In: Catarina Dutilh Novaes & Ole Thomassen Hjortland (eds.),
Insolubles and Consequences. Essays in Honour of Stephen Read.
London: College Publications 2012, pp. 203-211.
(Published online with permission of College Publications)

Paradoxes and Structural Rules

Peter Schroeder-Heister[†] [‡]

ABSTRACT. The derivation of many paradoxes can be blocked, if the application of structural rules is locally restricted in certain ways. This is shown independently for identity, contraction, and cut. Imposing such local restrictions is seen as an alternative to the global rejection of structural rules (notably contraction), which is no reasonable option given that structural rules are needed in mathematical reasoning.

Since the work of Fitch (1936) it is well known that the rule of contraction is crucial for the derivation of certain paradoxes. On the other hand, contraction is needed to formalize mathematical (and other) reasoning, so that giving up contraction altogether is not a viable way of avoiding them (see Read, 1994, pp. 162f, quoted from the 1995 paperback edition). This suggests to look for a restriction on contraction which is sufficient to block paradoxes without affecting 'normal' (non-paradoxical) reasoning. In this note we indicate what such a restriction might look like. As not only *contraction*, but also *identity* and *cut* are needed in paradoxical derivations, we extend our investigation to them. For each of these three structural rules, we propose provisos which formally prevent the derivation of paradoxes, without unduly restraining them as key principles of reasoning.

As our structural framework we use an intuitionistic sequent calculus whose sequents have the form $\Gamma \vdash C$ or the form $\Gamma \vdash$, where the antecedent Γ is a multiset of formulas and the succedent is either a formula C or empty. We use notations such as $\Gamma, A \vdash C$ or $A_1, \ldots, A_n \vdash C$ in the usual way. The empty succedent is supposed to represent falsity, and the empty antecedent is supposed to represent truth. The empty sequent $\vdash$ thus represents a contradiction. In our simplified framework we only need negation $\neg$ as a logical constant. In order to represent the paradoxes, we suppose that there is a constant R such that from $\neg R$ we can infer R and from R we can infer $\neg R$. As we are dealing only with the structural aspect of the logic of paradoxes, we are not interested in which way R is construed and how these paradoxical inference rules are obtained. In naive set theory R might be construed as the proposition $\{x : x \notin x\} \in \{x : x \notin x\}$, and the paradoxical rules are obtained from certain set theoretical principles. In a theory of definitional reflection R might be just a constant

[†]Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, 72076 Tübingen, Germany. Email: psh@uni-tuebingen.de.

[‡]This work goes back to investigations on definitional reasoning, which was carried out together with Lars Hallnäs. The preparation of this paper was supported by the French-German ANR-DFG project "Hypothetical Reasoning — Its Proof-Theoretic Analysis" (HYPOTHESES) (DFG Schr 275/16-2).

given by the definitional clause $R \Leftarrow \neg R$, and the paradoxical rules are obtained as rules of definitional closure and definitional reflection with respect to this definition (Hallnäs, 1991; Hallnäs and Schroeder-Heister, 1990/91; Schroeder-Heister, 1993; Schroeder-Heister, 2012a). In a theory of arbitrary propositional operators R might be a nullary logical constant for which these paradoxical rules are the introduction and elimination rules (Tennant, 1982; Read, 2010).

Thus our system contains the following rules:

Structural rules

Identity: $\dfrac{}{A \vdash A}\,(\mathrm{Id})$ Contraction: $\dfrac{\Gamma, A, A \vdash C}{\Gamma, A \vdash C}\,(\mathrm{Contr})$

Cut: $\dfrac{\Gamma \vdash A \qquad A, \Delta \vdash C}{\Gamma, \Delta \vdash C}\,(\mathrm{Cut})$ Weakening: $\dfrac{\Gamma \vdash C}{\Gamma, A \vdash C}$ $\dfrac{\Gamma \vdash}{\Gamma \vdash A}$

Negation rules

$$\frac{\Gamma \vdash A}{\Gamma, \neg A \vdash}\,(\neg\vdash) \qquad \frac{\Gamma, A \vdash}{\Gamma \vdash \neg A}\,(\vdash\neg)$$

Paradoxical rules

$$\frac{\Gamma, \neg R \vdash C}{\Gamma, R \vdash C}\,(R\vdash) \qquad \frac{\Gamma \vdash \neg R}{\Gamma \vdash R}\,(\vdash R)$$

Then the derivation of contradiction looks as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{R \vdash R}\,(\mathrm{Id})}{R, \neg R \vdash}\,(\neg\vdash)}{R, R \vdash}\,(R\vdash)}{R \vdash}\,(\mathrm{Contr})}{\vdash \neg R}\,(\vdash\neg)}{\vdash R}\,(\vdash R) \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{R \vdash R}\,(\mathrm{Id})}{R, \neg R \vdash}\,(\neg\vdash)}{R, R \vdash}\,(R\vdash)}{R \vdash}\,(\mathrm{Contr})}{\vdash}\,(\mathrm{Cut}) \tag{1}$$

This derivation uses all rules of our system with the exception of weakening. If a proper paradox is expected to allow the generation of any sequent whatsoever, we could add applications of weakening at the end of this derivation to obtain $\Gamma \vdash C$ for any Γ and C. However, as these applications would not affect the actual derivation of the empty sequent, we do not deal with weakening here. The legitimacy of weakening is a central matter in investigations of paraconsistency, which is a different way of dealing with the paradoxes, since it does not challenge the derivability of contradictions.

1 Identity

Obviously, dropping the rule of identity would block the derivation of the paradox. In fact, it would prevent us from starting any derivation whatsoever, as

identity is the only axiom available. This is why identity is also called the rule of "initial sequents". However, there is a reasonable restriction on identity which can block the paradox without affecting 'standard' derivations.

Proviso: Identity $A \vdash A$ may only be used if no right- and left-introduction rules for A are available.

The rationale behind this restriction is the following. Identity permits the introduction of a formula A in an *unspecific* way: No matter what A looks like, we can start with A as an assertion, which depends on itself as an assumption. However, there might be *specific* ways to introduce A, which depend on the meaning of A. The proviso then says that we must use these specific ways, if they are available. For example, if A has the form $\neg B$, then, rather than starting with $\neg B \vdash \neg B$ as an identity axiom, we request that $\neg B$ be introduced according to its specific meaning, *i.e.*, according to the specific introduction rules for $\neg$, yielding

$$\dfrac{\dfrac{B \vdash B}{B, \neg B \vdash}\,(\neg\vdash)}{\neg B \vdash \neg B}\,(\vdash\neg)$$

In the standard sequent calculus, this restriction corresponds to the common proviso that identity sequents must be atomic, so that nonatomic sequents of the form $A \vdash A$ can only be derived using the specific rules governing the constants occurring in A. From a more philosophical point of view, we might consider the distinction between atomic and non-atomic sentences to be the distinction between sentences for which no specific meaning rules are given in the system, and those for which specific meaning rules are available. The proviso then says that identity only applies to the unspecific case, whereas the specific case is always handled by the meaning rules. As in our context specific rules are available for R, R is not considered atomic in this sense, so identity cannot be used for R. The sequent $R \vdash R$ should rather be reduced to $\neg R \vdash \neg R$ via the derivation

$$\dfrac{\dfrac{\neg R \vdash \neg R}{R \vdash \neg R}\,(R\vdash)}{R \vdash R}\,(\vdash R)$$

However, since $\neg R$ is not atomic either, this derivation must be reduced to

$$\dfrac{\dfrac{\dfrac{\dfrac{R \vdash R}{R, \neg R \vdash}\,(\neg\vdash)}{\neg R \vdash \neg R}\,(\vdash\neg)}{R \vdash \neg R}\,(R\vdash)}{R \vdash R}\,(\vdash R)$$

leading us to what we started with and showing that we cannot initiate the paradoxical derivation.

Conclusion: There is a plausible restriction of identity which blocks the paradoxes without affecting non-paradoxical reasoning. In the presence of restricted identity, the rules of contraction and cut can be used without any restriction.[1]

[1] This way of restricting identity was originally developed in the context of logic program-

2 Contraction

Disallowing contraction would prevent the derivation of the paradox. However, in view of the frequent use of contraction in mathematical and other reasoning, this cannot be considered a viable strategy. By adopting our distinction between specific and unspecific ways of using a sentence, we can argue for a restriction on contraction which blocks the paradox without precluding sensible uses of contraction.

If we look at the premiss $R, R \vdash$ of the application of contraction in (1), we observe that the two occurrences of R result from the application of different rules. The left occurrence of R in $R, R \vdash$ comes from the identity axiom $R \vdash R$. It may therefore be called an *unspecific occurrence* of R, since identity applies unspecifically to any formula independent of its shape. The right occurrence of R in $R, R \vdash$ is the result of an introduction rule for R. Since this introduction rule is specific for R, this occurrence may be called a *specific* occurrence of R. By boxing the unspecific occurrences of R and encircling the specific occurrences of R, this situation may be depicted as follows:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\boxed{R} \vdash R}\,(\mathrm{Id})}{\boxed{R}, \neg R \vdash}\,(\neg\vdash)}{\boxed{R}, \textcircled{R} \vdash}\,(R\vdash)}{R \vdash}\,(\mathrm{Contr})}{\vdash \neg R}\,(\vdash\neg)}{\vdash R}\,(\vdash R) \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{\boxed{R} \vdash R}\,(\mathrm{Id})}{\boxed{R}, \neg R \vdash}\,(\neg\vdash)}{\boxed{R}, \textcircled{R} \vdash}\,(R\vdash)}{R \vdash}\,(\mathrm{Contr})}{\vdash}\,(\mathrm{Cut})
$$

The contraction step identifies the boxed R with the encircled R. If we argue that it makes a difference of whether we use R unspecifically (in an identity sequent) or specifically (according to its meaning, as the result of a left-introduction rule), we are led to the following restriction of contraction.

Proviso: Contraction may only be used if the two occurrences of the sentence A which are contracted into a single one, are either both specific or both unspecific. We must never contract a specific with an unspecific occurrence of A.

The significance of this proposal rests of course on the claim that the 'standard' uses of contraction (in mathematics and elsewhere) do *not* use the form of contraction prohibited by the proviso. This needs still to be verified — here we

ming with definitional reflection by Kreuger (1994). This is why we call it "Kreuger's rule" or "Kreuger's axiom". In logic programming with definitional reflection we extend logics by inference rules governing atoms which, as the common logical rules, split into right-introduction and left-introduction rules. This means that these atoms are atomic with respect to logic, but not atomic in the sense of being irreducible, because meaning-giving rules are associated with them. Therefore it is not unnatural to require that the unspecific initial sequents $A \vdash A$ be allowed only in the case in which A is not further reducible to any other sentence. Formally, it can be shown that by restricting initial sequents in the way indicated, the resulting system allows the elimination of cuts and is therefore consistent. For the simplified system considered above this is easy to show, but this result holds for the general system with definitional rules for atoms as well. See Schroeder-Heister (1994).

take it for granted. Secondly, the notion of unspecific vs. specific occurrences must be made precise in such a way that paradoxes are indeed excluded by observing the proviso.

When working out this idea in detail, it turns out that the distinction between specific and unspecific occurrences of assumptions is not clear enough as it stands. An unspecific assumption can easily be turned into a specific one by using intermediate derivation steps. For example, if the critical part

$$\begin{array}{l}
\dfrac{}{\boxed{R} \vdash R}\ (\mathrm{Id}) \\[2ex]
\dfrac{\boxed{R} \vdash R}{\boxed{R}, \neg R \vdash}\ (\neg\vdash) \\[2ex]
\dfrac{\boxed{R}, \neg R \vdash}{\boxed{R}, \textcircled{R} \vdash}\ (R\vdash) \\[2ex]
\dfrac{\boxed{R}, \textcircled{R} \vdash}{R \vdash}\ (\mathrm{Contr})
\end{array}$$

of (1) is replaced with

$$\begin{array}{l}
\dfrac{}{R \vdash R}\ (\mathrm{Id}) \\[2ex]
\dfrac{R \vdash R}{R, \neg R \vdash}\ (\neg\vdash) \\[2ex]
\dfrac{R, \neg R \vdash}{\neg R \vdash \neg R}\ (\vdash\neg) \\[2ex]
\dfrac{\neg R \vdash \neg R}{\neg R \vdash R}\ (\vdash R) \\[2ex]
\dfrac{\neg R \vdash R}{\textcircled{R} \vdash R}\ (R\vdash) \\[2ex]
\dfrac{\textcircled{R} \vdash R}{\textcircled{R}, \neg R \vdash}\ (\neg\vdash) \\[2ex]
\dfrac{\textcircled{R}, \neg R \vdash}{\textcircled{R}, \textcircled{R} \vdash}\ (R\vdash) \\[2ex]
\dfrac{\textcircled{R}, \textcircled{R} \vdash}{R \vdash}\ (\mathrm{Contr})
\end{array}$$

i.e., if a derivation of $R \vdash R$ is added on top, then the boxed occurrences of R become encircled, as they result from an introduction of R according to the rule $(R\vdash)$. This means that both occurrences of R, which are contracted, are specific occurrences of R and thus appear to satisfy the proviso.

To deal with such cases, we propose to introduce an indexing discipline. With every formula occurrence in a derivation, a natural number is associated as a *meaning index*, which is increased if a meaning rule (left- or right-introduction rule) is applied. Contraction is then prohibited, if the meaning indices of the sentences involved differ. In the above example, the right R undergoing contraction receives a higher meaning index under this discipline than the left one, as it results from a single application of $(\vdash R)$ plus a single application of $(R\vdash)$, whereas the left R results from a single application of $(R\vdash)$ alone. The result is a kind of type theory with respect to meaning, according to which the application of a meaning rules increases the type. Attaching these indices for R as superscripts yields

$$
\begin{array}{c}
\hline R \vdash R \quad (\mathrm{Id}) \\
\hline R, \neg R \vdash \quad (\neg\vdash) \\
\hline \neg R \vdash \neg R \quad (\vdash\neg) \\
\hline \neg R \vdash R^1 \quad (\vdash R) \\
\hline R^1 \vdash R^1 \quad (R\vdash) \\
\hline R^1, \neg R^1 \vdash \quad (\neg\vdash) \\
\hline R^1, R^2 \vdash \quad (R\vdash) \\
\hline R \vdash \quad (\mathrm{Contr})
\end{array}
$$

demonstrating that contraction equalizes occurrences of R with unequal index.

Conclusion: There is a plausible restriction of contraction which blocks the paradoxes without affecting non-paradoxical reasoning. In the presence of restricted contraction, the rules of identity and cut can be used without any restriction.[2]

3 Cut

Derivation (1) uses cut in its last step. Without cut the empty sequent is not derivable. It is a characteristic feature of paradoxes that derivations generating them do not permit the elimination of cut. This corresponds to the fact that in natural deduction, derivations of paradoxes are not normalisable, as Prawitz (1965, Appendix B) has first observed. Therefore disallowing cut blocks the derivation of the paradox. If we insist on using cut-free systems, we can discard with the paradoxes.

However, insisting on cut-free systems (or, equivalently, on systems in which cut is admissible) asks for too much. In advanced mathematical theories, in fact: already in arithmetic, we do need cut to carry out significant proofs.[3] It is therefore worthwhile to look for a restriction of cut, which prevents paradoxes but still allows for standard mathematical reasoning.

In order to formulate such a restriction, we propose to use a type system in the style of the Curry-Howard correspondence, but adapted to the sequent calculus. A sequent $A_1, \ldots, A_n \vdash A$ now takes the form $x_1 : A_1, \ldots, x_n : A_n \vdash t : A$ for variables $x_1, \ldots, x_n$ and a term t. Terms are generated using certain constructors and selectors which depend on the constants available, in our case case just negation and R. Furthermore, certain reduction principles for terms are available which tell which terms are considered equal. Terms which cannot be further reduced are considered normal. Intuitively, normal terms codify 'real' proofs, whereas non-normal terms denote the proofs codified by their normal forms (if they exist and are unique). When using term annotations, we cannot deal with empty succedents, so we use the falsity constant $\bot$ instead.

[2]The idea of restricting contraction in this way was proposed in Schroeder-Heister (2004).

[3]This frequently happens, for example, when, in order to prove a statement A, we prove by induction a stronger statement A' because A' is needed as the induction hypothesis. Specialising A' to A means to apply a cut which is not eliminable.

In our context it is of particular importance to see that cut becomes a substitution rule:

$$\frac{\Gamma \vdash t : A \qquad \Delta, x : A \vdash s : B}{\Gamma, \Delta \vdash s[x/t] : B}$$

When performing this substitution, it may happen that the newly created term $s[x/t]$ is not normal and not normalisable, even though the terms s and t are in normal form. This is exactly what is taking place in the paradoxical derivation. In the term-annotated system the last step of (1) takes the form

$$\frac{\vdash t : R \qquad x : R \vdash s : \bot}{\vdash s[x/t] : \bot}$$

where the term $s[x/t]$ generated for absurdity is not normalisable, while s and t are normal. (For more details see the appendix.) This leads to the following restriction on the form of cut:

Proviso: Cut may only be used if it does not create, by means of substitution, a non-normalisable term.

The idea behind this proviso is that we should only generate terms which represent "real" proofs. Denoting by $t\downarrow$ the fact that t is normalisable, we would then formulate the restricted cut rule as

$$\frac{\Gamma \vdash t : A \qquad \Delta, x : A \vdash s : B}{\Gamma, \Delta \vdash s[x/t] : B}\ s[x/t]\downarrow$$

This way of proceeding presupposes, of course, that we use a type system and no longer just operate with the sentences (= types) alone. Furthermore, which is more involved, we must specify how to establish $t\downarrow$ for a term t, as this is not a syntactic property of the given proof figure or of the formal system used. So the restricted form of cut should more appropriately be formulated as

$$\frac{\Gamma \vdash t : A \qquad \Delta, x : A \vdash s : B \qquad s[x/t]\downarrow}{\Gamma, \Delta \vdash s[x/t] : B}$$

with $s[x/t]\downarrow$ as a premiss of its own. In the end this leads to a sort of free type theory where one has to add rules by means of which we can prove that a term denotes.

Conclusion: There is a plausible restriction of cut which blocks the paradoxes without affecting non-paradoxical reasoning. In the presence of restricted cut, the rules of identity and contraction can be used without any restriction.[4]

4 Summary

For each of the structural rules of identity, contraction and cut, we considered certain restricted forms which suffice for 'ordinary' mathematical reasoning,

[4]This restriction of cut was proposed in Schroeder-Heister (2012b). There the fact that the sequent calculus makes the use of substitution explicit in the form of the cut rule, is seen as a conceptual advantage of this format of deductive reasoning as compared to natural deduction.

but which are insufficient for the derivation of paradoxes. Even though we did not decisively argue in favour of a particular one of these restrictions, this shows that investigating constraints on the application of structural rules is a promising research programme.

Appendix

In the term-annotated system, where we use the absurdity constant $\bot$ to represent the empty succedent, it is more convenient to consider $\neg A$ to be an abbreviation for $A \to \bot$. The term-annotated rules for implication and R are as follows.

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B} \qquad \frac{\Gamma \vdash t : A}{\Gamma, x : A \to B \vdash App(x, t) : B}$$

together with the reduction principle

$$App(\lambda x.t, s) \rhd t[x/s] \tag{2}$$

which is the same as β-contraction, and

$$\frac{\Gamma \vdash t : R \to \bot}{\Gamma \vdash rt : R} \qquad \frac{\Gamma, x : R \to \bot \vdash t : C}{\Gamma, y : R \vdash t[x/r'y] : C}$$

together with the reduction principle

$$r'rt \rhd t \ . \tag{3}$$

With term-annotations the derivation (1) takes the form

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{x : R \vdash x : R}{x : R, y : R \to \bot \vdash App(y, x) : \bot}}{x : R, z : R \vdash App(r'z, x) : \bot}}{x : R \vdash App(r'x, x) : \bot}}{\vdash \lambda x.App(r'x, x) : R \to \bot}}{\vdash r\lambda x.App(r'x, x) : R} \qquad \dfrac{\dfrac{\dfrac{x : R \vdash x : R}{x : R, y : R \to \bot \vdash App(y, x) : \bot}}{x : R, z : R \vdash App(r'z, x) : \bot}}{x : R \vdash App(r'x, x) : \bot}}{\vdash App(r'r\lambda x.App(r'x, x), r\lambda x.App(r'x, x)) : \bot} \tag{4}$$

It can easily be checked that the term generated for absurdity in the end sequent of (4) reduces to itself, *i.e.* the reduction sequence based on reductions (2) and (3) (there is only a single such sequence) loops. For further discussion see Schroeder-Heister (2012b).

BIBLIOGRAPHY

Fitch, F. B. (1936). A system of formal logic without an analogue to the Curry W Operator. *Journal of Symbolic Logic*, 1(3):92–100.

Hallnäs, L. (1991). Partial inductive definitions. *Theoretical Computer Science*, 87:115–142.

Hallnäs, L. and Schroeder-Heister, P. (1990/91). A proof-theoretic approach to logic programming: I. Clauses as rules. II. Programs as definitions. *Journal of Logic and Computation*, 1:261–283, 635–660.

Kreuger, P. (1994). Axioms in Definitional Calculi. In Dyckhoff, R., editor, *Extensions of Logic Programming. 4th International Workshop, ELP'93 (St. Andrews, U.K., March/April 1993). Proceedings (Lecture Notes in Computer Science, Vol. 798)*, pages 196–205. Springer, Berlin.
Prawitz, D. (1965). *Natural Deduction: A Proof-Theoretical Study.* Almqvist & Wiksell, Stockholm.
Read, S. (1994). *Thinking About Logic.* Oxford University Press, Oxford.
Read, S. (2010). General-elimination harmony and the meaning of the logical constants. *Journal of Philosophical Logic*, 39(5):557–576.
Schroeder-Heister, P. (1993). Rules of definitional reflection. In *Proceedings of the 8th Annual IEEE Symposium on Logic in Computer Science (Montreal 1993)*, pages 222–232. IEEE Press, Los Alamitos.
Schroeder-Heister, P. (1994). Cut Elimination for Logics with Definitional Reflection and Restricted Initial Sequents. Unpublished manuscript. (See the author's homepage).
Schroeder-Heister, P. (2004). On the Notion of Assumption in Logical Systems. In Bluhm, R. and Nimtz, C., editors, *Selected Papers Contributed to the Sections of GAP5, Fifth International Congress of the Society for Analytical Philosophy, Bielefeld, 22-26 September 2003*, pages 27–48. Mentis, Paderborn. `www.gap5.de/proceedings`.
Schroeder-Heister, P. (2012a). Proof-theoretic semantics. In Zalta, E., editor, *Stanford Encyclopedia of Philosophy.* `http://plato.stanford.edu`.
Schroeder-Heister, P. (2012b). Proof-theoretic semantics, self-contradiction, and the format of deductive reasoning. *Topoi*, 31(1):77–85.
Tennant, N. (1982). Proof and paradox. *Dialectica*, 36:265–296.