Refutation of a modal logic system for reasoning about the degree of blameworthiness

Abstract: We evaluate a modal logic system for the blameable coalition of an outcome if there is a strategy to prevent it and where the degree of blameworthiness is measured by costs of prevention or sacrifice. Of eight axioms, three are *not* tautologous, hence refuting the approach. We do not consider the claimed technical result of a completeness theorem.

We assume the method and apparatus of Meth8/VŁ4 with Tautology as the designated *proof* value, **F** as contradiction, N as truthity (non-contingency), and C as falsity (contingency). For results, the 16-valued truth table is row-major and horizontal, or repeating fragments of 128-tables, sometimes with table counts, for more variables. (See ersatz-systems.com.)

LET p , q , r , s , t ,
lc_phi ϕ, lc_psi ψ, Blameable, $^{\text{degree}}$, cost,
u , v , w , x, y, z:
$_{C}$ coalition (small), $_{D}$ coalition (large), Statement (N), x, y, z;
~ Not, ¬ ; + Or, ∨, ∪; - Not Or; & And, ∧, ∩; \ Not And;
> Imply, greater than, →, ⊢, ↦; < Not Imply, less than, ∈;
= Equivalent, ≡, ⊨, :=, ⟺, ↔; @ Not Equivalent, ≠;
% possibility, for one or some, ∃, ◊, M; # necessity, for every or all, ∀, □, L;
(z=z) T as tautology; (z@z) **F** as contradiction, ∅, Null;
(%z<#z) C as contingency, Δ, ordinal 1;
(%z>#z) N as non-contingency, ∇, ordinal 2;
~(y < x) (x ≤ y), (x ⊆ y).

From: Cao, R.; Naumov, P. (2019). The limits of morality in strategic games.
arxiv.org/pdf/1901.08467.pdf rui.cao@alumni.ubc.ca, pgn2@cornell.edu

3 Axioms: In addition to the propositional tautologies in language Φ, our logical system contains the following axioms:

1. Truth:

$N\phi \rightarrow \phi$ and $B^{s}_{C}\phi \rightarrow \phi$, (1.1)

((w&p)>p)&(((r&s)&(u&p))>p) ; TTTT TTTT TTTT TTTT (1.2)

2. Distributivity:

$N(\phi \rightarrow \psi) \rightarrow (N\phi \rightarrow N\psi)$, (2.1)

(w&(p>q))>((w&p)>(w&q)) ; TTTT TTTT TTTT TTTT (2.2)

3. Negative Introspection:

$\neg N\phi \rightarrow N\neg N\phi$, (3.1)

~(w&p)>(w&~(w&p)) ;

TTTT TTTT TTTT TTTT(8), **FFFF FFFF FFFF FFFF**(8) (3.2)

4. None to Blame:

$\neg B^{s}_{\varnothing}\phi$, (4.1)

~((r&s)&((z@z)&p))=(p=p); TTTT TTTT TTTT TTTT (4.2)

5. Monotonicity:

$B^{s}_{C}\phi \rightarrow B^{t}_{D}\phi$, where $C \subseteq D$ and $s \leq t$, (5.1)

(~(v<u)&~(t<s))>(((r&s)&(u&p))>((r&t)&(u&p))) ;
TTTT TTTT TTTT TTTT(2), TTTT TTTT TTTT T**F**TT(1),
TTTT TTTT TTTT TTTT(3), TTTT TTTT TTTT T**F**TT(1),
TTTT TTTT TTTT TTTT(1) (5.2)

6. Joint Responsibility:

if $C \cap D = \varnothing$, then $NB^{s}_{C}\phi \wedge NB^{t}_{D}\psi \rightarrow (\phi \vee \psi \rightarrow B^{s+t}_{C\cup D}(\phi \vee \psi))$, (6.1)

((u&v)=(z@z))>
(((~w&((r&s)&(u&p)))&(~w&((r&t)&(v&q))))>((p+q)>
((r&((s+t)&(u+v)))&(p+q)))) ; TTTT TTTT TTTT TTTT (6.2)

7. Blame for Cause:

$N(\phi \rightarrow \psi) \rightarrow (B^{s}_{C}\psi \rightarrow (\phi \rightarrow B^{s}_{C}\phi))$, (7.1)

(w&(p>q))>(((r&s)&(u&q))>(p>((r&s)&(u&p)))) ;
TTTT TTTT TTTT TTTT (7.2)

8. Fairness:

$B^{s}_{C}\phi \rightarrow N(\phi \rightarrow B^{s}_{C}\phi)$. (8.1)

((r&s)&(u&p))>(w&(p>((r&s)&(u&p)))) ;
TTTT TTTT TTTT TTTT(2), TTTT TTTT TTTT T**F**T**F**(2),
TTTT TTTT TTTT TTTT(2), TTTT TTTT TTTT T**F**T**F**(2),
TTTT TTTT TTTT TTTT(8) (8.2)

For Eqs. 1.2-8.2 as rendered, the three 3.2, 5.2, and 8.2 are not tautologous, hence refuting the proposed system. The claimed technical result of a completeness theorem is not evaluated.