+1}-term;
- (f) if t is an L_{i+1}-term then $T(t)$ is an atomic L_{i+1}-wff.

(ii) $\Omega_i = \emptyset$ for all $i \in \omega$;

(iii) Δ contains, for each language L_i the set of classical ND rules, plus $\mathcal{R}_{up.}$ and

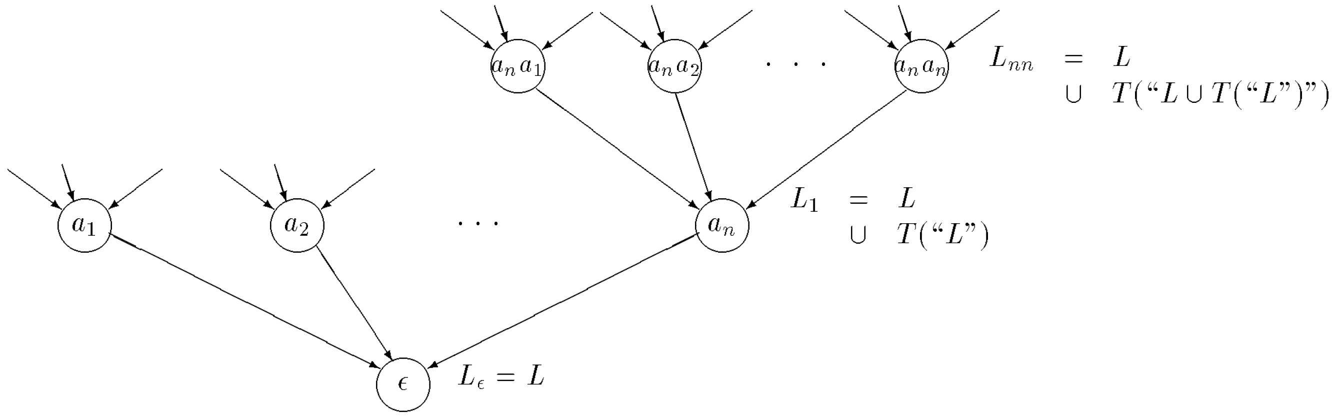

Figure 6: Structure and languages of MK(n)

Remark 3.9 In MK(n) all the meta-theories are identical. Namely

$$\vdash_{\mathrm{MK(n)}} \langle A,\ ia_k\rangle \text{ if and only if } \vdash_{\mathrm{MK(n)}} \langle A,\ ia_h\rangle$$

for each $i \in \mathcal{A}^*$ and $a_k, a_h \in \mathcal{A}$. Furthermore

$$\langle T(\text{``}A\text{''}),\ ia_k\rangle \vdash_{\mathrm{MK(n)}} \langle T(\text{``}A\text{''}),\ ia_h\rangle$$

This makes MK(n) not very interesting. More interesting versions of MK(n) can be obtained by considering *partial* metatheories, *i.e.* metatheories which partially describe the derivability relation of their object theory.

Example 3.6 (MBK(n)) MBK(n) is the basic system for the representation of multi-agent beliefs [GS93, GS91, GSGF92]. The idea underlying the formalization of multiagent belief is that there is a set of agents $\mathcal{A} = \{a_1, a_2 \ldots a_n\}$, who have beliefs about the world, beliefs about their own beliefs and beliefs about the other agents' beliefs. The basic structure of MBK(n) is reported in figure 7.

Definition 3.7 (MBK(n)) *Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ be a set of n symbols (agents), and $\mathcal{A}^*$ the set of finite (possible empty) sequences of symbols in $\mathcal{A}$. Let $\epsilon \in \mathcal{A}^*$ denote the empty sequence. Let L be a propositional language and $\mathcal{S} = \{S_i\}_{i \in I}$ a family of formal systems. Then MBK(n) is the MR system M$\mathcal{S}$ such that:*

(i) *I is the set $\mathcal{A}^*$;*

(ii) *$ia_k \prec i$, for each $i \in I$ and $a_k \in \mathcal{A}$;*

(iii) *$\bullet_{ia_k}$ is B_k for each $i \in I$ and $a_k \in \mathcal{A}$;*

(iv) *$S_i = C_L$.*

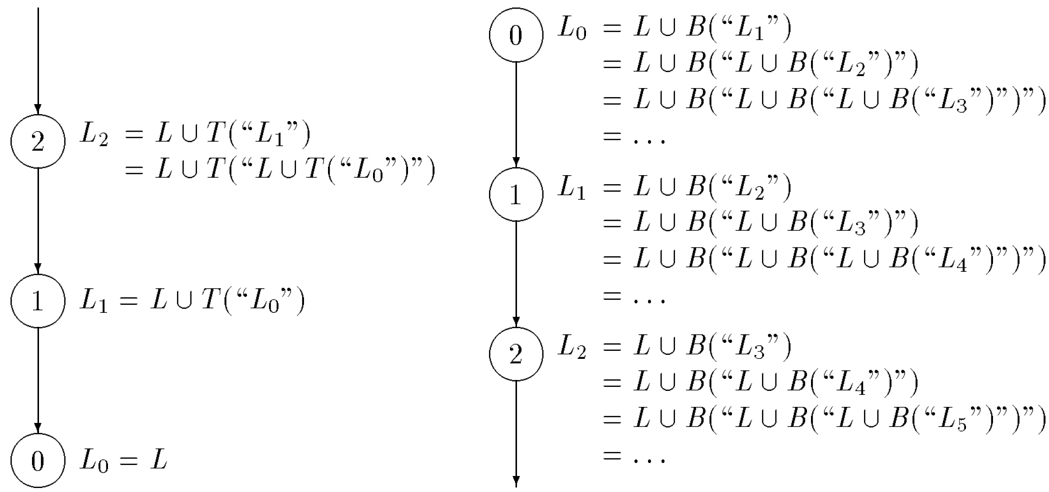

Figure 5: Structures and languages of MK and MBK

Example 3.4 (MBK) MBK is the basic ML system for the representation of beliefs [GS91, GSGF92, GS92]. The idea underlying the formalization of belief is that there is an agent a who has beliefs about the world and beliefs about its own beliefs. Given a proposition A about the state of the world, $B(\text{“}A\text{”})$ means that A itself is believed by a or, in other words, that A holds in a's view of the world. Similarly $B(\text{“}B(\text{“}A\text{”})\text{”})$ means that $B(\text{“}A\text{”})$ is believed by a, *i.e.* A holds in a's view of its beliefs of the world, and so on. In MBK each view is modeled as a theory with its own language. MBK's basic structure is reported in figure 5.

Definition 3.5 (MBK) *MBK is defined the same as MK except that $\prec$ and the predicates $\bullet_i$ are defined as follows:*

(iii) $i + 1 \prec i$ *for all* $i \in \omega$*;*

(iv) $\bullet_i$ *is* B*;*

Example 3.5 (MK(n)) MK can de actually generalized to consider the case where each theory has multiple metatheories (see figure 6).

Definition 3.6 (MK(n)) *Let $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$ be a set of n symbols (metaviews), and $\mathcal{A}^*$ the set of finite (possible empty) sequences of symbols in $\mathcal{A}$. Let $\epsilon \in \mathcal{A}^*$ denote the empty sequence. Let L be a propositional language and $\mathcal{S} = \{S_i\}_{i \in I}$ a family of formal systems. Then MK(n) is the MR system $\mathcal{MS}$ such that:*

(i) I *is the set* $\mathcal{A}^*$*;*

(ii) $i \prec ia_k$ *for each* $i \in I$ *and* $a_k \in \mathcal{A}$*;*

(iii) $\bullet_i$ *is* T *for each* $i \in I$*;*

(iv) $S_i = C_L$ *for each* $i \in I$*.*

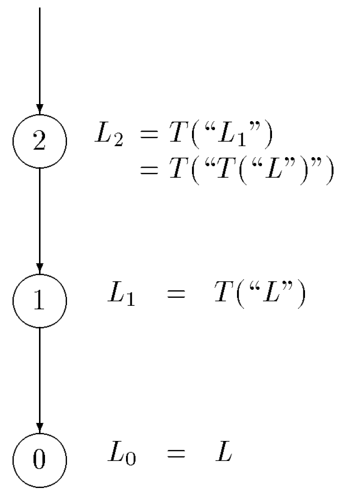

Figure 4: Structures and languages of MPK

(i) *I is the set ω of natural numbers;*

(ii) $i \prec i+1$ *for all* $i \in \omega$*;*

(iii) $\bullet_i$ *is* T*;*

(iv) $S_i = \begin{cases} C_L & \textit{if } i = 0; \\ C_\emptyset & \textit{if } i > 0; \end{cases}$

Example 3.3 (MK) MK is the MR system studied in detail in this paper. Analogously to MPK, MK is an ML system for the formalization of metatheoretic reasoning [GS93, GT92, GT91]. Indeed MK is the same as MPK except for the fact that the language of each metatheory contains the language of its object theory. Hence in MK, if the language of the theory i is L, then the language of its metatheory (*i.e.* the $(i+1)$ theory) is $L \cup T(\text{“}L\text{”})$ (see figure 5).

Definition 3.4 (MK) *Let L be a propositional language, and $\mathcal{S} = \{S_i\}_{i \in I}$ a family of formal systems. Then MK is the MR system $M\mathcal{S}$ such that:*

(i) *I is the set ω of the natural numbers;*

(ii) $i \prec i+1$ *for all* $i \in \omega$*;*

(iii) $\bullet_i$ *is* T*;*

(iv) $S_i = C_L$.

Remark 3.8 In definition 3.4, $L_i \subseteq L_{i+1}$. Certain results, like the equivalence with modal logics, are more easily formalizable under this hypothesis. The results for MK can be suitable specialized to the case with smaller languages. [GSS92] presented a version of MK with $L_i \not\subseteq \mathrm{L}_{i+1}$.

(i) *for each $i \in I$, L'_i is the minimal language such that:*

   (a) $L_i \subseteq L'_i$*;*

   (b) if $A \in L_i$ and $i \prec j$, then "A" is a constant symbol of L_j;

   (c) if $i \prec j$, then $\bullet_i$ is a unary predicate symbol of L_j;

   (d) L'_i is closed under the formation rules of L_i;

(ii) $\Omega'_i = \Omega_i$*;*

(iii) *Δ contains for each $i \in I$ the multi-language version of the rules Δ_i with index i, and the bridge rules $\mathcal{R}_{up.i}$ and $\mathcal{R}_{dn.i}$ for each $j \in I$ such that $i \prec j$.*

Remark 3.6 In general, combining ND systems in an MR system causes the addition of new symbols (*e.g.* $\bullet_i$, "A") to the languages of the ND systems. If the axioms Ω_i of a ND system S_i are schemata (rather than being sets), then the axioms Ω'_i (*i.e.* the corresponding axioms in the ML system) are the schemata Ω_i applied to L'_i. The same argument applies to the deductive machinery.

Remark 3.7 Definition 3.2 states that an MR system can be defined by specifying:

(i) a non empty set I;

(ii) the relation $\prec$ on I;

(iii) a unary predicate $\bullet_i$ for each $i \in I$;

(iv) For each $i \in I$ a formal system $S_i = \langle L, \Omega, \Delta \rangle$, where L is a (possibly empty) first order language.

We conclude this section with the definition of some significative examples of MR systems, that is MPK, MK, MBK, MK(n) and MBK(n).

Notation 3.2 Let C_L be the classical natural deduction system (see [Pra65]) on a (subset of a) first order language L. Let $C_\emptyset$ be the classical natural deduction system on the language without any nonlogical symbol.

Example 3.2 (MPK) MPK is the basic propositional system for the formalization of theorem proving with metatheories [GT92, GT91]. In metatheoretic reasoning, one usually starts with the object theory and then defines its metatheory, its metametatheory and so on. Analogously, in MPK (see figure 4), the bottom theory, labeled with 0, is any object theory with language L, the theory labeled with 1 is its simplest metatheory, the theory labeled with 2 is its simplest metametatheory and so on. The *simplest* metatheory of an object theory with language L, is a theory with the minimal linguistic and axiomatic requirements. The language of such a metatheory is $T(\text{"}L\text{"})$, that is the propositional language whose only atomic wffs are $\{T(\text{"}A\text{"}) : A \text{ is an } L\text{-wff}\}$. Its set of axioms is the empty set.

Definition 3.3 (MPK) *Let L be a propositional language and $\mathcal{S} = \{S_i\}_{i \in I}$ a family of formal systems. Then MPK is the MR system $M\mathcal{S}$ such that:*

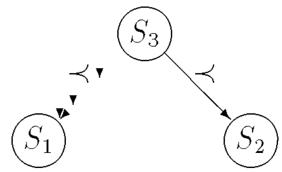

Figure 3: Structure of $\langle \mathcal{S}, \prec \rangle$

Example 3.1 Let $\mathcal{S} = \{S_1, S_2, S_3\}$ be a set of ND systems such that:

$$S_1 = \langle L_1, \Omega_1, \Delta_1 \rangle$$
$$S_2 = \langle L_2, \Omega_2, \Delta_2 \rangle$$
$$S_3 = \langle L_3, \Omega_3, \Delta_3 \rangle$$

and $\prec$ a binary relation on $\{1, 2, 3\}$ such that:

$$1 \prec 3$$
$$2 \prec 3$$

The structure of the ordered set $\langle \mathcal{S}, \prec \rangle$ of ND systems is shown in figure 3.
The MR system $\mathrm{M}\mathcal{S} = \langle \{L'_i\}_{i \in \{1,2,3\}}, \{\Omega'_i\}_{i \in \{1,2,3\}}, \Delta' \rangle$ generated from the family $\langle \mathcal{S}, \prec \rangle$ has the same structure as that reported in figure 3 and is such that:

(*a*) $L'_1 = L_1$, $L'_2 = L_2$ and L'_3 contains L_3, the names for the L_1-wffs and L_2-wffs and the predicates $\bullet_1$ and $\bullet_2$.

(*b*) Ω'_i is the same as Ω_i.

(*c*) The deductive machinery Δ' contains, for each $i \in \{1, 2, 3\}$, the multi-language version of the rules in Δ_i with index i and the bridge rules:

$$\frac{\langle A,\ 1 \rangle}{\langle \bullet_1(\text{“}A\text{”}),\ 3 \rangle}\,\mathcal{R}_{up.1} \quad \frac{\langle A,\ 2 \rangle}{\langle \bullet_2(\text{“}A\text{”}),\ 3 \rangle}\,\mathcal{R}_{up.2} \quad \frac{\langle \bullet_1(\text{“}A\text{”}),\ 3 \rangle}{\langle A,\ 2 \rangle}\,\mathcal{R}_{dn.1} \quad \frac{\langle \bullet_2(\text{“}A\text{”}),\ 3 \rangle}{\langle A,\ 2 \rangle}\,\mathcal{R}_{dn.2}$$

$\mathrm{M}\mathcal{S}$ inherits properties from the source ND systems. For instance the following hold:

$$\text{If } A \vdash_{S_1} B \text{ then } \langle A,\ 1 \rangle \vdash_{\mathrm{M}\mathcal{S}} \langle B,\ 1 \rangle \tag{9}$$
$$\text{If } A \vdash_{S_1} B \text{ then } \langle \bullet_1(\text{“}A\text{”}),\ 3 \rangle \vdash_{\mathrm{M}\mathcal{S}} \langle \bullet_1(\text{“}B\text{”}),\ 3 \rangle \tag{10}$$

Generalizing the previous example, we give the following definition.

Definition 3.2 (MR system generated from a family of ND systems) *Let $\mathcal{S} = \{S_i = \langle L_i, \Omega_i, \Delta_i \rangle\}_{i \in I}$ a (at worst countable) family of first order ND systems, and $\prec$ a relation on I satisfying condition (iii) of definition 3.1. The MR system $M\mathcal{S} = \langle \{L'_i\}_{i \in I}, \{\Omega'_i\}_{i \in I}, \Delta \rangle$ is defined as follows:*

they respect conditions (5) and (6). In the following proposition we prove that any two theories i and j, of any MR system such that $i \prec j$, constitute an O/M pair.

Proposition 3.1 *Let MS be an MR system and $i, j \in I$ such that $i \prec j$, then for any set of L_i-wffs $G \cup \{A, B\}$:*

(i) $\vdash_{\mathrm{MS}} \langle \bullet_i(\text{“}A\text{”}),\ j\rangle$ *if and only if* $\vdash_{\mathrm{MS}} \langle A,\ i\rangle$*;*

(ii) *if* $\langle G,\ i\rangle \vdash_{\mathrm{MS}} \langle B,\ i\rangle$ *then* $\langle \bullet_i(\text{“}G\text{”}),\ j\rangle \vdash_{\mathrm{MS}} \langle \bullet_i(\text{“}B\text{”}),\ j\rangle$*.*

Proof (i) If $\vdash_{\mathrm{MS}} \langle \bullet_i(\text{“}A\text{”}), j\rangle$, then there exists a proof Π in MS of $\langle \bullet_i(\text{“}A\text{”}), j\rangle$. Applying $\mathcal{R}_{dn.i}$ to the conclusion of Π we obtain a proof of $\langle A,\ i\rangle$. The viceversa is a special case of (ii) when $G = \emptyset$.

(ii) $\langle G,\ i\rangle \vdash_{\mathrm{MS}} \langle B,\ i\rangle$ implies that there exists a deduction Π in MS of $\langle B,\ i\rangle$ from $\langle G,\ i\rangle$. The formula tree

$$\cfrac{\cfrac{\langle \bullet_i(\text{“}C\text{”}),\ j\rangle}{(\langle C,\ i\rangle)}\ \mathcal{R}_{dn.i}}{\cfrac{\Pi}{\cfrac{\langle B,\ i\rangle}{\langle \bullet_i(\text{“}B\text{”}),\ j\rangle}\ \mathcal{R}_{up.i}}} \tag{7}$$

is obtained by writing $\langle \bullet_i(\text{“}C\text{”}), j\rangle$ on the top of every undischarged assumption $\langle C,\ i\rangle \in \langle G,\ i\rangle$ of Π, and by writing $\langle \bullet_i(\text{“}B\text{”}),\ j\rangle$ under the conclusion of Π. We show that the formula tree (7) is a deduction of $\langle \bullet_i(\text{“}B\text{”}),\ j\rangle$ from $\langle \bullet_i(\text{“}G\text{”}),\ j\rangle$. Every application of $\mathcal{R}_{up.k}$ in Π, whose premise depends on an assumption $\langle C,\ i\rangle \in \langle G,\ i\rangle$, is allowed in (7). In fact, if h is the index of the consequence of $\mathcal{R}_{up.k}$, we have that $i \not\prec^+ h$ (for the applicability of $\mathcal{R}_{up.k}$ in Π) and $i \prec j$ (by hypothesis) imply that $j \not\prec^+ h$, which guarantees the applicability of $\mathcal{R}_{up.k}$ in Π'. Furthermore the last application of $\mathcal{R}_{up.i}$ to $\langle B,\ i\rangle$ which derives $\langle \bullet_i(\text{“}B\text{”}),\ j\rangle$, is allowed as $j \not\prec^+ j$. Q.E.D.

Remark 3.4 MR systems are the multilanguage counterpart of normal modal logics in the sense that their provability relations can be put in “correspondence” with the provability relations of normal modal logics. In some cases (but not always) this correspondence amounts to an isomorphism under the “natural” mapping of a first order language into a modal language (theorem 7.1). In general the correspondence is an embedding of the MR's provability relation into the provability relation of modal K. In the following sections we will be more precise. At this point it is useful to notice that, in an MR system, for any $i, j \in I$ such that (i) $i \prec j$, (ii) *modus ponens* is admissible in i and (iii) the fact that the deduction theorem holds in j implies that:

$$\vdash_{\mathrm{MS}} \langle \bullet_i(\text{“}A \supset B\text{”}) \supset (\bullet_i(\text{“}A\text{”}) \supset \bullet_i(\text{“}B\text{”})),\ j\rangle \tag{8}$$

i.e. the formalization of *modus ponens* of j is provable in i. Under the obvious interpretation (read $\bullet_i(\text{“}A\text{”})$ as $\Box A$), (8) is the axiom characterizing normal modal logics.

Remark 3.5 Definition 3.1 suggests a general way to build MR systems starting from a set $\mathcal{S}$ of first order ND systems and a binary relation $\prec$ on them. Intuitively this requires the transformation of each system S_i into its corresponding multi-language version with index i, and the addition to S_i of $\bullet_j$ and the names of the wffs of S_j, for each $j \prec i$.

(i) *I is a not empty countable set with cardinality less than or equal to $\aleph_0$;*

(ii) *each language L_i is (a subset of) a first order language;*

(iii) *a recursive binary relation $\prec$ is defined on I, such that $\prec^+$ is not reflexive;*

(iv) *for each $i, j \in I$, if $i \prec j$, then L_j contains* names *for all the L_i-wffs; i.e. for each L_i-wff A there exists a L_j-term, denoted by "A", such that for each L_i-wff A, "A" is recursively computable from A and for each L_i-wff B distinct from A, "A" is distinct from "B";*

(v) *for each $i, j \in I$, if $i \prec j$, then L_j contains a unary predicate denoted by $\bullet_i$;*

(vi) *the only bridge rules are, for each $i, j \in I$, such that $i \prec j$,*

$$\frac{\langle A,\ i\rangle}{\langle \bullet_i(\text{“}A\text{”}),\ j\rangle}\ \mathcal{R}_{up.i} \qquad \frac{\langle \bullet_i(\text{“}A\text{”}),\ j\rangle}{\langle A,\ i\rangle}\ \mathcal{R}_{dn.i}$$

where $\mathcal{R}_{up.}$ is applicable if and only if $\langle A,\ i\rangle$ does not depend from any assumption $\langle B,\ k\rangle$ with $k \prec^+ j$.

The class of all the MR systems is called $\mathcal{MR}$.

Remark 3.2 Condition (i) of definition 3.1 ensures that the set of languages is finite or at worst enumerable; this restriction is justified by the fact that we want to prevent from having in the language an uncountable set of predicates $\bullet_i$ (see request (v)). Request (ii) is justified mainly on pragmatic grounds. So far we have never needed more expressiveness than that provided by a first order language. Restriction (iii) ensures that the relation $\prec$ is computable. This guarantees that the applicability of $\mathcal{R}_{up.i}$ and $\mathcal{R}_{dn.i}$, is decidable. $\prec$ induces an *acyclic* graph on I, which prevents theories from being self-referential. (iv) and (v) are the minimal expressivity conditions under which the language L_j can be used to provide a metatheory for the provability relation of any object theory expressed in a language L_i, see [GSS92]. Condition (vi) ensures that, for any pair of theories i and j, such that $i \prec j$, the metatheory j is *sound* and *closed* with respect to the derivability relation of the object theory i.

Remark 3.3 (Soundness and Closure) In the previous remark we have mentioned the notions of soundness and closure for a metatheory with respect to the derivability relation of an object theory. We need to be more precise about this point. Reporting what has been argued in [GSS92], for a metatheory M to be sound, the properties of an object theory O, expressed by theorems of M must, indeed, hold in O. In the case of theoremhood, $\bullet_i$ becomes T and the requirement is:

$$\vdash_{\mathrm{M}} T(\text{“}A\text{”}) \textit{ if and only if } \vdash_{\mathrm{O}} A \tag{5}$$

The closure of M on the derivability relation $\vdash_{\mathrm{O}}$, is formalized by the following condition (which is a generalization of the condition described in [GSS92]):

$$\textit{if } G \vdash_{\mathrm{O}} B \textit{ then } T(\text{“}G\text{”}) \vdash_{\mathrm{M}} T(\text{“}B\text{”}) \tag{6}$$

were G is a set of wffs of the language of the object theory, and $T(\text{“}G\text{”})$ is the set $\{T(\text{“}A\text{”}) : A \in G\}$. We say that two theories constitute an O/M pair if and only if

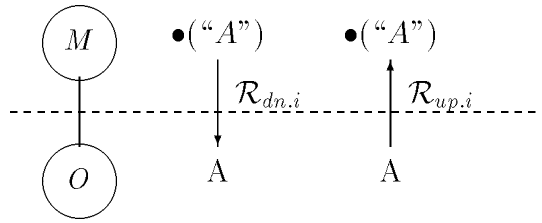

Figure 2: The class $\mathcal{MR}$

Proposition 2.1 $\vdash_{L_i}$ *is a derivability relation: i.e. for each set of L_i-wffs $G \cup F \cup \{A\}$:*

(RX) $A \vdash_{L_i} A$;
(LM) *if* $G \vdash_{L_i} A$, *then* $G, F \vdash_{\mathbf{L}} A$;
(CUT) *if* $G \vdash_{L_i} A$ *and* $F, A \vdash_{L_i} B$, *then* $G, F \vdash_{L_i} B$.

3 The Class $\mathcal{MR}$

Informally, each member of $\mathcal{MR}$ has the following properties:

(a) the languages are ordered in a hierarchy;
(b) each language in the hierarchy has names for the wffs of the level below;
(c) any two adjacent languages in the hierarchy, say M and O, are linked only by two bridge rules of the form:

$$\frac{\langle A,\ O\rangle}{\langle \bullet(\text{“}A\text{”}),\ M\rangle}\ \mathcal{R}_{up.} \qquad \frac{\langle \bullet(\text{“}A\text{”}),\ M\rangle}{\langle A,\ O\rangle}\ \mathcal{R}_{dn.}$$

where “•” is a unary predicate. $\mathcal{R}_{up.}$ is called reflection up, $\mathcal{R}_{dn.}$ reflection down.

Figure 2 gives a graphical representation of the language structure of an MR system.

Remark 3.1 This hierarchical structure with multiple languages is somehow similar to Tarski's [Tar36]. Intuitively, the main difference is that here all the metatheories are formalized and communicate via bridge rules. A very similar multi-theory version of reflection up and reflection down was introduced in [GS89].

Notation 3.1 If $\prec$ is a binary relation on a set I, then $\prec^+$ denotes the transitive closure of $\prec$ on I, $\prec^*$ denotes the transitive and reflexive closure of $\prec$ on I.

Definition 3.1 (MR system) *The ML system $MS = \langle \{L_i\}_{i \in I}, \{\Omega_i\}_{i \in I}, \Delta\rangle$, is an MR system if and only if:*

While the threads of the deduction of item (i′) of example 2.5 are:

$$\begin{array}{rcl} \tau_1' & = & \langle A,\ 1\rangle, \langle A,\ 3\rangle, \langle B,\ 3\rangle \\ \tau_2' & = & \langle A \supset B,\ 2\rangle, \langle A \supset B,\ 3\rangle, \langle B,\ 3\rangle \end{array}$$

Notice that the number of distinct threads of a deduction is the same as the number of leaves.

Definition 2.10 (Discharged assumption) *Let $\langle A,\ i\rangle$ be an assumption of a deduction Π. $\langle A,\ i\rangle$ is* undischarged *in a deduction or* discharged *at a certain occurrence in a deduction, according to the following rules:*

(i) *$\langle A,\ i\rangle$ is undischarged in $\langle A,\ i\rangle$*

(ii) *if Π is*

$$\frac{\begin{array}{ccc} \Pi_1 & & \Pi_n \\ \langle B_1,\ j_1\rangle & \ldots & \langle B_n,\ j_n\rangle \end{array}}{\langle B,\ j\rangle}\ \iota$$

where $\alpha = \langle\langle B_1, j_1\rangle, \ldots, \langle B_n, j_n\rangle, \langle B, j\rangle\rangle$ and $\langle A, i\rangle$ is undischarged in Π_k, then $\langle A,\ i\rangle$ is discharged at $\langle B,\ j\rangle$ by α if and only if it is an element of $d_k(\iota)(\alpha)$ (the k-th discharging function of ι), otherwise $\langle A,\ i\rangle$ is undischarged in Π.

An assumption is discharged in Π *if and only if it is discharged at some formula occurrence of Π.*

It is easy to see that an assumption in Π is either discharged or undischarged.

Remark 2.10 Any deduction can be seen as composed of subdeductions in distinct languages, obtained by repeated applications of i-rules, concatenated by the application of bridge-rules.

The deductive machinery in a logic determines the derivability relation in that logic. Since deductions in ML systems are allowed to have assumptions and conclusion in distinct languages, a new notion of derivability relation which spans multiple languages is necessary. The notion of derivability relation for ML systems, which is a generalization of the single language derivability relation, is introduced by the following definition.

Definition 2.11 (Multi-Language Derivability Relation) *A* Multi-Language Derivability Relation *(*ML-D.R.*) on a family of languages $\mathbf{L} = \{L_i\}_{i\in I}$ is a binary relation between sets of formulas and formulas such that:*

(ML–RX) *$\langle A,\ i\rangle \vdash_{\mathbf{L}} \langle A,\ i\rangle$;*

(ML–LM) *if $\Gamma \vdash_{\mathbf{L}} \langle A,\ i\rangle$, then $\Gamma, \Sigma \vdash_{\mathbf{L}} \langle A,\ i\rangle$;*

(ML–CUT) *if $\Gamma \vdash_{\mathbf{L}} \langle A,\ i\rangle$ and $\Sigma, \langle A,\ i\rangle \vdash_{\mathbf{L}} \langle B,\ i\rangle$, then $\Gamma, \Sigma \vdash_{\mathbf{L}} \langle B,\ i\rangle$.*

For each $L_i \in \mathbf{L}$, $\vdash_{L_i}$ is the set $\{G \vdash_{L_i} A : \langle G,\ i\rangle \vdash_{\mathbf{L}} \langle A,\ i\rangle\}$. Furthermore the i-theory is the set $\{A \in L_i : \vdash_{L_i} A\}$.

- if $L_1 \neq L_2$ then ML_3 is weaker than ML'_3;

This highlights the fact that the relation between languages in an ML system plays a crucial role in the formalization of a problem. For example, in the formalization of belief, we are often interested in non-omniscient agents whose reasoning capabilities are bounded by their languages. In a single language approach,a limitation in the language must be explicitly formalized by means of a meta-theoretic predicate or a modal operator, whose meaning is given by a particular axiomatization (see for example the "*awareness*" operator introduced by Fagin and Halpern in [FH88]). In the multi-language approach the limitation of reasoning capabilities is directly induced by the choice of the languages. No explicit axiomatization is needed. An in depth discussion and several examples of ML systems for non-omniscient believers can be found in [GSGF92].

Remark 2.9 Indexes are not part of the languages of an ML system. They are, rather, a "metanotation" useful proof-theoretically to keep track of the locality of the reasoning. This is a key point which makes clear the difference between ML systems and logics where the language allows for indexed elements. Similar, but still quite different from ML Systems, are Labeled Deductive Systems (LDS), introduced by D. Gabbay in [Gab90]. In such systems an inference rule is applied to a set of *Labeled formulas* $t_1 : A_1, \ldots, t_n : A_n$ and to conditions on labels (*e.g.* $t_1 < t_2 < \ldots < t_n$), and it returns a Labeled formula $t : A$. Though our definition of inference rule is quite similar, the role of indexes in ML systems is very different from that of labels in LDS. In ML systems indexes encode the metalevel information of the localization of the reasoning, while, in LDS, labels keep track of the metalevel information of the derivation process (*e.g.* the number of times an assumption is exploited, the number of steps of a deduction, the current possible world). Furthermore, rules of inference in LDS accept both object level premises (Labeled formulas) and metalevel premises (conditions on labels), while rules of inference in ML systems are applied only to formulas.

Definition 2.8 (Thread) *A* thread *of a deduction* Π *is a sequence* $\delta = \langle A_1, i_1\rangle, \ldots \langle A_n, i_n\rangle$ *such that:*

(i) $\langle A_1,\ i_1\rangle$ *is a top formula of* Π*;*
(ii) $\langle A_n,\ i_n\rangle$ *is the bottom formula of* Π*;*
(iii) *for every* $1 \leq k < n$*,* $\langle A_{k+1},\ i_{k+1}\rangle$ *is a consequence of an application* α *and* $\langle A_k,\ i_k\rangle$ *is a premise of* α*.*

Definition 2.9 (Assumption) $\langle A, i\rangle$ *is an assumption of a deduction* Π *in MS, if and only if, it is a top formula of* Π *and it is not an axiom of MS.*

Example 2.6 The threads of the deduction of item (i) of example 2.4 are:

$$\begin{aligned}\tau_1 &= \langle A,\ 1\rangle, \langle A,\ 3\rangle \\ \tau_2 &= \langle A,\ 2\rangle, \langle A \supset A,\ 2\rangle, \langle A,\ 3\rangle\end{aligned}$$

Definition 2.7 (Theorem) *A deduction* Π *is a* deduction *in MS* of $\langle A,\, i\rangle$ *from* Γ *if and only if* Π *is a deduction in MS of* $\langle A,\, i\rangle$ *depending on* Γ *or some subset of* Γ*.*

$\langle A,\, i\rangle$ *is* derivable from Γ *in MS—abbreviated as* $\Gamma \vdash_{\mathrm{MS}} \langle A,\, i\rangle$*—if and only if there is a deduction in MS of* $\langle A,\, i\rangle$ *from* Γ*.*

A deduction of $\langle A,\, i\rangle$ *from the empty set is a* proof *of* $\langle A,\, i\rangle$*.* $\langle A,\, i\rangle$ *is* provable in *(*a theorem of *) MS—abbreviated as* $\vdash_{\mathrm{MS}} \langle A,\, i\rangle$*—if and only if there exists a proof in MS of* $\langle A,\, i\rangle$*.*

Example 2.4 (Example 2.1 – continued) Let

$$\frac{\begin{array}{c}[\langle A,\, i\rangle]\\ \langle B,\, i\rangle\end{array}}{\langle A\supset B,\, i\rangle}\supset\mathrm{I}_i$$

be an i-rule of ML_3 for $i = 1, 2$. The following are some deductions in ML_3.

(i) $\langle A,\, 1\rangle \vdash_{\mathrm{ML}_3} \langle A,\, 3\rangle$, if and only if A is an L_1-wff, an L_2-wff and an L_3-wff;

(ii) $\langle A,\, 2\rangle \vdash_{\mathrm{ML}_3} \langle A,\, 3\rangle$: if and only if A is an L_1-wff, an L_2-wff and an L_3-wff.

$$\frac{\langle A,\, 1\rangle \quad \dfrac{\langle A,\, 2\rangle}{\langle A\supset A,\, 2\rangle}\supset\mathrm{I}_2}{\langle A,\, 3\rangle}\mathrm{MMP}_{1,2} \qquad \frac{\langle A,\, 2\rangle \quad \dfrac{\langle A,\, 1\rangle}{\langle A\supset A,\, 1\rangle}\supset\mathrm{I}_1}{\langle A,\, 3\rangle}\mathrm{MMP}_{2,1}$$

Example 2.5 (Example 2.2 – continued) The following are some deductions in ML_3':

(i′) $\langle A,\, 1\rangle, \langle A\supset B,\, 2\rangle \vdash_{\mathrm{ML}_3'} \langle B,\, 3\rangle$;

(ii′) $\langle A,\, 2\rangle, \langle A\supset B,\, 1\rangle \vdash_{\mathrm{ML}_3'} \langle B,\, 3\rangle$.

$$\frac{\dfrac{\langle A,\, 1\rangle}{\langle A,\, 3\rangle}1{\subseteq}3 \quad \dfrac{\langle A\supset B,\, 2\rangle}{\langle A\supset B,\, 3\rangle}2{\subseteq}3}{\langle B,\, 3\rangle}\supset\mathrm{E}_3 \qquad \frac{\dfrac{\langle A,\, 2\rangle}{\langle A,\, 3\rangle}2{\subseteq}3 \quad \dfrac{\langle A\supset B,\, 1\rangle}{\langle A\supset B,\, 3\rangle}1{\subseteq}3}{\langle B,\, 3\rangle}\supset\mathrm{E}_3$$

Remark 2.7 The systems presented in examples 2.1 and 2.2 may seem equivalent, but they are not. It is indeed true that ML_3' is stronger than ML_3, as conditions (i′) and (ii′) of example 2.5 state that the ML_3's bridge rules $\mathrm{MMP}_{1,2}$ and $\mathrm{MMP}_{2,1}$ are admissible rules of ML_3'. On the other hand, (i) (resp. (ii)) of example 2.4 is not generally strong enough to state that $1{\subseteq}2$ (resp. $2{\subseteq}3$) is admissible in ML_3. If, for example, $L_1 \cap L_2 = \emptyset$, then $\mathrm{MMP}_{1,2}$ and $\mathrm{MMP}_{2,1}$ are never applicable, as there is no wff $A \in L_1$ such that $A \supset B \in L_2$ and viceversa. If $L_1 = L_2$ then ML_3 is equivalent to ML_3' (provided that $\supset\mathrm{E}_3$ is admissible in ML_2').

Remark 2.8 In examples 2.1 and 2.2 we have two systems ML_3 and ML_3' composed of three languages L_1, L_2 and L_3 such that:

- if $L_1 = L_2$ then ML_3 is equivalent to ML_3';

version of a unary ND rule ι. ι_i is applicable to a premiss $\langle A, i\rangle$ depending on the set of assumptions $\langle G, i\rangle$, if and only if it is applicable to $\langle A, i\rangle$ depending on $\Gamma \cup \langle G, i\rangle$, for any set of wffs Γ with index $\neq i$. Notice that the applications of the multi-language version of a ND rule are always localized to a single index i, only involving wffs belonging to L_i. Viceversa any rule which respects this property of "locality" can be described as the multi-language version of a ND rule.

Definition 2.4 (i-rule and bridge rule) *A multi-language inference rule is an i-rule if and only if it is the multi-language version of some ND inference rule at i. A multi-language inference rule is a bridge rule if and only if it is not an i-rule.*

Remark 2.6 Notice that definition 2.4 slightly refines remark 2.4. In particular, the rules which have restrictions involving formulas with index different from that of the conclusion are bridge rules.

Definition 2.5 (Formula-tree) *A* formula-tree *in an ML system MS is defined recursively as follows:*

(i) $\langle A, i\rangle$ *is a formula tree;*
(ii) *if* $\Pi_1 \ldots \Pi_n$ *are formula trees then* $\frac{\Pi_1 \ldots \Pi_n}{\langle A, i\rangle}$ *is a formula tree;*
(iii) *nothing else is a formula tree.*

The top formulae *of a formula tree* Π *are the leaves of* Π *and the* bottom formula *of* Π *is the root of* Π.

Deductions are formula-trees built by starting from a finite number of assumptions and axioms, possibly belonging to distinct languages, and by applying a finite number of inference rules.

Definition 2.6 (Deduction ... depending on) *A formula-tree is a* deduction in MS *of* $\langle A, i\rangle$ depending on *a set of formulas according to the following rules:*

(i) *if* $\langle A, i\rangle$ *is an axiom of MS, i.e.* $A \in \Omega_i$, *then* $\langle A, i\rangle$ *is a deduction in MS of* $\langle A, i\rangle$ *depending on the empty set;*
(ii) *if* $\langle A, i\rangle$ *is not an axiom of MS, i.e.* $A \notin \Omega_i$, *then* $\langle A, i\rangle$ *is a deduction in MS of* $\langle A, i\rangle$ *depending on* $\{\langle A, i\rangle\}$*;*
(iii) *if* Π_k *is a deduction of* $\langle A_k, i_k\rangle$ *depending on* Γ_k *for any* $(1 \leq k \leq n)$, *then*

$$\frac{\Pi_1 \ \ldots \ \Pi_n}{\langle A, i\rangle}\ \iota$$

is a deduction in MS of $\langle A, i\rangle$ *depending on* $\mathbf{d}_\iota^\alpha(\Gamma_1, \ldots \Gamma_n)$, *if and only if:*

(a) $\alpha = \langle\langle A_1, i_1\rangle, \ldots, \langle A_n, i_n\rangle, \langle A, i\rangle\rangle$ *is an application of* ι*;*
(b) ι is applicable *i.e.:* $\langle\langle\Gamma_1, \langle A_1, i_1\rangle\rangle, \ldots, \langle\Gamma_n, \langle A_n, i_n\rangle\rangle\rangle \notin rest(\iota)$.

Definition 2.3 (Multi-Language version of ND inference rules) *If ι is an n-ary ND inference rule, that can be described as a tuple $\langle \rho(\iota), d_1(\iota), \ldots, d_n(\iota), rest(\iota)\rangle$, then its multi-language version with index i is $\iota_i = \langle \rho(\iota_i), d_1(\iota_i), \ldots, d_n(\iota_i), rest(\iota_i)\rangle$ where:*

(i) $\langle\langle A_1, i\rangle, \ldots, \langle A_n, i\rangle\langle A, i\rangle\rangle \in \rho(\iota_i)$ *if and only if* $\langle A_1, \ldots, A_n, A\rangle \in \rho(\iota)$
(ii) $\langle B, i\rangle \in d_k(\iota_i)(\langle A_1, i\rangle, \ldots, \langle A_n, i\rangle, \langle A, i\rangle)$ *if and only if* $B \in d_k(\iota)(A_1, \ldots, A_n, A)$.
(iii) $\langle\langle \Gamma_1, \langle A_1, i\rangle\rangle, \ldots, \langle \Gamma_n, \langle A_n, i\rangle\rangle\rangle \in rest(\iota_i)$ *if and only if* $\langle\langle G_1, A_1\rangle, \ldots, \langle G_n, A_n\rangle\rangle \in rest(\iota)$, *where* G_k $(1 \leq k \leq n)$ *is the set of all the wffs with index i occurring in* Γ_k.

Example 2.3 The ND rule $\exists$E for the elimination of the existential quantifier $\exists x$ and the substitution of x for a parameter p (as given in [Pra65])

$$\frac{\exists x A \qquad \begin{array}{c}[A^x_p]\\ B\end{array}}{B}\ \exists\mathrm{E}$$

is a 2-ary rule which can be described as a tupla $\langle \rho(\exists\mathrm{E}), d_1(\exists\mathrm{E}), d_2(\exists\mathrm{E}), rest(\exists\mathrm{E})\rangle$, where:

$$\begin{array}{rcl} \rho(\exists\mathrm{E}) & = & \{\langle \exists x A, B, B\rangle : \text{for any wffs } A, B\} \\ d_1(\exists\mathrm{E})(\exists x A, B, B) & = & \emptyset \\ d_2(\exists\mathrm{E})(\exists x A, B, B) & = & \{A^x_p\} \\ rest(\exists\mathrm{E}) & = & \{\langle\langle G_1, \exists x A\rangle, \langle G_2 \cup \{A^x_p\}, B\rangle\rangle : \\ & & p \text{ occurs in } G_2 \cup \{B, \exists x A\}\} \end{array}$$

Its multi-language version with index i, $\exists\mathrm{E}_i$,

$$\frac{\langle \exists x A,\ i\rangle \qquad \begin{array}{c}[\langle A^x_p,\ i\rangle]\\ \langle B,\ i\rangle\end{array}}{\langle B,\ i\rangle}\ \exists\mathrm{E}_i$$

is defined as $\langle \rho(\exists\mathrm{E}_i), d_1(\exists\mathrm{E}_i), d_2(\exists\mathrm{E}_i), rest(\exists\mathrm{E}_i)\rangle$ where:

$$\begin{array}{rcl} \rho(\exists\mathrm{E}_i) & = & \{\langle\langle \exists x A,\ i\rangle, \langle B,\ i\rangle, \langle B,\ i\rangle\rangle : \\ & & \text{for any } L_i\text{-wffs } A, B \text{ of } L_i\} \\ d_1(\exists\mathrm{E}_i)(\langle \exists x A,\ i\rangle, \langle B,\ i\rangle, \langle B,\ i\rangle) & = & \emptyset \\ d_2(\exists\mathrm{E}_i)(\langle \exists x A,\ i\rangle, \langle B,\ i\rangle, \langle B,\ i\rangle) & = & \{\langle A^x_p,\ i\rangle\} \\ rest(\exists\mathrm{E}_i) & = & \{\langle\langle \Gamma_1, \langle \exists x A,\ i\rangle\rangle, \langle \Gamma_2 \cup \{\langle A^x_p,\ i\rangle\}, \langle B,\ i\rangle\rangle\rangle : \\ & & p \text{ occurs in a wff } \langle C,\ i\rangle \in \Gamma_2, \cup\{\langle \exists x A,\ i\rangle, \langle B,\ i\rangle\}\} \end{array}$$

Remark 2.5 Item (i) and (ii) of definition 2.3 state that the multi-language version of a ND rule derives and discharges the same formulas as the corresponding ND rule. Item (iii) states that the restrictions on the applicability of the multi-language version of a ND rule at i depend only on the wffs with index equal to i. Let ι_i be the multi-language

Figure 1: ML_3 and ML_3'

(i) L_1, L_2 and L_3 are first order languages such that $L_1 \cup L_2 \subseteq L_3$;

(ii) $\Omega_i \subseteq L_i$;

(iii) Δ contains any set of i-rules ($i = 1, 2, 3$) and the following bridge rules:

$$\frac{\langle A,\, 1\rangle \quad \langle A \supset B,\, 2\rangle}{\langle B,\, 3\rangle}\ \mathrm{MMP}_{1,2} \qquad \frac{\langle A,\, 2\rangle \quad \langle A \supset B,\, 1\rangle}{\langle B,\, 3\rangle}\ \mathrm{MMP}_{2,1} \tag{3}$$

A graphical representation of ML_3 is given in figure 1. ML_3 can be seen as modeling a situation with three agents 1, 2, and 3. For each agent i ($i = 1, 2, 3$) the set of i-rules represents i's reasoning capability. $\mathrm{MMP}_{1,2}$ and $\mathrm{MMP}_{2,1}$ represent the fact that 1 and 2 use their mutual knowledge and "communicate" the result to 3. Notice that 1 cannot communicate anything to 3 without agreeing with 2.

Example 2.2 Let ML_3' be the same as ML_3 except that $\mathrm{MMP}_{1,2}$ and $\mathrm{MMP}_{2,1}$ are substituted by the following bridge rules

$$\frac{\langle A,\, 1\rangle}{\langle A,\, 3\rangle}\ 1{\subseteq}3 \qquad \frac{\langle A,\, 2\rangle}{\langle A,\, 3\rangle}\ 2{\subseteq}3 \tag{4}$$

Similarly to above, $1{\subseteq}3$ and $2{\subseteq}3$ can be taken to represent the fact that 1 and 2 communicate their knowledge to 3, this time independently. The result is that 3 is the union of 1 and 2.

As it can also be noticed from the examples above, the usual single language inference rules can be "rewritten" into the multi-language format. Thus, for instance, the multi-language version of

$$\frac{A \quad A \supset B}{B}\ {\supset}\mathrm{E} \quad \text{in the language } i \text{ is:} \quad \frac{\langle A,\, i\rangle \quad \langle A \supset B,\, i\rangle}{\langle B,\, i\rangle}\ {\supset}\mathrm{E}_i$$

A ND system can be seen as a particular case of an ML system where I is a singleton. Moreover any family $\{S_i\}_{i \in I}$ of ND systems can be seen as an ML system. However, the syntax of such ND systems must be slightly modified by adding suitable indexes to axioms and inference rules. For this and other reasons it is useful to define the notion of multi-language version of a ND inference (deduction) rule.

(ii) n discharging functions*:*
$d_1(\iota), \ldots, d_n(\iota)$ are n recursive functions where $d_k(\iota) : \rho(\iota) \perp\!\!\rightarrow 2^{\cup_{i \in I}\langle L_i, i\rangle}$ for each $1 \leq k \leq n$.

(iii) *an* applicability restriction*:*
$rest(\iota)$ is a recursive set of n-tuples of the form $\langle\langle\Gamma_1, \langle A_1, i_1\rangle\rangle, \ldots, \langle\Gamma_n, \langle A_n, i_n\rangle\rangle\rangle$ where each Γ_k is a set of wffs.

$\langle\langle A_1,\, i_1\rangle, \ldots, \langle A_n,\, i_n\rangle, \langle A,\, i\rangle\rangle$ is an application of ι if and only if it is in $\rho(\iota)$.

Remark 2.3 Definition 2.2 explicitly mentions concepts such as "discharging functions" and "applicability restriction". These concepts are extensively used in the paper. In this paper we consider rules whose discharging functions always return either the empty set (no assumption is discharged), or a singleton (only the assumptions of the specified formula are discharged).

Notation 2.2 Inference rules are informally written as:

$$\frac{\langle A_1,\, i_1\rangle \ldots \langle A_n,\, i_n\rangle}{\langle A,\, i\rangle}\ \iota \qquad (1)$$

or as:

$$\frac{\langle A_1,\, i_1\rangle \quad \ldots \quad \langle A_n,\, i_n\rangle \quad \begin{array}{c}[\langle B_1,\, j_1\rangle]\\ \langle A_{n+1},\, i_{n+1}\rangle\end{array} \quad \cdots \quad \begin{array}{c}[\langle B_m,\, j_m\rangle]\\ \langle A_{n+m},\, i_{n+m}\rangle\end{array}}{\langle A,\, i\rangle}\ \delta \qquad (2)$$

(1) represents a rule whose discharging functions return the empty set, (its applications do not discharge any assumption). (2) represents a rule such that the first n discharging functions return the empty set and each d_k ($n < k \leq n+m$) returns the singleton $\{\langle B_k,\, j_k\rangle\}$. We write the assumptions discharged by d_k within square brackets and put them over the k-th premiss.

Notation 2.3 For any application α of an n-ary inference rule ι, we write $\mathbf{d}_\iota^\alpha(\Gamma_1, \ldots \Gamma_n)$, in place of $\cup_{1 \leq k \leq n}(\Gamma_k \perp d_k(\iota)(\alpha))$. $\mathbf{d}_\iota^\alpha(\Gamma_1, \ldots \Gamma_n)$ is the set of undischarged assumptions after the application α of ι to n premisses depending on $\Gamma_1, \ldots, \Gamma_n$.

Remark 2.4 We can intuitively distinguish two kinds of inference rules:

- rules whose assumptions, premises and conclusions belong to the same language L_i. These rules, called *i-rules*, allow to draw consequences inside a theory;
- rules whose assumptions, premises or conclusions belong to distinct languages. These rules, called *bridge rules*, allow to export results from one or more theories to another.

Let us consider some examples of ML systems.

Example 2.1 Let ML_3 be $\mathrm{ML}_3 = \langle\{L_i\}_{1 \leq i \leq 3}, \{\Omega_i\}_{1 \leq i \leq 3}, \Delta\rangle$ where:

2 ML Systems

The goal of this section is to briefly introduce ML systems. More detailed discussions can be found in [Giu91, Giu92]. A formal system with multiple languages is a natural extension of the notion of axiomatic formal system. An axiomatic formal system S is usually described as a triple consisting of a language (which we take to be a set of *well formed formulas* or *wffs*), a set of axioms and a set of inference rules, *i.e.*, $S = \langle L, \Omega, \Delta\rangle$. The generalization is to take many languages and many sets of axioms while keeping one set of inference rules. We thus have the following definition:

Definition 2.1 (Multi Language System) *Let I be a set of indices, $\{L_i\}_{i\in I}$ a family of languages and $\{\Omega_i \subseteq L_i\}_{i\in I}$ a family of sets of wffs. A* Multi-Language Formal System (ML System) *MS is a triple $\langle \{L_i\}_{i\in I}, \{\Omega_i\}_{i\in I}, \Delta\rangle$ where $\{L_i\}_{i\in I}$ is the* Family of Languages, *$\{\Omega_i\}_{i\in I}$ is the* Family of Axioms *and Δ is the* Deductive machinery *of MS.*

Remark 2.1 A *family* is a set with repetitions. $\{L_i\}_{i\in I}$ and $\{\Omega_i\}_{i\in I}$ can be constructed as the codomain of two functions f_L and f_Ω, respectively, with domain I.

Remark 2.2 Two different wffs, might have the same intuitive meaning; *i.e.* they are two different ways of representing the same proposition. This happens also in a "one-language" system ($A \wedge B$, and $B \wedge A$ can be taken as an example). In an ML system it may also happen that two occurrences of the same wff (in two distinct languages) have different intuitive meanings.

Notation 2.1 For some $i, j \in I$, it might be the case that $L_i \cap L_j \neq \emptyset$. We write $\langle A, i\rangle$ to mean A and that A is a well formed formula of L_i. We say that $\langle A, i\rangle$ is a wff and that A is an L_i-wff. If G is a set of L_i-wffs for some $i \in I$, $\langle G, i\rangle$ denotes the set of wffs $G \times \{i\}$. $\mathbf{L}_I$ denotes the set of all wffs, namely $\cup_{i\in I}\langle L_i, i\rangle$.

Even if definition 2.1 is more general, allowing arbitrary formal languages and deductive machinery, we concentrate on first order languages and adopt a suitable modification of the natural deduction (ND) formalism defined in [Pra65].

The deductive machinery Δ is a set of inference rules. We define inference rules as a generalization of Prawitz's ND inference rules. We describe them as complex formal objects composed of three elements. A *relation* between premisses and conclusion which states which conclusions can be drawn from a given n-tupla of premisses. *n discharging functions*, one for each premiss, each computing which assumptions are discharged by the application of the rule. A *restriction relation* which defines the applicability of the rule.

Definition 2.2 (Inference rule) *An n-ary inference rule ι from $i_1, \ldots i_n$ to i, is composed of:*

(i) *an $n+1$-ary* relation*:*
$\rho(\iota)$ *is a recursive subset of* $(\langle L_{i_1}, i_1\rangle \times \ldots \times \langle L_{i_n}, i_n\rangle \times \langle L_i, i\rangle)$;

1 Introduction

[Giu91, Giu92] introduce and define a new kind of formal systems, called *MultiLanguage systems* (*ML systems*). ML systems are formal systems allowing multiple distinct languages, each language being associated with its own theory, and inference rules whose premises and consequences belong to distinct languages. [GS93] ([GS92] is a shorter version) introduces and defines the class $\mathcal{MR}$ of ML systems. The members of this class, called MR systems, are shown in [GS93] to be theoretically, representationally and implementationally adequate as the basic propositional systems for the formalization of theorem proving with metatheories and for the representation of propositional attitudes.

This paper is roughly divided in two parts. In the first part we develop the proof theory of the $\mathcal{MR}$ class. We concentrate on MK, one of the simplest but, at the same time, one of the most important $\mathcal{MR}$ systems. The results can be easily generalized to the other MR systems. We study the following issues:

Propagation of deductions through the languages. Deductions in ML systems span across languages. Some of the basic properties of MK concern how whole deductions can be moved across languages and how assumptions and conclusions can be shifted from a language to another.

Consistency. We provide a syntactic proof of the consistency of MK, and of other MR systems defined as extensions of MK. We show that inconsistency in MK behaves "unusually". For instance, it is impossible to achieve global inconsistency with a finite set of axioms.

Normal forms. We define three kinds of normal form. A *weak normal form* involving languages – this form minimizes the number of languages deductions go through. A *normal form* and a *strong normal form* involving wffs – these forms minimize the complexity of the wffs occurring in a deduction.

In the second part we study the relationship between MR systems and modal logics. We prove that the set of theorems of various modal logics can be embedded, under the standard bijective mapping between a modal and a first order language (mapping the modal operator into a unary predicate), into that of the corresponding MR systems.

The paper is structured as follows. Section 2 gives some basic notions concerning ML systems. Section 3 introduces the $\mathcal{MR}$ class, gives some important properties and defines some significative instances of MR systems. Section 4 proves some basic properties of MK. In particular it proves the deduction theorem, the cut rule and the weak normal form, which yields consistency. Section 5 proves the normal form, the strong normal form, the subformula property and some consequences. Section 6 shows how new ML systems can be defined by enlarging languages and adding axioms. Section 7 proves the equivalence between MK and modal K, defines new ML systems and proves their equivalence with various modal systems (*e.g.* T, K4, K45, S4, S5 and G). This section ends with the description of the axiom free versions of the ML systems considered.

Contents

Hierarchical Meta-Logics – Some Proof Theoretical Results *

Fausto Giunchiglia[1,2] Luciano Serafini[1]

[1]IRST – Istituto per la Ricerca Scientifica e Tecnologica
Povo, 38050 Trento, Italy
[2]University of Trento, Via Inama 5, 38100 Trento, Italy
fausto@irst.it, serafini@irst.it
phone: +39 461 814440, fax: +39 461 810851

Abstract

Multi-Language systems (*ML systems*) are formal systems allowing multiple distinct logical languages. In this paper we study an important class of ML systems, called $\mathcal{MR}$, such that its elements, called MR systems, are composed of a partially ordered set of first order languages, each language containing names and meta predicates for the language below. In the first part we develop the proof theory of a particular MR system called MK. The proof theoretic properties of the other MR systems can be easily deduced from MK's. In the second part we study the relationship between the $\mathcal{MR}$ class and modal logics. We prove that the set of theorems of various modal logics can be embedded, under the standard bijective mapping between a modal and a first order language, into that of the corresponding MR systems.

*This work has been done at IRST as part of the MAIA project (subproject No. 222B) (described in "An integrated approach to Artificial Intelligence" by L. Stringa, IRST technical report No 9012-11).

ISTITUTO PER LA RICERCA SCIENTIFICA E TECNOLOGICA
I 38100 TRENTO ⊥ LOC. PANTÉ DI POVO ⊥ TEL. 0461⊥314444
TELEX 400874 ITCRST ⊥ TELEFAX 0461⊥302040

HIERARCHICAL META-LOGICS
SOME PROOF THEORETICAL RESULTS

F. Giunchiglia, L. Serafini

March 1993
Technical Report # 9303-11

ISTITUTO TRENTINO DI CULTURA