A DESCENDING HIERARCHY OF REFLECTION PRINCIPLES

Albert Visser

Department of Philosophy
University of Utrecht

Logic Group
Preprint Series
No. 34

Department of Philosophy

University of Utrecht

A DESCENDING HIERARCHY OF REFLECTION PRINCIPLES

Albert Visser

Department of Philosophy
University of Utrecht

April 1988

Department of Philosophy
University of Utrecht
Heidelberglaan 2
3584 CS Utrecht
The Netherlands

A Descending Hierarchy of Reflection Principles

Albert Visser

1 Introduction

This paper can best be viewed as a portrait in miniature of a fascinating structure: a *descending* hierarchy of reflection principles. *Ascending* hierarchies of reflection principles are amply studied, e.g. in Feferman's great paper *Transfinite Recursive Progressions of Axiomatic Theories* (Feferman[1962]). The problem adressed in the study of ascending hierarchies is: what is the a priori implicit in mathematical reasoning? Ascending hierarchies can be viewed as developing by continuing reflection in idealised time. Descending hierarchies, in contrast, rest on presumption: in T_0 we claim to know the reflection principle for T_1, in T_1 we claim reflection for T_2, etcetera. Thus theories in a descending hierarchy are thoroughly *ungrounded*. Here we reach my motivation for writing this paper, which is: to provide a case-study of ungroundedness of provability-statements.

Ungroundedness is a notion which is much studied in connection with truth. The idea originated with Herzberger (see Herzberger[1970]). Kripke in his classical *Outline of a theory of truth* (Kripke[1975]) rightly stresses the importance of the notion, even if it turns out that there is more to the semantical paradoxes than ungroundedness. In his recent paper *Ungroundedness in classical languages* (McCarthy[1988]) Timothy McCarthy studies and defines ungroundedness also for e.g. arithmetical predicates that play the role of a truth-predicate (for example the well known Σ_1 truth-predicate for Σ_1-sentences). His definition involves an idealised evaluation procedure (as given in Kripke[1975]) for propositions involving truth. (In the paper McCarthy constructs a (classical model for a) descending hierarchy of truth-predicates.) There is no evident generalisation of McCarthy's work to proof, simply since we have no idealised evaluation procedure for propositions involving proof. This is, in my opinion, one of the great problems left open by the Metamathematical Tradition. Thus we are left with just our intuitions. For example I feel that the usual (literal) Henkin-sentence is true, but ungrounded, but to analyse this idea would be a subtle matter. With the descending hierarchy the ungroundedness is more evident: it can be made visible by looking for a possible justification of T_0. One would like to ask "How do you know reflection for T_1?", "Because I know reflection for T_2!","How do you know reflection for T_2?", etcetera (1).

Finally let me mention one important connection: a descending hierarchy of consistency statements has since long been known to logicians. My impression is that it was discovered independently by many people in the field. The hierarchy studied here is constructed in essentially the same way as this descending hierarchy. The consistency statements involved are all true. They serve as an important example in Provability Logic: they show that uniqueness and explicitness of Gödelian Fixed Points is lost as soon as one considers infinite sets of fixed point equations. The example shows at the same time the inadequacy of upwards wellfounded Kripke Models for Provability

Logic as soon as one turns to infinite sets of formulas.

2 Conventions

In this paper I try to keep the notational burdens light. Thus Quine quotes and numeral underlinings are left out at places where it is clear that they must be there anyway. For example I write $\mathrm{proof}_{PA}(x,A)$ instead of $\mathrm{proof}_{PA}(x,\ulcorner\underline{A}\urcorner)$. Free variables in formulas within prov are read according to the usual Kaplanesque convention for quantifying in: thus e.g. $\forall x\ \mathrm{prov}_{PA}(A(x))$ means (assuming that A contains only x free): $\forall x\ \mathrm{prov}_{PA}(\mathrm{sub}(x,\ulcorner\underline{x}\urcorner, \ulcorner\underline{A(x)}\urcorner))$. It is easily shown that this convention does no harm as long as we have only terms for provably recursive functions in our language (provably recursive, that is, in the weakest theory around, which is in our case PA).

RFN(W) is in this paper the uniform reflection principle for W. For example the instance of RFN(W) corresponding to A(x) with only x free is: $\forall x\ (\mathrm{prov}_W(A(x))\rightarrow A(x))$.

3 The descending hierarchy

We start with the definition of the Descending Hierarchy:

3.1 Definition

i) Find using the recursion theorem a sequence of theories T_n such that:

   T_n=PA+RFN(T_{n+1}) if for every $k\leq n$ not $\mathrm{Proof}_{T0}(k,\perp)$,

   T_n=PA otherwise.

   As is easily seen the axiom sets of the T_n can be chosen to be primitive recursive. We write $\Box_x A$ for the formalisation of: $T_x\vdash A$.

ii) Define $T_\omega:=\bigcap\{T_k|k\in\omega\}$ (The intersection here is intended as intersection of theorem sets). Let $\Box_\omega A$ be the formalisation of $T_\omega\vdash A$. Thus we have: $PA\vdash \forall A\ (\Box_\omega A\leftrightarrow \forall x\Box_x A)$.

iii) U_0:=PA, U_{n+1}:=PA+RFN(U_n). We write $\Delta_x A$ for the formalisation of $U_x\vdash A$.

iv) Define $U_\omega:=\bigcup\{U_k|k\in\omega\}$. Let $\Delta_\omega A$ be the formalisation of $U_\omega\vdash A$. Thus we have: $PA\vdash \forall A\ (\Delta_\omega A\leftrightarrow \exists x\Delta_x A)$.

In an ascending hierarchy like the U_k the union of the elements of the hierarchy is the natural target theory. In our descending hierarchy both T_0 and T_ω could be considered as target. T_0 is the strongest theory reached by the non-wellfounded sequence of theories. T_ω 'quantifies out' the 'unjustified' reflection principles of the theories above. As we will see this doesn't save T_ω from falsehood.

Our next step is to collect some basic facts:

3.2 Facts

(i) $PA\vdash \forall x,y,A\ ((x<y\wedge \Box_y A)\rightarrow \Box_x A)$.

(ii) $PA\vdash \forall x,y,A\ ((x<y\wedge \Delta_x A)\rightarrow \Delta_y A)$.

(iii) $PA+\mathrm{con}T_0 \vdash \forall x,y,A(\Delta_x A\rightarrow \Box_y A)$.

(iv) $PA+\mathrm{con}T_0 \vdash \forall A\ (\Delta_\omega A\rightarrow \Box_\omega A)$.

(v) $PA+\mathrm{incon}T_0\vdash \exists x\ (T_0=U_x\wedge T_x=U_0)$.

(vi) $PA+\mathrm{incon}T_0\vdash T_\omega=PA$.

Proof: (i), (ii), (v) and (vi) are left to the reader. (iv) is an immediate consequence of (iii). We prove (iii). Reason in $PA+\mathrm{con}T_0$:

We show by induction on x that for all x,y,A: $(\Delta_x A\rightarrow \Box_y A)$ and $\Delta_0(\Delta_x A\rightarrow \Box_y A)$. The case that $x=0$ is trivial. Let $x=z+1$. We find by the Induction Hypothesis: $\Box_y(\Delta_z B\rightarrow \Box_{y+1}B)$, hence $\Box_y(\Delta_z B\rightarrow B)$. Thus T_y proves the axioms of U_x, so: $(\Delta_x A\rightarrow \Box_y A)$ and $\Delta_0(\Delta_x A\rightarrow \Box_y A)$. □

In our first theorem we prove two equiconsistency results (over PA) for T_0 and T_ω.

3.3 Theorem

(i) $PA\vdash \mathrm{con}T_0\leftrightarrow \mathrm{con}U_\omega$

(ii) $PA\vdash \mathrm{con}T_\omega\leftrightarrow \mathrm{con}PA$

Proof: To prove (i) it is sufficient to show: $PA\vdash \mathrm{con}T_0\leftrightarrow \forall x\ \mathrm{con}U_x$. Reason in PA:

"$\rightarrow$" Immediate by 3.2(iii).

"$\leftarrow$" Suppose $\mathrm{incon}T_0$. Then for some z $T_0=U_z$. Hence for some z $\mathrm{incon}U_z$. □(PA)

Note that 3.3(i) implies that T_0 is really consistent. To prove (ii): clearly: $PA+\mathrm{con}T_0\vdash \mathrm{con}PA$, $PA+\mathrm{con}T_0\vdash \mathrm{con}T_\omega$, hence $PA+\mathrm{con}T_0\vdash \mathrm{con}T_\omega\leftrightarrow \mathrm{con}PA$. On the other hand $PA+\mathrm{incon}T_0\vdash T_\omega=PA$ and so: $PA+\mathrm{incon}T_0\vdash \mathrm{con}T_\omega\leftrightarrow \mathrm{con}PA$. □

I feel 3.3(ii) is rather surprising, given the fact that 3.2(iv) and 3.3(i) imply that: $U_\omega\subseteq T_\omega$.

In 3.4-3.7 we show that T_ω is false: it proves a false Σ_1-sentence.

3.4 Definition

Find by the Gödel Diagonalisation Lemma a formula G(x) such that:

$$PA\vdash G(x)\leftrightarrow \exists y\geq x\neg\Box_y G(y+1).$$

3.5 Theorem

$PA+conT_0 \vdash \forall y \Box_y G(y+1)$.

Proof: reason in $PA+conT_0$:

Consider T_y. Because $conT_0$, we have: $T_y=PA+RFN(T_{y+1})$. Reason in T_y:

Suppose $\neg G(y+1)$, i.e. $\forall z \geq y+1 \Box_z G(z+1)$ (*). By 3.2(i): $\forall z \geq y+1 \Box_{y+1} G(z+1)$. Thus by $RFN(T_{y+1})$: $\forall z \geq y+1 G(z+1)$ and so: $\forall z \geq y+1 \exists u \geq z+1 \neg \Box_u G(u+1)$ (**). By (**) there is a $u \geq y+1$ with $\neg \Box_u G(u+1)$. But by (*): $\Box_u G(u+1)$. Contradiction. Conclude: $G(y+1)$. □

3.6 Theorem

$PA+conT_0 \vdash \forall x \neg G(x)$.

Proof: immediate from 3.5. □

3.7 Theorem

$PA \vdash \Box_\omega inconT_0$.

Proof: By 3.5 $PA+conT_0 \vdash \forall y \Box_y G(y+1)$. The formalisation of 3.6 plus contraposition gives us: $PA \vdash \forall y \Box_y (\exists z G(z) \rightarrow inconT_0)$. Combining we find: $PA+conT_0 \vdash \forall y \Box_y inconT_0$. On the other hand it is immediate by Σ-completeness that $PA+inconT_0 \vdash \forall y \Box_y inconT_0$. □

Note that 3.7 implies that each T_n 'thinks' itself to be one of the U_k-s, without knowing which one.

We have seen that $U_\omega \subseteq T_\omega$ and that $T_\omega \vdash inconT_0$. How much more does T_ω prove compared to U_ω? Note that T_ω is prima facie Π_2 and that U_ω is Σ_1. We show that T_ω is complete Π_2 and thus proves much more than U_ω.

3.8 Lemma

$PA+conT_0 \vdash \forall x,A \Box_x \forall y>x(\Box_y A \rightarrow A)$.

Proof: this is immediate from 3.2(i). □

3.9 Theorem

Let $S(x,y)$ be any Σ_1-formula, then there is a Σ_1-formula $R(x,y)$, such that:

$PA+conT_0 \vdash \forall x(\forall y S(x,y) \leftrightarrow \Box_\omega \forall z R(x,z))$.

Proof: Find by the Gödel Diagonalosation Lemma a Σ_1-formula $R(x,y)$ such that:

$$PA\vdash R(x,y) \leftrightarrow (S(x,y)\vee\Box_0\bot)\leq\Box_yR(x,y).$$

Reason in PA+conT_0:

First suppose $S(x,y)$ and $\neg R(x,y)$. It follows that $(\Box_yR(x,y))<(S(x,y)\vee\Box_0\bot)$ (*). From (*) we have on the one hand: $\Box_yR(x,y)$, on the other: $\Box_y((\Box_yR(x,y))<(S(x,y)\vee\Box_0\bot))$ (**) (by Σ-completeness). (**) yields $\Box_y\neg R(x,y)$. Conclude $\Box_y\bot$ and hence $\Box_0\bot$, quod non. So we find: $R(x,y)$.

Now suppose $\forall v\leq yS(x,v)$. It follows that $\forall v\leq yR(x,v)$ and hence by Σ-completeness: $\Box_y\forall v\leq yR(x,v)$. We show that also: $\Box_y\forall v>yR(x,v)$. Reason in T_y:

Suppose $v>y$ and $\neg R(x,v)$. Since $\Box_0\bot$, it follows that $(\Box_vR(x,v))<(S(x,v)\vee\Box_0\bot)$ and hence $\Box_vR(x,v)$. By lemma 3.8, we find: $R(x,v)$. Contradiction. Conclude: $R(x,v)$. $\Box(T_y)$

Conclude: $\Box_y\forall zR(x,z)$.

From the above we have: if $\forall yS(x,y)$, then $\forall y\,\Box_y\forall zR(x,z)$.

For the converse assume that for some y $\neg S(x,y)$. It is clearly sufficient to show $\neg\Box_yR(x,y)$. Suppose $\Box_yR(x,y)$. We have: $\Box_yR(x,y)$, $\neg S(x,y)$ and $\neg\Box_0\bot$. Hence: $(\Box_yR(x,y))<(S(x,y)\vee\Box_0\bot)$. So by Σ-completeness: $\Box_y((\Box_yR(x,y))<(S(x,y)\vee\Box_0\bot))$. Conclude: $\Box_y\neg R(x,y)$, ergo $\Box_y\bot$ and hence $\Box_0\bot$, quod non. $\Box$

3.10 Questions

(i) By the Gödel Diagonalisation Lemma there are formulas $H_0(x)$ and $H_1(x)$ such that:

$$PA\vdash H_0(x)\leftrightarrow \exists y\geq x\Box_yH_0(y+1),$$
$$PA\vdash H_1(x)\leftrightarrow \forall y\geq x\Box_yH_1(y+1).$$

Describe the behaviour of H_0 and H_1.

(ii) What happens if we restrict ourselves to Σ-reflection in the definition of the T_k? (It is easy to see that restriction to Π-reflection gives us theories that are all true!)

NOTES

1) Of course a formal proof is just a syntactical object. A sentence is provable in a formal system or it isn't. There can be no question of ungroundedness here. Intuitions of ungroundedness only emerge when we try to see proofs in a formal system as the forms in which we try to articulate the gaining of insights. The crucial notion of conceptual priority or 'being created earlier in conceptual time' really applies not to the syntactical objects -they are just there- but to the real proofs of which they are the forms.

REFERENCES

Feferman, S., 1962, Transfinite Recursive Progressions of Axiomatic Theories, *JSL 27,* 259-316.

Herzberger, H.G., 1970, Paradoxes of Grounding in Semantics, *The Journal of Philosophy 67,* 145-167.

Kripke, S., 1975, Outline of a Theory of Truth, *The Journal of Philosophy 72,* 690-716.

McCarthy, T., 1988, Ungroundedness in Classical Languages, *JPL 17,* 61-74.

Smorynski, C., 1985, *Self-Reference and Modal Logic,* Springer, New York.

Logic Group Preprint Series
Department of Philosophy
University of Utrecht
Heidelberglaan 2
3584 CS Utrecht
The Netherlands

nr. 1 C.P.J. Koymans, J.L.M. Vrancken, *Extending Process Algebra with the empty process*, September 1985.
nr. 2 J.A. Bergstra, *A process creation mechanism in Process Algebra,* September 1985.
nr. 3 J.A. Bergstra, *Put and get, primitives for synchronous unreliable message passing,* October 1985.
nr. 4 A.Visser, *Evaluation, provably deductive equivalence in Heyting's arithmetic of substitution instances of propositional formulas*, November 1985.
nr. 5 G.R. Renardel de Lavalette, *Interpolation in a fragment of intuitionistic propositional logic,* January 1986.
nr. 6 C.P.J. Koymans, J.C. Mulder, *A modular approach to protocol verification using Process Algebra* , April 1986.
nr. 7 D. van Dalen, F.J. de Vries, *Intuitionistic free abelian groups,* April 1986.
nr.8 F. Voorbraak, *A simplification of the completeness proofs for Guaspari and Solovay's R*, May 1986.
nr.9 H.B.M. Jonkers, C.P.J. Koymans & G.R. Renardel de Lavalette, *A semantic framework for the COLD-family of languages,* May 1986.
nr.10 G.R. Renardel de Lavalette, *Strictheidsanalyse,* May 1986.
nr.11 A. Visser, *Kunnen wij elke machine verslaan? Beschouwingen rondom Lucas' argument*, July 1986.
nr.12 E.C.W. Krabbe, *Naess's dichotomy of tenability and relevance,* June 1986.
nr.13 Hans van Ditmarsch, *Abstractie in wiskunde, expertsystemen en argumentatie,* Augustus 1986
nr.14 A.Visser, *Peano's Smart Children, a provability logical study of systems with built-in consistency* , October 1986.
nr.15 G.R.Renardel de Lavalette, *Interpolation in natural fragments of intuitionistic propositional logic,* October 1986.
nr.16 J.A. Bergstra, *Module Algebra for relational specifications,* November 1986.
nr.17 F.P.J.M. Voorbraak, *Tensed Intuitionistic Logic,* January 1987.
nr.18 J.A. Bergstra, J. Tiuryn, *Process Algebra semantics for queues,* January 1987.
nr.19 F.J. de Vries, *A functional program for the fast Fourier transform,* March 1987.
nr.20 A. Visser, *A course in bimodal provability logic,* May 1987.
nr.21 F.P.J.M. Voorbraak, *The logic of actual obligation, an alternative approach to deontic logic,* May 1987.
nr.22 E.C.W. Krabbe, *Creative reasoning in formal discussion,* June 1987.
nr.23 F.J. de Vries, *A functional program for Gaussian elimination*, September 1987.
nr.24 G.R. Renardel de Lavalette, *Interpolation in fragments of intuitionistic propositional logic,* October 1987.(revised version of no. 15)
nr.25 F.J. de Vries, *Applications of constructive logic to sheaf constructions in toposes*, October 1987.
nr.26 F.P.J.M. Voorbraak, *Redeneren met onzekerheid in expertsystemen*, November 1987.
nr.27 P.H. Rodenburg, D.J. Hoekzema, *Specification of the fast Fourier transform algorithm as a term rewriting system,* December 1987.

nr.28 D. van Dalen, *The war of the frogs and the mice, or the crisis of the Mathematische Annalen*, December 1987.

nr.29 A. Visser, *Preliminary Notes on Interpretability Logic*, January 1988.

nr.30 D.J. Hoekzema, P.H. Rodenburg, *Gauß elimination as a term rewriting system*, January 1988.

nr. 31 C. Smorynski, *Hilbert's Programme*, January 1988.

nr. 32 G.R. Renardel de Lavalette, *Modularisation, Parameterisation, Interpolation*, January 1988.

nr. 33 G.R. Renardel de Lavalette, *Strictness analysis for POLYREC, a language with polymorphic and recursive types*, March 1988.

nr. 34 A. Visser, *A Descending Hierarchy of Reflection Principles*, April 1988.