

On the Logic of Category Definitions

Marcus Kracht
Mathematisches Institut
FU Berlin
1000 Berlin 33, Germany

February 7, 2008

In their paper on category structures, [Gazdar *et al.*, 1988] define a constraint language L_C for categories and a logic Λ_C of admissible category structures.¹

The intuitive idea is that for a constraint ϕ expressed in L_C , ϕ is a **nontrivial** constraint if and only if $\Lambda_C \not\models \phi$; and it is a **satisfiable** constraint if and only if $\Lambda_C \not\models \neg\phi$. From a practical point of view it is therefore important to know whether Λ_C is decidable and even better that the decision can be given in a time bounded by a recursive function on the length of ϕ . However, the remarks made in their paper only suffice to show that the modal fragment of Λ_C ² contains **S4.Grz** = $\mathsf{K}(\Box p \rightarrow p, \Box p \rightarrow \Box\Box p, \Box(\Box(p \rightarrow \Box p) \rightarrow p) \rightarrow p)$, which does not show that this fragment is decidable. In this note, I will establish both that the modal fragment of Λ_C and Λ_C itself are decidable, and I will prove it in that order. As a result, I will also axiomatize Λ_C . Thus I show first that the modal reduct of Λ_C , which I call Λ_M , is decidable. This paper will be rather hard going for anyone not acquainted with modal logic. We advise the reader to have [Gazdar *et al.*, 1988] at hand while reading this paper, or better still, to read it once through beforehand. For the modal logics we refer the reader to [Boolos, 1994], [Harel, 1984], and [Segerberg, 1971], but in principle any introduction to modal logic will provide enough background

¹Unfortunately, they do not distinguish between the language L_C and the logic, which defines a subset of that language, namely the set of its theorems. We make this distinction here by calling the logic as well as the set of theorems it defines Λ_C .

²We define a logic as a set of rules, which are pairs Δ/ϕ , where Δ is the set of premises of that rule and ϕ its consequence. Modus Ponens thus takes the form $\phi \rightarrow \psi, \psi/\phi$. Rules are closed under substitution. Axioms are rules because we can take $\Delta = \emptyset$. The modal fragment of Λ_C is then simply the subpart of rules that only involve modal formulas, i.e. no type 1 features.

to be able to understand the gist of the arguments. Without going into too many details of the construction, I will show that there is an easy way to give a proof that in fact shows that $\Lambda_C = \text{S4.Grz}$ using the structure of the models those logics admit. Intuitively, categories correspond to Kripke models. For let α be a category. Then α defines a set of categories W , which is obtained by successively applying type 1 features to α .³ An accessibility relation \triangleleft is defined via $\alpha \triangleleft \beta$ iff $f(\alpha) = \beta$ for some type 1 feature f . This accessibility relation is irreflexive, intransitive, finite, and defines a tree-structure on W . Most importantly, it is cycle free. Thus, if we look at the reflexive and transitive closure \triangleleft^* of \triangleleft , it is again finite and has no non-trivial cycles. It therefore is an S4.Grz -structure (see, e.g., [Boolos, 1994]). Conversely, an S4.Grz -structure $\langle W, \triangleleft \rangle$ which is a tree can be represented as a category. If we then take a model $\langle W, \triangleleft, \text{val} \rangle$ based on that frame, where $\text{val} : X \rightarrow \wp(W)$ maps a finite set of propositional variables into $\wp(W)$, we can code this model by adding a type 0 feature f_p for each $p \in X$ that takes values \top or \perp . Thus the resulting category α not only codes the successor function by means of type 1 features, but also the valuation val . W is in one-to-one correspondence ϕ with the set Γ of categories generated by α . We then make the following definitions:

Let β be in Γ :

1. $\beta \vDash f_p : \top$ iff $\phi(\beta) \in \text{val}(p)$ iff $\phi(\beta) \vDash p$
2. $\beta \vDash f_p : \perp$ iff $\phi(\beta) \in \text{val}(p)$ iff $\phi(\beta) \vDash \neg p$

It is easy to see that for any modal formula P with variables in X , the corresponding translation τ induced by $\mapsto f_p : \top$ satisfies $\beta \vDash \tau(P)$ iff $\phi(\beta) \vDash P$. The logic of $\langle W, \triangleleft^* \rangle$ therefore coincides with the logic of all categories that differ from α only in the assignment of type 0 features. To conclude, the logic of categories as defined in [Gazdar *et al.*, 1988] coincides with the logic of all finite, reflexive, transitive trees. It is easily seen that the finite, reflexive, transitive trees generate the class of finite models for S4.Grz . Thus the logic of categories is the logic of the finite models of S4.Grz which, since S4.Grz has the finite model property, is identical to S4.Grz (*end of proof*)

A few remarks are in order:

1. I used a purely semantical argument, which in this case is the most direct way, because it is fairly easy to see why we get just the models we get, though there is some footwork to be done.

³Remember that type 1 features take propositions as values, whereas type 0 features only take atoms.

2. Alternatively I could have built a canonical model out of a category structure Σ , whose worlds are the categories that Σ admits and whose accessibility relation is as defined above for categories. The proof is essentially the same.
3. The idea of encoding frames and valuations into a single structure has also been explored in [Fagin and Vardi, 1985].
4. In [Rautenberg, 1983] a simple tableau calculus for S4.Grz is given which shows that S4.Grz consistency is effectively decidable, and that the decision procedure is primitive recursive. Furthermore, the size of a tableau is bounded by a function of the number $\sigma(P)$ of subformulas of P , or, more precisely, the theoremhood of P can be decided with a tableau of length $\leq 2^{7+6\sigma(P)}$. Given the proof, the same holds for Λ_M , since the translation procedure reduces the size of a formula. So we have the same bound for Λ_C .
5. In [Gazdar *et al.*, 1988] another logic is mentioned which arises from restricting the number of type 1 features to 1. The resulting logic is equal to S4.3.Grz = S4.Grz($\diamond p \wedge \diamond q. \rightarrow .\diamond(p \wedge \diamond q) \vee \diamond(q \wedge \diamond p) \vee \diamond(p \wedge q)$), the logic of all finite linear orders, as can be seen in the same way. Since finitely generated S4.3.Grz models are finite, this logic is decidable as well.

I will now proceed to the full case. Before I embark on the proof, let me remark on a few things. First, although each particular category structure contains only a finite number of features and values, L_C contains infinitely many of them. As regards the type 0 features, this causes no problem, since we treat $f : a$ as a proposition and we allow ourselves infinitely many of those. However, type 1 features will create some problems that are not very serious but have to be dealt with carefully. Second, as we defined a translation of L_C into modal logic, we will now define a translation of L_C into elementary propositional dynamic logic (EPDL) so that every type 1 feature has a program associated with it whose interpretation is an accessibility relation between categories. This translation is harmless and allows us to forget about type 0 features altogether.

1. $\tau(f : \alpha) = p_{\langle f:a \rangle}$ where f is of type 0
2. $\tau(f_i : \phi) = \langle \gamma_i \rangle \tau(\phi)$ where f_i is of type 1
3. $\Box \phi = [\alpha] \phi$

α is a program that by definition contains all other programs; that is, if two categories are related by γ_i , they are also related by α . However, although the intuition is that α is the reflexive and transitive closure of all the γ_i s this fact is not expressible in EPDL nor in L_C because it requires a formula of infinite length. But, as it turns out, this is a harmless deficiency of our language. The image of Λ_C under the translation into EPDL will be called Ξ .

I will now give a full axiomatization of Ξ . As we saw earlier, the axioms governing the behaviour of $[\alpha]$ are exactly the axioms for S4.Grz, since $[\alpha]$ is the old \Box . Similar reasoning will reveal that the $[\gamma_i]$ behave alike, and the corresponding logic is the logic otherwise known as K.Alt₁, which stands for "only one alternative". The accessibility relation for K.Alt₁ allows a world to have at most one successor. Although it has more models, K.Alt₁ is the logic of all irreflexive, linear and finite frames, which shows that it is the logic we are looking for. If we take all this together with the observation that α includes γ_i , we get the following axioms for Ξ :

$$\begin{array}{ll}
 \text{Grz}_\alpha & \Xi \vdash [\alpha]([\alpha](\phi \rightarrow [\alpha]\phi) \rightarrow \phi)\phi \\
 (1) \quad \text{Alt}_{1_i} & \Xi \vdash \langle \gamma_i \rangle (\phi \wedge \psi) \leftrightarrow (\langle \gamma_i \rangle \phi \wedge \langle \gamma_i \rangle \psi) \\
 \text{Mix}_i & \Xi \vdash \langle \gamma_i \rangle \rightarrow \langle \alpha \rangle
 \end{array}$$

Note that Ξ is not finitely axiomatisable and so Λ_C isn't either. Note also that if it weren't for the axiom(s) called Mix_i, life would be very easy for us now. Since the axioms for the various programs are independent, the finite model property for each of those programs individually would yield the finite model property for the whole logic by simple induction on the number of programs. Thus let us call the logic without Mix_i Ξ^\ominus . Also, since we have a tableau calculus for S4.Grz and a tableau calculus for K.Alt₁, we have a calculus for Ξ^\ominus as well, simply by putting all tableau rules together. The tableau rule for γ_i would look as follows: from $\Gamma; \langle \gamma_i \rangle \phi$ step to $\Gamma^\diamond; \phi$, where $\Gamma^\diamond := \{\psi : \langle \gamma_i \rangle \psi \in \Gamma \text{ or } [\gamma_i] \in \Gamma\}$. I shall spare the reader an exact specification of the tableau rules and refer him to [Rautenberg, 1983] again. Note that the length of the tableaux for K.Alt₁ is bounded by the size of ψ so that the actual size of the tableau is at most $2^{n\sigma(\phi)}$, where n is the length of ϕ and $\sigma(\phi)$ the number of subformulas of ϕ . This bound could be sharpened somewhat but we ignore this point. What I want to show is how the fact that Ξ^\ominus can be shown to have all those properties can be made to explain why Ξ must have those properties, too. What the reader should understand at this point is that tableaux are a way of systematically constructing a model for a formula (if it is consistent) and showing inconsistency by exhausting all possible choices, of which there are only finitely many.

What I do now is boost up a model construction procedure for Ξ^\ominus to a model construction procedure for Ξ . This is done as follows: Suppose we want to construct a Ξ -model for ϕ . Since we do not know how to do this we construct a Ξ^\ominus -model instead. However, this model might be deficient by not respecting Mix_i . Therefore we add a finite set ϕ^\sharp of instances of Mix_i which will ensure that Mix_i is respected for subformulas of ϕ . The Ξ^\ominus -model can then safely be turned into a Ξ -model. Let me therefore define the modal degree d of a formula.

$$\begin{aligned}
d(\phi) &= 0, \text{ if } \phi \text{ is a propositional variable or constant} \\
d(\neg\phi) &= d(\phi) \\
(2) \quad d(\phi \wedge \psi) &= \max(d(\phi), d(\psi)) \\
d([\alpha]\phi) &= d(\phi) + 1 \\
d([\gamma_i]\phi) &= d(\phi) + 1
\end{aligned}$$

Furthermore, let (ϕ) be the set of subformulas of ϕ . Then define

$$\begin{aligned}
(3) \quad \zeta &:= \bigcup \langle \gamma_i : \gamma_i \text{ occurs in } \phi \rangle \cup \alpha \\
\zeta^1 &:= \zeta \\
\zeta^{n+1} &:= \zeta; \zeta^n \\
\phi^\sharp &:= \{ [\zeta^{d(\phi)}] (\langle \gamma_i \rangle \chi \rightarrow \langle \alpha \rangle \chi : \chi \in (\phi), \gamma_i \text{ occurs in } \phi) \}
\end{aligned}$$

Then we say that a Ξ -tableau for ϕ is simply a Ξ^\ominus -tableau for $\phi; \phi^\sharp$. Let us see how ϕ^\sharp makes everything right for us. The failure of Ξ^\ominus is to allow models for $\langle \gamma_i \rangle \psi \wedge [\alpha] \neg \psi$. For suppose we build a simple Ξ^\ominus -tableau for ϕ , and we encounter a line $\Gamma; \langle \gamma_i \rangle \psi \wedge [\alpha] \neg \psi$. In the next step we get $\Gamma; \langle \gamma_i \rangle \psi; [\alpha] \neg \psi$ and then $\Gamma^\diamond; \psi$. But if we added ϕ^\sharp then Γ would necessarily contain a formula yielding $\langle \gamma_i \rangle \psi \rightarrow \langle \alpha \rangle \psi$, which would close this branch of the tableau. Thus a Ξ^\ominus -tableau for $\phi; \phi^\sharp$ results in a model $\langle W, \triangleleft, \text{val} \rangle$ in which, though α is an independent program, for every subformula ψ of ϕ , if $s, \text{val} \vDash \langle \gamma_i \rangle \psi$, then also $s, \text{val} \vDash \langle \alpha \rangle \psi$. Thus it is easy to see that if we now reinterpret the α relation \triangleleft_α as the reflexive, transitive closure of α and the γ_i , we get a Ξ -model $\langle W, \triangleleft, \text{val} \rangle$ for ϕ , which obviously is of the same size. Thus, if Ξ^\ominus has the finite model property, Ξ also has the finite model property, and if the decision procedure for Ξ^\ominus is bounded a priori by a function on the length of ϕ , the same holds for Ξ . Obviously, the bound is much higher than for Ξ because of the ϕ^\sharp , but this is the price we have to pay. (*end of proof*)

Given the above proof we can now change the tableau calculus for Ξ by redefining the rule for γ_i to the following: from $\Gamma; \langle \gamma_i \rangle \psi$ infer $\Gamma^\diamond; \Gamma^\square; \psi$ where

$\Gamma^\diamond := \{\psi : \langle \gamma_i \rangle \in \Gamma \text{ or } [\gamma_i] \in \Gamma\}$ and $\Gamma^\square := \{[\alpha]\psi : [\alpha]\psi \in \Gamma\} \cup \{\psi : [\alpha]\psi \in \Gamma\}$. The proof method I used for proving decidability for $\Xi(\Lambda_C)$ from the decidability Ξ^\ominus is explained in full detail in [Kracht, 1993]. The same method applies to the case when we only allow a single type I feature, since it only requires the base logic—in this case **S4.3.Grz**—to have the finite model property. The resulting logic **$\Xi.3.\text{Rec}$** can be (finitely) axiomatized as follows:

$$\begin{array}{ll}
\text{Grz}_\alpha & \Xi.3 \vdash [\alpha]([\alpha](\phi \rightarrow [\alpha]\phi) \rightarrow \phi) \rightarrow \phi \\
\mathbf{3}_\alpha & \Xi.3 \vdash (\langle \alpha \rangle \phi \wedge \langle \alpha \rangle \psi) \rightarrow \langle \alpha \rangle (\phi \wedge \langle \alpha \rangle \psi) \vee \langle \alpha \rangle (\psi \wedge \langle \alpha \rangle \phi) \\
& \vee \langle \alpha \rangle (\phi \wedge \psi) \\
(4) \quad \text{Alt}_1, 1 & \Xi.3 \vdash \langle \gamma_1 \rangle \phi \wedge \langle \gamma_1 \rangle \psi \leftrightarrow \langle \gamma_1 \rangle (\phi \wedge \psi) \\
\text{Mix}_1 & \Xi.3 \vdash \langle \gamma_1 \rangle \rightarrow \langle \alpha \rangle \phi \\
\text{Rec} & \Xi.3 \vdash \langle \alpha \rangle \phi \rightarrow \phi \vee \langle \gamma_1 \rangle \langle \alpha \rangle \phi
\end{array}$$

The model construction requires some care since we do not have an unlimited resort of extra features, but it can be done in the same spirit. This shows decidability for **$\Xi.3.\text{Rec}$** and hence for the corresponding logic mentioned in the paper. An alternative formulation of Ξ could be given with the help of propositional dynamic logic using the star operator $*$. We would then simply have an axiom $\langle \gamma_1^* \rangle \phi \leftrightarrow \langle \alpha \rangle \phi$, that says nothing else but that α is the reflexive and transitive closure of γ_1 . This trick would also work if we restrict Ξ to any finite number of features. But the star does not gain us much for Ξ itself since we still could not express the fact that α is the reflexive and transitive closure of the γ_i s because we have infinitely many of them.

ACKNOWLEDGMENTS

This paper was written while I was at the Centre for Cognitive Science in Edinburgh. I wish to thank Jaap van der Does for encouraging me to write this proof down and for proofreading it. I also wish to thank G. Gazdar for remarks on an earlier version of the paper and an anonymous referee for further suggestions.

REMARKS (FEBRUARY 2008)

The original version contains a mistake in the axiomatisation of the 1-feature version of Ξ . Ξ cannot express the fact that α is the closure of γ_i . However, if we restrict the number of features we can. The added axiom **3** is not enough since it still does not guarantee that. We need to add the recursion axiom **Rec** (which makes **3** redundant in presence of **Alt**₁ for γ_1). The proof method (reduction of consequence relations) is best explained in [Kracht, 1999], not in the paper

Kracht (1988), which later appeared as [Kracht, 1993]. The proof of the finite model property of Ξ is pretty terse (like most of that paper).

References

- [Boolos, 1994] George Boolos. *The Logic of Provability*. Cambridge University Press, Cambridge, 1994.
- [Fagin and Vardi, 1985] Ron Fagin and Moshe Vardi. An Internal Semantics for Modal Logics: Preliminary Report. Technical report, 1985. CSLI-Report No. 85-25.
- [Gazdar *et al.*, 1988] G. Gazdar, G. K. Pullum, R. Carpenter, E. H. Klein, T. E. Hukari, and R. D. Levine. Category structures. *Journal of Computational Linguistics*, 14-1:1–14, 1988.
- [Harel, 1984] David Harel. Dynamic logic. In Dov M. Gabbay and Franz Guenther, editors, *Handbook of Philosophical Logic*, volume 2. Reidel, 1984.
- [Kracht, 1993] Marcus Kracht. Splittings and the finite model property. *The Journal of Symbolic Logic*, 58:139–157, 1993.
- [Kracht, 1999] Marcus Kracht. *Tools and Techniques in Modal Logic*. Number 142 in Studies in Logic. Elsevier, Amsterdam, 1999.
- [Rautenberg, 1983] Wolfgang Rautenberg. Modal tableau calculi and interpolation. *Journal of Philosophical Logic*, 12:403–423, 1983.
- [Seegerberg, 1971] Krister Segerberg. An essay in classical modal logic. Mimeo-graph, Uppsala, 1971.