Normalization for Cubical Type Theory

Jonathan Sterling   Carlo Angiuli
Carnegie Mellon University

LICS '21

□TT: cubical type theory

□TT is an extension of *Martin-Löf's Intensional Type Theory* by an interval:

- a new sort $\Gamma \vdash \mathbb{I}$ and context extension $\Gamma, i : \mathbb{I} \longrightarrow \Gamma$
- with endpoints $\Gamma \vdash 0, 1 : \mathbb{I}$
- and potentially further structure: $r \sqcup s, r \sqcap s, \sim r$ [Coh+17]

Why? A new way to think about equality (paths) as *figures* of shape $\mathbb{I}$.

$$(a =_A b) \coloneqq \{p : \mathbb{I} \to A \mid p(0) \equiv a \land p(1) \equiv b\}$$

Supports function extensionality, type extensionality (univalence), and effective quotients.

Computation in □TT

Unlike HoTT, cubical type theory has good computational properties.

Theorem (Cubical canonicity [AFH18; Hub18])

If $\mathbb{I}^n \vdash M : \mathsf{bool}$ *is a closed n-cube of booleans, then either* $\mathbb{I}^n \vdash M \equiv \mathsf{tt} : \mathsf{bool}$ *or* $\mathbb{I}^n \vdash M \equiv \mathsf{ff} : \mathsf{bool}$*.*

Therefore **□TT** can be used as a programming language [Ang+21], and we have multiple implementations, *e.g.* Cubical Agda, `redtt`, `cooltt` [Red18; Red20; VMA19].

Canonicity is only about computation in *purely cubical* contexts $i, j, k : \mathbb{I}$.
Implementation requires computation in *arbitrary* contexts, *i.e.* normalization.

In what contexts do we compute?

$$\Psi \ \Vdash \ M : A \ \Downarrow \ V$$

In what contexts do we compute?

element

$$\Psi \Vdash M : A \Downarrow V$$

In what contexts do we compute?

element

$$\Psi \Vdash M : A \Downarrow V$$

observation

In what contexts do we compute?

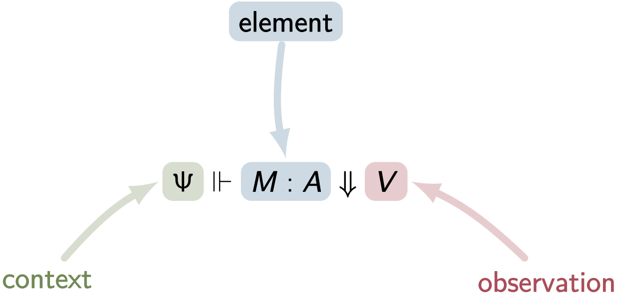
element

$$\Psi \Vdash M : A \Downarrow V$$

context

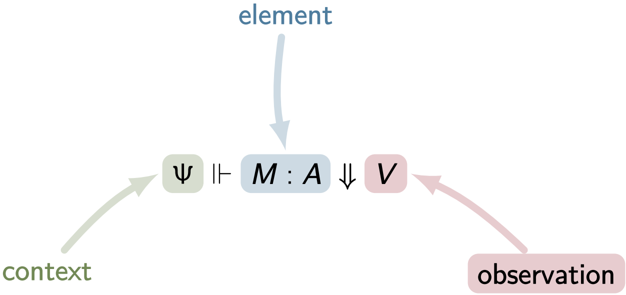
observation

In what contexts do we compute?



canonicity: $A \in \{\mathsf{nat}\}$; **normalization:** $A \in \{\Psi \vdash \mathit{type}\}$

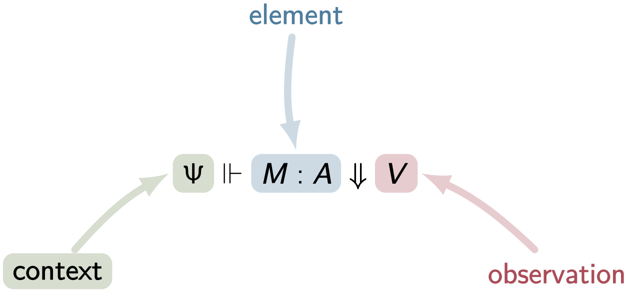
In what contexts do we compute?



canonicity: $A \in \{\mathsf{nat}\}$; **normalization:** $A \in \{\Psi \vdash \mathit{type}\}$

canonicity: $V \in \mathbb{N}$; **normalization:** $V \in \{\Psi \Vdash^{\beta\eta}_{\mathbf{nf}} A\}$

In what contexts do we compute?



canonicality: $A \in \{\mathrm{nat}\}$; **normalization:** $A \in \{\Psi \vdash \mathit{type}\}$

canonicity: $V \in \mathbb{N}$; **normalization:** $V \in \left\{\Psi \Vdash^{\beta\eta}_{\mathbf{nf}} A\right\}$

canonicity: $\Gamma \in \{\cdot\}$; **cubical canonicity:** $\Gamma \in \{\mathbb{I}^n \mid n \in \mathbb{N}\}$; **normalization:** $\Gamma \in \{\vdash \mathit{ctx}\}$

Stability of observation *sans* $\mathbb{I}$

$$x : \mathsf{nat} \Vdash x : \mathsf{nat} \Downarrow \mathfrak{var}(x)$$

Stability of observation *sans* $\mathbb{I}$

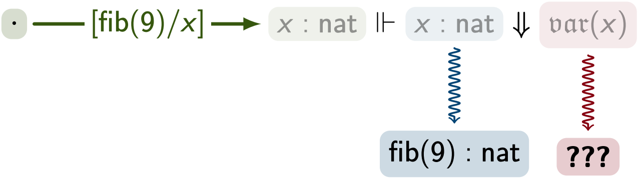
$$\cdot \xrightarrow{[\mathsf{fib}(9)/x]} x : \mathsf{nat} \Vdash x : \mathsf{nat} \Downarrow \mathfrak{var}(x)$$

Stability of observation *sans* $\mathbb{I}$

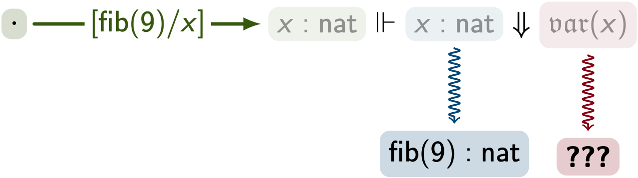
$$\cdot \xrightarrow{[\mathsf{fib}(9)/x]} x : \mathsf{nat} \Vdash x : \mathsf{nat} \Downarrow \mathfrak{var}(x)$$

$$x : \mathsf{nat} \rightsquigarrow \mathsf{fib}(9) : \mathsf{nat}$$

Stability of observation *sans* $\mathbb{I}$

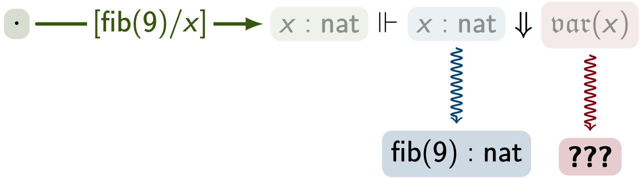


Stability of observation *sans* $\mathbb{I}$



In **ITT**, neutral observations (elimination forms blocked on variables) are closed under *renaming*, but not full substitution.

Stability of observation *sans* $\mathbb{I}$



In **ITT**, neutral observations (elimination forms blocked on variables) are closed under *renaming*, but not full substitution.

Therefore, computation takes place in the Artin gluing of $\mathrm{Pr}(\mathcal{C}) \longrightarrow \mathrm{Pr}(\mathcal{R})$ where $\mathcal{R} : \mathbf{Cat}_{/\mathcal{C}}$ is category of contexts and **renamings** (*c.f.* Kripke logical relations).

Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for $\square\mathbf{TT}$. The problem is the interval:

Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for □TT. The problem is the interval:

$$p : \mathsf{fib} =_{\mathsf{nat}\to\mathsf{nat}} \mathsf{fib}, i : \mathbb{I} \Vdash (p \mathbin{@} i)\, 9 : \mathsf{nat} \Downarrow \mathfrak{app}(\mathfrak{papp}(\mathfrak{var}(p), i), \mathfrak{su}^{9}(\mathfrak{ze}))$$

Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for **□TT**. The problem is the interval:

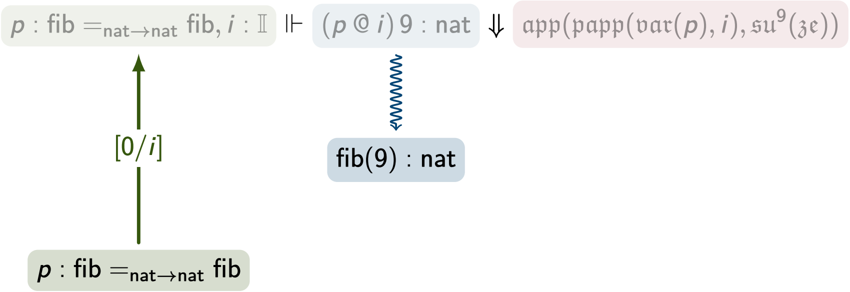
$$p : \mathsf{fib} =_{\mathsf{nat}\to\mathsf{nat}} \mathsf{fib}, i : \mathbb{I} \Vdash (p \mathbin{@} i)\, 9 : \mathsf{nat} \Downarrow \mathfrak{app}(\mathfrak{papp}(\mathfrak{var}(p), i), \mathfrak{su}^9(\mathfrak{ze}))$$

$$\uparrow [0/i]$$

$$p : \mathsf{fib} =_{\mathsf{nat}\to\mathsf{nat}} \mathsf{fib}$$

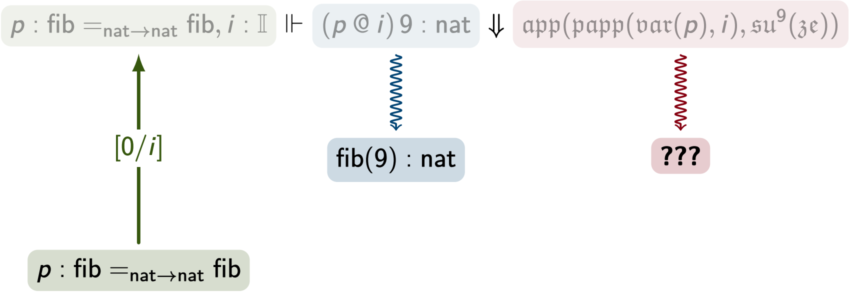
Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for **□TT**. The problem is the interval:



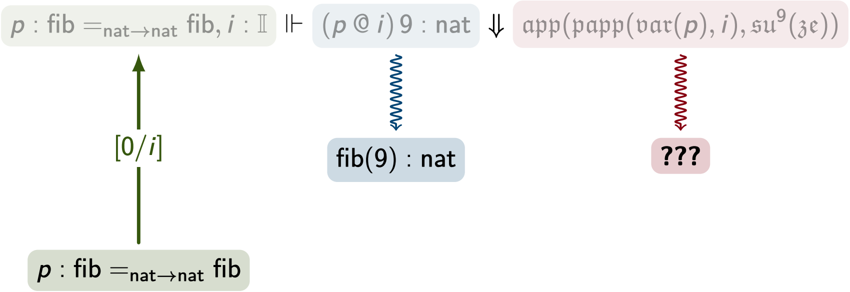
Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for □TT. The problem is the interval:



Instability of observation *avec* $\mathbb{I}$

Unfortunately, removing the substitutions for which neutral observations are unstable is not possible for □TT. The problem is the interval:



We cannot remove $[0/i], [1/i]$ from the category of contexts and renamings because we need $\mathbb{I}$ to restrict to something *representable* in $\mathrm{Pr}(\mathcal{R})$, *c.f.* **tininess** criterion [Lic+18].

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease 'being neutral'.

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease ‘being neutral’.

Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease ‘being neutral’.

Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

$$\partial(\mathfrak{var}(x)) = \bot$$

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease 'being neutral'.

Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

$$\partial(\mathfrak{var}(x)) = \bot$$
$$\partial(\mathfrak{app}(E, M)) = \partial E$$

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease ‘being neutral’.

Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

$$\begin{aligned}
\partial(\mathfrak{var}(x)) &= \bot \\
\partial(\mathfrak{app}(E, M)) &= \partial E \\
\partial(\mathfrak{fst}(E)) &= \partial E
\end{aligned}$$

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease 'being neutral'.

Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

$$\begin{aligned}\partial(\mathfrak{var}(x)) &= \bot\\ \partial(\mathfrak{app}(E, M)) &= \partial E\\ \partial(\mathfrak{fst}(E)) &= \partial E\\ \partial(\mathfrak{papp}(E, r)) &= \partial E \vee (r = 0) \vee (r = 1)\end{aligned}$$

The power of dialectical thinking

Thesis: neutrals need to have a cubical substitution action.

Antithesis: positive neutrality is not a cubical notion: under face maps $[\epsilon/i]$ a neutral observation can cease 'being neutral'.

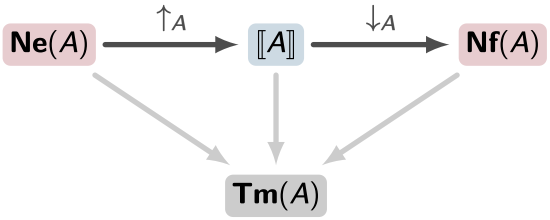
Synthesis: the conditions under which a given neutral **de**stabilizes *are* cubical. Given a neutral form $e : A$, write ∂e for this *frontier of instability*.

$$\begin{aligned}
\partial(\mathfrak{var}(x)) &= \bot \\
\partial(\mathfrak{app}(E, M)) &= \partial E \\
\partial(\mathfrak{fst}(E)) &= \partial E \\
\partial(\mathfrak{papp}(E, r)) &= \partial E \vee (r = 0) \vee (r = 1)
\end{aligned}$$

Therefore, we define an inductive family $\mathbf{Ne}_\phi(A) \rightarrowtriangle \mathbf{Tm}(A)$ of neutrals e with $\partial e = \phi$. Traditional neutrals = $\mathbf{Ne}_\bot(A)$; to model destabilization, $\mathbf{Ne}_\top(A) \cong \mathbf{Tm}(A)$.

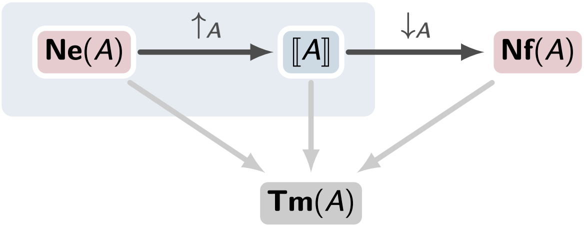
The standard **Tait yoga**

Tait [Tai67] introduced the famous *saturation yoga* for normalization.



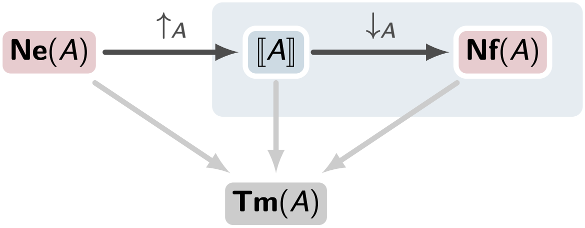
The standard **Tait yoga**

Tait [Tai67] introduced the famous *saturation yoga* for normalization.



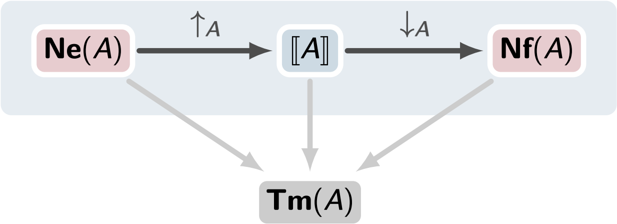
The standard **Tait yoga**

Tait [Tai67] introduced the famous *saturation yoga* for normalization.



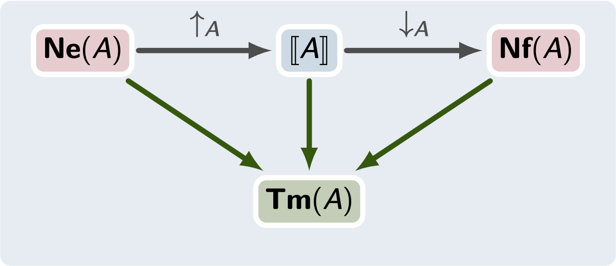
The standard **Tait yoga**

Tait [Tai67] introduced the famous *saturation yoga* for normalization.

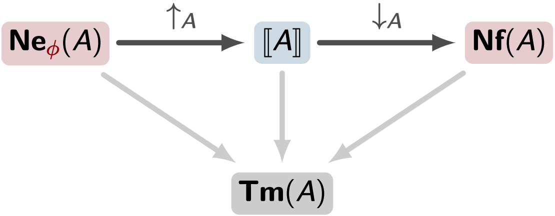


The standard **Tait yoga**

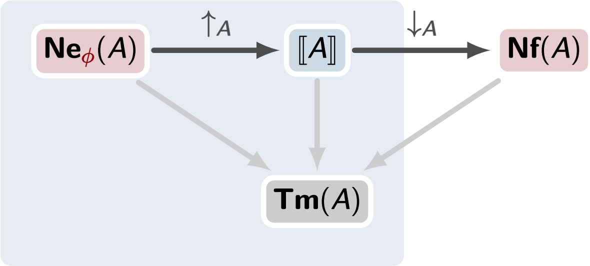
Tait [Tai67] introduced the famous *saturation yoga* for normalization.



Tait's yoga with unstable neutrals

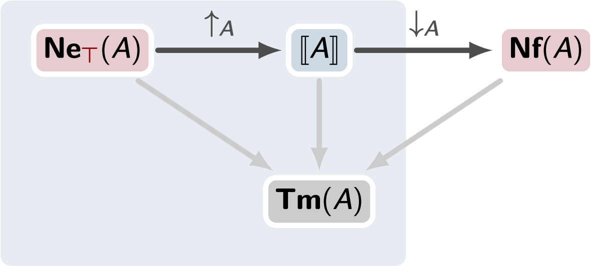


Tait's yoga with unstable neutrals



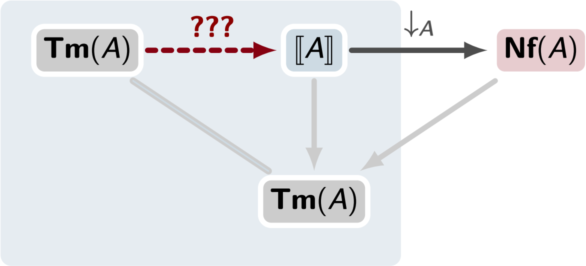
What if $\phi = \top$?

Tait's yoga with unstable neutrals



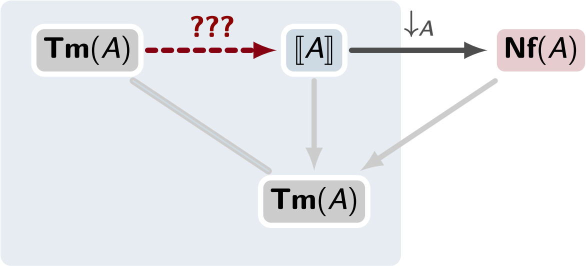
What if $\phi = \top$?

Tait's yoga with unstable neutrals



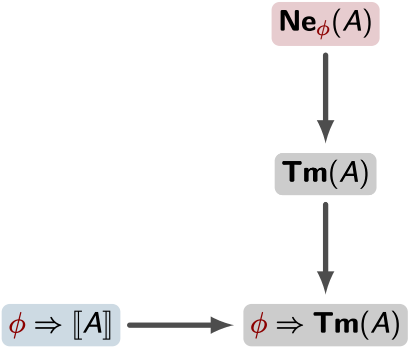
What if $\phi = \top$?

Tait's yoga with unstable neutrals

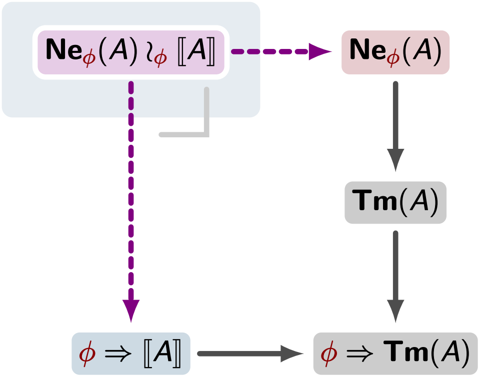


What if $\phi = \top$? We must strengthen the induction hypothesis.

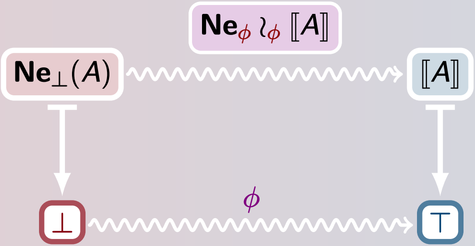
Stabilization of neutrals



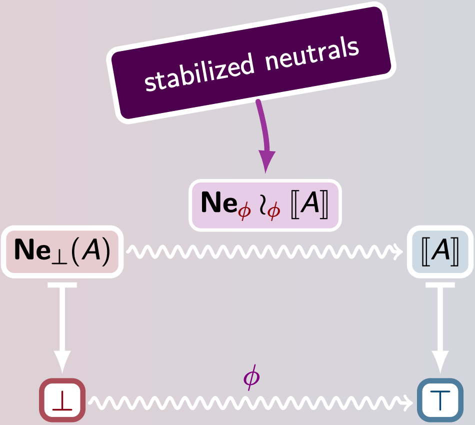
Stabilization of neutrals



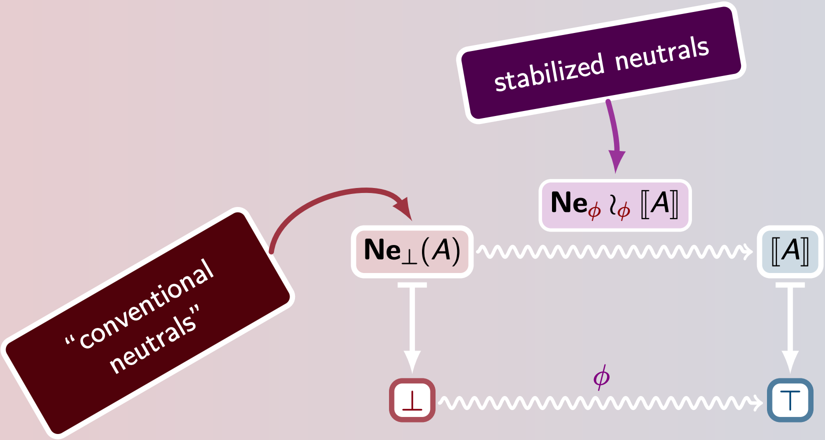
*Unstable neutrals are **glued together** with compatible computability data along their frontiers of instability.*



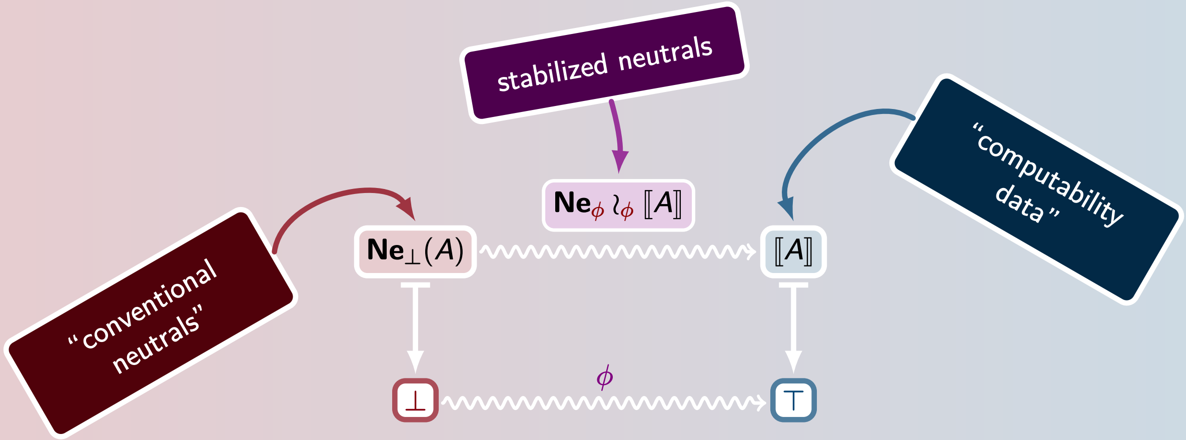
$\mathbf{Ne}_{\phi} \wr_{\phi} [\![A]\!]$
$\mathbf{Ne}_{\perp}(A)$
$[\![A]\!]$
ϕ
$\perp$
$\top$



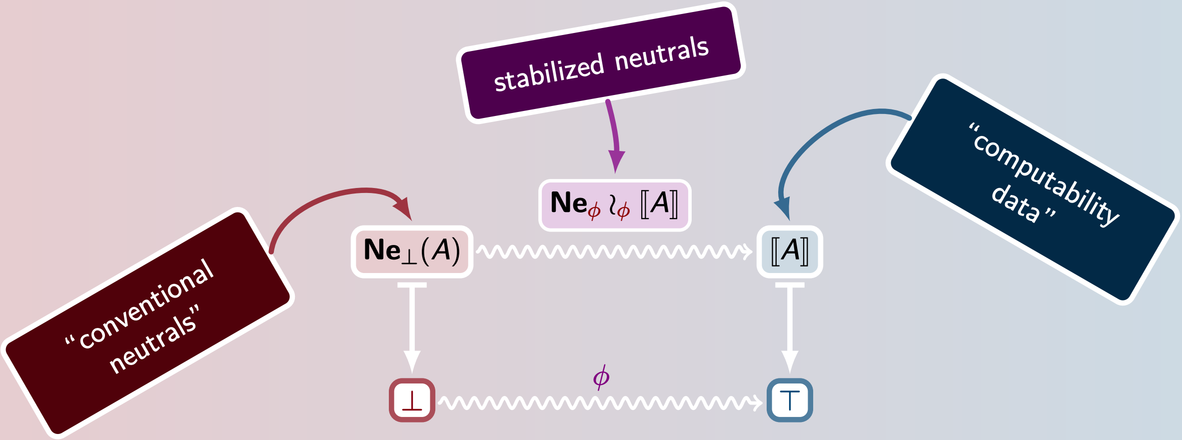
stabilized neutrals
$\mathbf{Ne}_{\phi} \wr_{\phi} [\![A]\!]$
$\mathbf{Ne}_{\perp}(A)$
$[\![A]\!]$
ϕ
$\perp$
$\top$



stabilized neutrals
$\mathbf{Ne}_{\phi} \wr_{\phi} [\![A]\!]$
“conventional neutrals”
$\mathbf{Ne}_{\perp}(A)$
$[\![A]\!]$
ϕ
$\perp$
$\top$

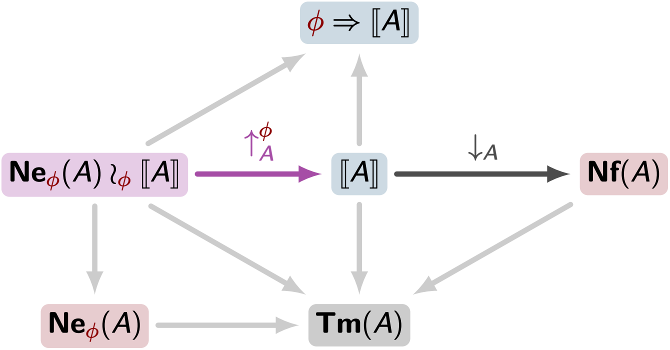


stabilized neutrals
$\mathbf{Ne}_\phi \wr_\phi [\![A]\!]$
"computability data"
"conventional neutrals"
$\mathbf{Ne}_\perp(A)$
$[\![A]\!]$
ϕ
$\perp$
$\top$

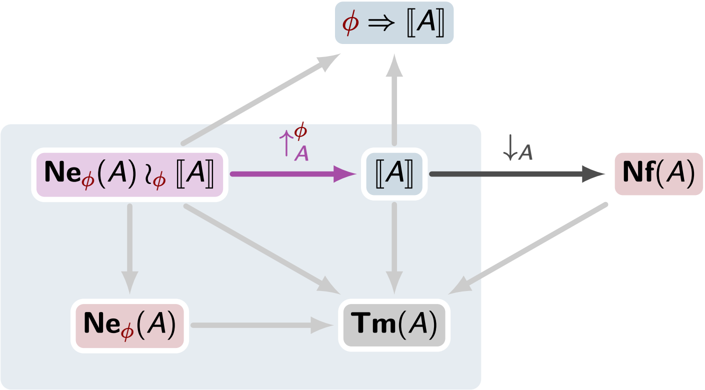


Stabilization interpolates between neutrals and computability data.

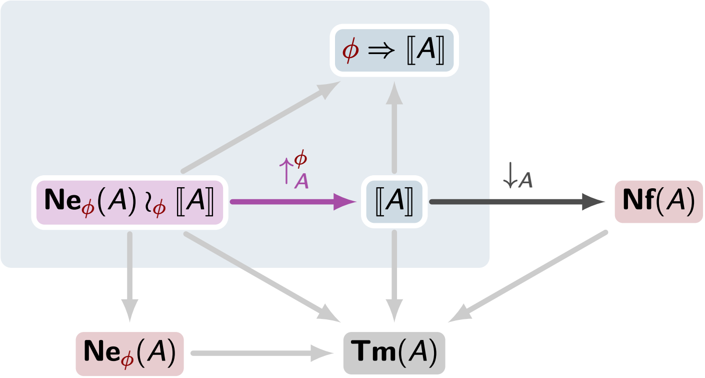
The stabilized Tait yoga



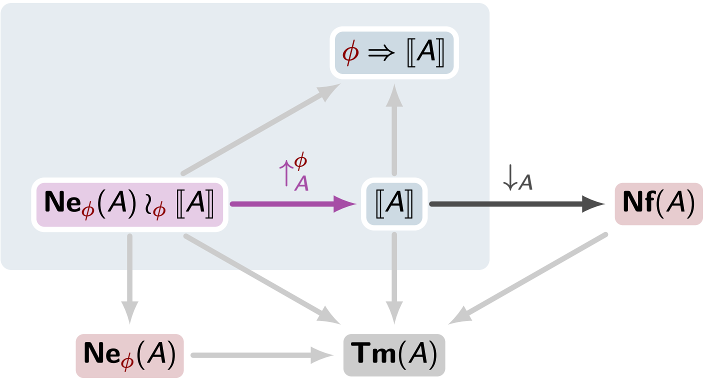
The stabilized Tait yoga



The stabilized Tait yoga



The stabilized Tait yoga



Theorem. Every type is closed under the **stabilized** Tait yoga.

Summary of results

For univalent **□TT** without universes, we have proved the following results:

1. Every type and every term has a *unique* normal form.
2. Judgmental equality of types and terms is decidable.
3. Type constructors (*e.g.* Π) are injective.
4. Type checking is decidable (corollary of 1–3).

Forthcoming: S. has extended this result to **□TT** with a countable hierarchy of univalent universes [Ste21].

References I

[AFH18] Carlo Angiuli, Kuen-Bang Hou (Favonia), and Robert Harper. “Cartesian Cubical Computational Type Theory: Constructive Reasoning with Paths and Equalities”. In: *27th EACSL Annual Conference on Computer Science Logic (CSL 2018)*. Ed. by Dan Ghica and Achim Jung. Vol. 119. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, 6:1–6:17. ISBN: 978-3-95977-088-0. DOI: 10.4230/LIPIcs.CSL.2018.6. URL: http://drops.dagstuhl.de/opus/volltexte/2018/9673.

[AHS95] Thorsten Altenkirch, Martin Hofmann, and Thomas Streicher. “Categorical reconstruction of a reduction free normalization proof”. In: *Category Theory and Computer Science*. Ed. by David Pitt, David E. Rydeheard, and Peter Johnstone. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, pp. 182–199. ISBN: 978-3-540-44661-3.

[AK16] Thorsten Altenkirch and Ambrus Kaposi. “Normalisation by Evaluation for Dependent Types”. In: *1st International Conference on Formal Structures for Computation and Deduction (FSCD 2016)*. Ed. by Delia Kesner and Brigitte Pientka. Vol. 52. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016, 6:1–6:16. ISBN: 978-3-95977-010-1. DOI: 10.4230/LIPIcs.FSCD.2016.6. URL: http://drops.dagstuhl.de/opus/volltexte/2016/5972.

[Ang+21] Carlo Angiuli et al. “Internalizing Representation Independence with Univalence”. In: *Proceedings of the ACM on Programming Languages* 5.POPL (Jan. 2021). DOI: 10.1145/3434293.

References II

[Coh+17] Cyril Cohen et al. "Cubical Type Theory: a constructive interpretation of the univalence axiom". In: *IfCoLog Journal of Logics and their Applications* 4.10 (Nov. 2017), pp. 3127–3169. URL: http://www.collegepublications.co.uk/journals/ifcolog/?00019.

[Coq19] Thierry Coquand. "Canonicity and normalization for dependent type theory". In: *Theoretical Computer Science* 777 (2019). In memory of Maurice Nivat, a founding father of Theoretical Computer Science - Part I, pp. 184–191. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2019.01.015. arXiv: 1810.09367 [cs.PL].

[Fio02] Marcelo Fiore. "Semantic Analysis of Normalisation by Evaluation for Typed Lambda Calculus". In: *Proceedings of the 4th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming*. PPDP '02. Pittsburgh, PA, USA: ACM, 2002, pp. 26–37. ISBN: 1-58113-528-9. DOI: 10.1145/571157.571161.

[Hub18] Simon Huber. "Canonicity for Cubical Type Theory". In: *Journal of Automated Reasoning* (June 13, 2018). ISSN: 1573-0670. DOI: 10.1007/s10817-018-9469-1.

[Lic+18] Daniel R. Licata et al. "Internal Universes in Models of Homotopy Type Theory". In: *3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9-12, 2018, Oxford, UK*. 2018, 22:1–22:17. DOI: 10.4230/LIPIcs.FSCD.2018.22.

[Red18] The RedPRL Development Team. redtt. 2018. URL: https://www.github.com/RedPRL/redtt.

[Red20] The RedPRL Development Team. cooltt. 2020. URL: https://www.github.com/RedPRL/cooltt.

References III

[SA21] Jonathan Sterling and Carlo Angiuli. “Normalization for Cubical Type Theory”. In: *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science*. To appear. New York, NY, USA: ACM, 2021. arXiv: 2101.11479 [cs.LO].

[SH21] Jonathan Sterling and Robert Harper. “Logical Relations As Types: Proof-Relevant Parametricity for Program Modules”. In: *Journal of the ACM* (2021). To appear. arXiv: 2010.08599 [cs.PL].

[Ste21] Jonathan Sterling. “First Steps in Synthetic Tait Computability”. Forthcoming. PhD thesis. Carnegie Mellon University, 2021.

[Str98] Thomas Streicher. “Categorical intuitions underlying semantic normalisation proofs”. In: *Preliminary Proceedings of the APPSEM Workshop on Normalisation by Evaluation*. Ed. by O. Danvy and P. Dybjer. Department of Computer Science, Aarhus University, 1998.

[Tai67] W. W. Tait. “Intensional Interpretations of Functionals of Finite Type I”. In: *The Journal of Symbolic Logic* 32.2 (1967), pp. 198–212. ISSN: 00224812. URL: http://www.jstor.org/stable/2271658.

[VMA19] Andrea Vezzosi, Anders Mörtberg, and Andreas Abel. “Cubical Agda: A Dependently Typed Programming Language with Univalence and Higher Inductive Types”. In: *Proceedings of the 24th ACM SIGPLAN International Conference on Functional Programming*. ICFP ’19. Boston, Massachusetts, USA: ACM, 2019. DOI: 10.1145/3341691.