

Complejidad de cálculos profundos vía topología de espacios de funciones

Eduardo Dueñez

Universidad de Texas en San Antonio

Facultad de Estudios Superiores Acatlán, UNAM

22 de mayo de 2026



<https://github.com/edueney/public-site>

- 1 Introducción
- 2 Un marco matemático para el cómputo profundo
- 3 Existencia de equilibrios profundos
- 4 La dicotomía dócil/salvaje: clase de Baire y NIP
- 5 Una clasificación más fina: niveles de docilidad
- 6 Conclusión

¿Qué sucede a profundidad infinita?

Las redes neuronales son cada vez más **profundas**, año tras año:

AlexNet (2012): 8 capas $\longrightarrow$ ResNet (2015): 152 capas $\longrightarrow$ LLMs modernos: >100 bloques de transformer.

Experimento mental 1. **Profundidad** $\rightarrow \infty$ vista como **tiempo continuo** $t \rightarrow \infty$:

- **Neural ODEs** de Chen *et al.* [Che+18]: Reemplazar las **capas discretas** por una **EDO de tiempo continuo** para la función de evolución del estado $z(t) = z_\theta(t) = z(\theta, z_0; t)$ definida por la EDO **autónoma**:

$$\frac{d}{dt} z(t) = f_\theta(z(t)), \quad z(0) = z_0, \quad t \in [0, \infty), \quad \theta: \text{parámetro(s) real(es)}.$$

- Los “**estados de equilibrio profundo**” con $t \rightarrow \infty$ son los *equilibrios asintóticos* de las trayectorias.
 - Los *equilibrios profundos* para el parámetro θ son cualesquier estados $z^* = z^*(\theta)$ definidos **implícitamente** por:

$$f_\theta(z^*) = 0$$

- Se hallan (en la práctica) usando un *algoritmo iterativo de búsqueda de raíces* **RootFind** “de caja negra” para f_θ con **condición inicial** z_0 .

¿Qué sucede a profundidad infinita? (Continuación...)

Experimento mental 2. Profundidad $\rightarrow \infty$ como un problema de punto fijo bajo iteraciones (en tiempo discreto):

- **Deep Equilibrium Models (DEQ)** [BKK19]: Iterar capas con “**pesos compartidos**” (capas sucesivas con los mismos pesos) *ad infinitum*.
- Si la transformación de capa es f_θ :

$$z_{n+1} = f_\theta(z_n), \quad \text{dados } \theta, z_0: \text{parámetros reales,}$$

hasta alcanzar un punto fijo $z^* = z^*(\theta)$ definido implícitamente por

$$z^* = f_\theta(z^*).$$

- Un algoritmo de punto fijo de caja negra **FixFind** para f_θ con inicial z_0 dado encuentra iterativamente algún equilibrio $z^*(\theta)$.

Preguntas centrales de esta charla

¿Qué tipo de funciones $\theta \mapsto z^*(\theta; z_0)$ computa una red infinitamente profunda? ¿Cuándo es $z^*(\theta; z_0)$ efectivamente computable? ¿Cuándo es $z^*(\theta; z_0)$ una función dócil vs. salvaje?

Deep Equilibrium Models: un ejemplo trabajado

Fijemos la **función de activación**

$$f_{\theta}(z) = \tanh(\mathbf{w}z + \mathbf{b}).$$

Seguimos la convención *implícita/fea* de que θ denota “todos los parámetros ajustables”; en efecto, $\theta = (\mathbf{w}, \mathbf{b})$ aquí.

Paso hacia adelante: Iterar hasta converger

$$z_{n+1} = \tanh(\mathbf{w}z_n + b) \quad \text{partiendo de un } z_0 \text{ dado.}$$

- Existen 1, 2 ó 3 **puntos fijos** z^* si $w \neq 0$.
- Los puntos fijos pueden hallarse mediante un **algoritmo iterativo “de caja negra”** distinto de la definición recursiva de z_n .
- **Cuál** punto fijo z^* se encuentra a partir de z_0 **depende del algoritmo**.

Paso hacia atrás (diff. implícita):

$$\frac{\partial z^*}{\partial \mathbf{w}} = \frac{(1 - z^{*2}) z^*}{1 - \mathbf{w}(1 - z^{*2})}, \quad \text{etc.}$$

No se requiere retropropagación “infinita”.

- Por qué importa:
 - Memoria $O(1)$, sin importar la “profundidad efectiva”;
 - **El cómputo del gradiente es exacto en el límite**;
 - **Bai–Kolter–Koltun (2019)**: empíricamente estable.

BKK dejan abierto: la existencia de z^* y qué tipo de función es $\theta \mapsto z^*(\theta; z_0)$.

Los límites a profundidad infinita son funciones de clase de Baire 1

Las **redes estándar (de profundidad finita)** computan funciones **continuas**.

Los **límites a profundidad infinita** surgen como **límites puntuales** o **ultralímites puntuales** de sucesiones de funciones continuas: distinguir **límites** de **ultralímites** es **crítico**.

Definition (Clase de Baire 1)

Una **función real** g sobre un espacio topológico X se llama de **clase de Baire 1** (**Baire-1**, $B_1(X)$ o simplemente B_1) si es el **límite puntual** $g = \lim_{n \rightarrow \infty} f_n$ de una sucesión de *funciones continuas* f_n sobre X .

Baire-1

- Equilibrios profundos “dóciles”;
- Función de Thomae (regla);
- Algoritmo de búsqueda de raíces de Newton $(f; z_0) \mapsto \text{Newton}(f; z_0)$ para $z_0 \in \text{Fatou}$.

No Baire-1

- Equilibrios profundos “salvajes”;
- Función de Dirichlet $\mathbf{1}_{\mathbb{Q}}$;
- Indicadores del conjunto de Julia, p. ej., $(f; z_0) \mapsto \text{Newton}(f; z_0)$ para $z_0 \in \text{Julia}$.

Preguntas centrales (reformuladas): ¿Qué cálculos profundos son **genuinos límites**

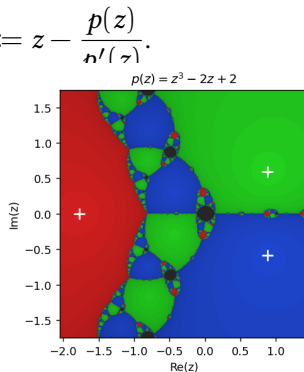
Un ejemplo recurrente: el método de Newton como cómputo profundo

Método de Newton: Resolver $p(z) = 0$ para un **polinomio** $p = p_\theta$ dado mediante aplicaciones sucesivas del *mapeo de Newton* $N_\theta = N_p$:

$$z_{n+1} = N_p(z_n), \quad z_0 \in \mathbb{C} \text{ dado}; \quad N_p(z) := z - \frac{p(z)}{p'(z)}.$$

El plano complejo extendido $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ se parte en:

- Dinámica **dócil**: **conjunto de Fatou**.
 - Abierto denso de puntos iniciales z_0 que admiten una vecindad donde una subsucesión de iterados de Newton converge puntualmente.
- Dinámica **salvaje**: **conjunto de Julia** (complemento de Fatou).
 - Puntos z_0 que *no* admiten vecindad alguna donde una subsucesión de iterados converja puntualmente.

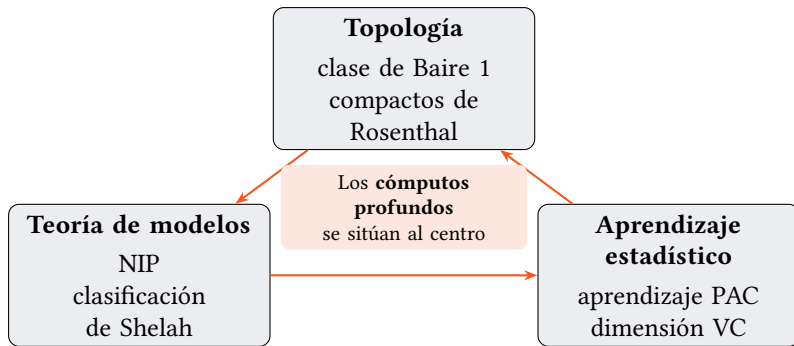


Fractal de Newton para $z^3 - 2z + 2$.

Todo **ultralímite** $z^*(z_0; \theta) = z_\mathcal{U}^*(z_0; \theta) = \mathcal{U}\lim_{n \rightarrow \infty} z_n$ es un **cómputo profundo** o

La piedra Rosetta: tres programas de clasificación

La misma dicotomía dócil/salvaje aparece **independientemente** en tres áreas:



Los **cómpu**tos **profundos** **dóci**les son simultáneamente:
Baire-1 (topología) $\iff$ NIP (teoría de modelos) $\iff$ PAC-aprendibles (estadística)

§2 Un marco matemático para el cómputo profundo

Estructuras de Estados de Cómputo (CSS) y Estructuras de Cómputo Composicional (CCS); el mapeo de tipos $\text{tp}: L \rightarrow \mathbb{R}^{\mathbb{N}}$.

§3 Existencia de equilibrios profundos

Cómputos profundos como ultralímites; el lema de Ellis–Numakura garantiza existencia para cómputos confinados.

§4 La dicotomía dócil/salvaje: clase de Baire y NIP

¿Cuándo está un cómputo profundo en B_1 ? El teorema de Bourgain–Fremlin–Talagrand como criterio computable.

§5 Una clasificación más fina: niveles de docilidad

La tricotomía de Todorčević para compactos de Rosenthal; niveles NIP y PAC-aprendibilidad; los fractales de Newton como ejemplo recurrente.

§6 Conclusión y problemas abiertos

Un modelo para el cómputo a profundidad infinita

Fijemos un **espacio de estados** L y una familia numerable $\mathcal{P} = (P_n)$ de **observables reales** $P_n: L \rightarrow \mathbb{R}$ (coordenadas, normas, salidas de capa, ...).

El **tipo** de $v \in L$ codifica todos sus **rasgos observables**:

$$\text{tp}(v) := (P_n(v))_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}.$$

El **espacio de tipos** $\bar{L} = \overline{\text{tp}(L)} \subseteq \mathbb{R}^{\mathbb{N}}$ (cerradura en la topología producto) añade los **“ultraestados” ideales** alcanzables sólo a *profundidad infinita*.

Un **cómputo** $\gamma: L \rightarrow L$ es cualquier transformación **continua** del estado; los cálculos se componen, formando un semigrupo Γ .

CSS y CCS (en breve)

- **CSS** = $\langle L, \mathcal{P} \rangle$: espacio de estados + observables.
- **CCS** = $\langle L, \Gamma \rangle$: CSS + semigrupo de cálculos.
- Si L es **compacto**, también lo es $\bar{L}$ (por Tychonoff): **no se escapa al infinito**.

Ejemplos recurrentes

- **DEQ.** $L = \mathbb{R}$; $\gamma_\theta(z) = \tanh(wz + b)$.
- **Newton.** $L = \hat{\mathbb{C}}$ (plano complejo extendido, **compacto**); $\gamma = N_p$.

Cómputos profundos: ultrálmites de iteraciones

Iterar un cómputo indefinidamente: $\gamma^0 = \text{id}$,

$$\gamma^1 = \gamma, \gamma^2 = \gamma \circ \gamma, \dots$$

Un **cómputo profundo** (ultracómputo) es cualquier **ultralímite**

$$\gamma^{(\mathcal{U})} := \mathcal{U}\lim_{n \rightarrow \infty} \gamma^n,$$

definido **puntualmente**:

$$\gamma^{(\mathcal{U})}(\nu) = \mathcal{U}\lim_n \gamma^n(\nu) \in \bar{L}.$$

- $\gamma^{(\mathcal{U})}(\nu)$ puede ser un **ultraestado** $\in \bar{L} \setminus L$.
- $\gamma^{(\mathcal{U})}$ es **típicamente discontinuo**.
- Cuando $L = \bar{L}$ es **compacto**: todo ultrafiltro $\mathcal{U}$ sobre $\mathbb{N}$ da un cómputo profundo “realizado” $\gamma^{(\mathcal{U})} : L \rightarrow L$.

Los cómputos finitos son **continuos** (B_0).

Los cómputos profundos *discontinuous* son $B_{\geq 1}$.

Clase de Baire 1 (B_1)

Una función $g: L \rightarrow \mathbb{R}$ es **Baire-1** si

$$g = \lim_{n \rightarrow \infty} f_n$$

puntualmente, con cada f_n **continua**.

Ser Baire-1 significa que g es “**apenas**” **discontinua**.

Preguntas centrales de esta charla:

- ¿Cuándo es un **cómputo profundo** **Baire-1**?
- ¿Qué sucesiones de cómputos producen cómputos profundos B_1 —pero **no** $B_{>1}$?

Existen equilibrios profundos — Ellis–Numakura

Definition (Equilibrio profundo)

Un cómputo profundo f_u de γ es un **equilibrio profundo (iterativo)** si es **idempotente**:

$$f_u \circ f_u = f_u.$$

Theorem ([Alv+24], vía Ellis–Numakura [Ell58])

Si la cerradura de la órbita $\overline{\{\gamma^n : n \geq 1\}}$ es **compacta**, existe al menos un equilibrio profundo.

▷ Todo semigrupo topológico compacto contiene un elemento **idempotente** —éste es el lema de Ellis–Numakura (argumento por lema de Zorn sobre subsemigrupos cerrados minimales). $\square$

Newton: $L = \widehat{\mathbb{C}}$ (compacto)

Cerradura $\overline{(N_p^n)_n}$ compacta
 $\Rightarrow$ **siempre existen EQP.**

B_1 en Fatou: $\forall \mathcal{U}$ existe una subsucesión (n_i)
t.q. $N_p^{(\mathcal{U})} = \lim_i N_p^{n_i}$ (en el conjunto de Fatou).

DEQ: la órbita queda en $[-1, 1]$

$|f_\theta(z)| = |\tanh(\mathbf{w}z + \mathbf{b})| \leq 1 \Rightarrow$ cerradura
 $\overline{(f_\theta^n)_n}$ **compacta** $\Rightarrow$ **existen EQP.**

Dinámica **dócil:** $\forall \mathcal{U}: (\theta; z) \mapsto f_\theta^{(\mathcal{U})}(z)$ es B_1 .

Bourgain–Fremlin–Talagrand: una dicotomía fundamental

Pregunta. Dada una iteración (γ^n) , ¿cuáles iterados profundos $\gamma^{(\mathcal{U})}$ son **Baire-1**?

Theorem (Dicotomía BFT [BFT78])

Para una familia $(\gamma^n)_{n \in \mathbb{N}}$ de funciones continuas sobre un espacio polaco, *localmente uniformemente acotada*, se cumple *exactamente una* de:

- (Dócil)** La cerradura $\overline{\{\gamma^n\}}$ en la topología puntual está contenida en B_1 : *todo iterado profundo* $\gamma^{(\mathcal{U})} = \lim_i \gamma^{n_i}$ es un *límite sub-sucesional*.
- (Salvaje)** La cerradura $\overline{\{\gamma^n\}}$ tiene un subconjunto *homeomorfo a* $\beta\mathbb{N}$ (la compactificación de Stone–Čech de $\mathbb{N}$).

- **Dócil.** Hay a lo más $2^{\aleph_0}$ iterados profundos distintos: cualquier $\mathcal{U}$ da un ultralímite $\gamma^{(\mathcal{U})} = \lim_i \gamma^{n_i}$ para alguna subsucesión (n_i) .
- **Salvaje.** Los iterados profundos forman un conjunto enorme de cardinalidad $2^{2^{\aleph_0}}$: ultrafiltros distintos dan cálculos profundos distintos **que no surgen como límites sucesionales** de ninguna subsucesión.

Patrones de independencia: el lado salvaje

¿Qué distingue lo **dócil** de lo **salvaje**? Un **patrón de independencia** en (γ^n) .

Definition (NIP, informal)

(γ^n) tiene la **Propiedad de Independencia (IP)** al umbral $a < b$ si para **toda** pareja finita disjunta $E, F \subseteq \mathbb{N}$ existe un estado v con

$$\gamma^n(v) \leq a \text{ para } n \in E, \quad \gamma^n(v) \geq b \text{ para } n \in F.$$

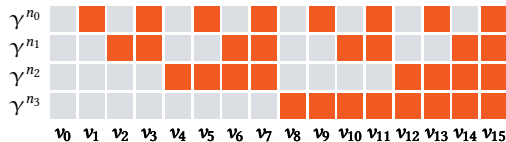
(γ^n) es **NIP** (“**No IP**”) si no la tiene.

NIP es la propiedad “**buena**” que *acota* la complejidad de (γ^n) .

Para una familia $\mathcal{F}$ de funciones $\{0, 1\}$ -valuadas:
 $\mathcal{F}$ tiene **IP** $\Leftrightarrow \dim \text{VC}(\mathcal{F}) = \infty$
 $\Leftrightarrow \mathcal{F}$ **fragmenta** conjuntos finitos

Imagen de IP

Un patrón de independencia de 4 filas (hipotético) para **todas** las particiones $E \sqcup F = \{n_1, n_2, n_3, n_4\}$.



Familia **IP** $\mathcal{F} = \{\gamma^{n_0}, \gamma^{n_1}, \gamma^{n_2}, \gamma^{n_3}\}$:
todos los $2^4 = 16$ patrones de signos.

Familias **dóviles** $\mathcal{F}$: Menos de $2^{\#\mathcal{F}}$ patrones.

Tres caras equivalentes de la docilidad (Teorema Principal)

Theorem (Resultado central del artículo)

Sea $(L, \mathcal{P}, \Gamma)$ una CCS con $\mathcal{P}$ numerable y $\Delta \subseteq \Gamma$ **una familia confinada** de cálculos continuos. Entonces son equivalentes:

- ❶ (Topología) $\bar{\Delta}$ es un **compacto de Rosenthal** (i.e., $\bar{\Delta} \subseteq B_1$);
- ❷ (Teoría de modelos) Para cada observable $P \in \mathcal{P}$, la familia $(P \circ \gamma)_{\gamma \in \Delta}$ es **NIP**;
- ❸ (Aprendizaje estadístico) Las “funciones-nivel” $\{0, 1\}$ -valuadas de Δ tienen **dimensión VC finita**, luego son **PAC-aprendibles**.

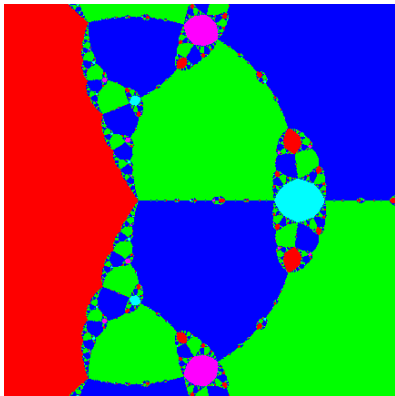
Más aún, en tal caso **todo** iterado profundo $\gamma^{(\omega)} \in \bar{\Delta}$ es un **límite puntual sucesional** de cálculos en Δ .

$$\text{Dócil} = \text{Baire-1} = \text{NIP} = \text{PAC-aprendible}$$

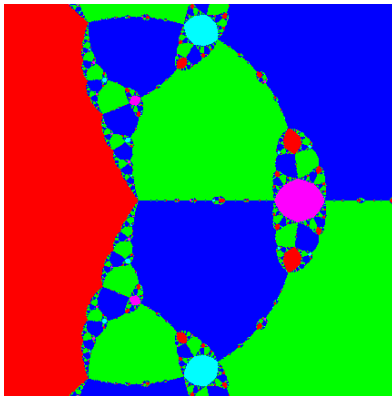
Eslogan. “El cálculo profundo es factible **si y sólo si** la familia subyacente es estadísticamente aprendible.”

Fractales de Newton: dóciles, pero con estructura

Recordemos. $p(z) = z^3 - 2z + 2$ tiene 3 raíces r_0, r_+, r_- + 2-ciclo $0 \leftrightarrow 1$ de N_p (no raíces de $p!$).



$n = 100$ (par)



$n = 101$ (impar)

Regiones R, V, A: cuencas de las tres raíces reales $r_0, r_{\pm}$.
Regiones cian / magenta: cuenca del 2-ciclo $\{0, 1\}$.
Cian $\Leftrightarrow$ magenta al pasar $n \rightarrow n + 1$: las subsucesiones par e impar $(N_p^{2i}), (N_p^{2i+1})$ convergen a equilibrios profundos **distintos** f_{par} y f_{non} .

El mundo dócil es más rico que “un solo límite”

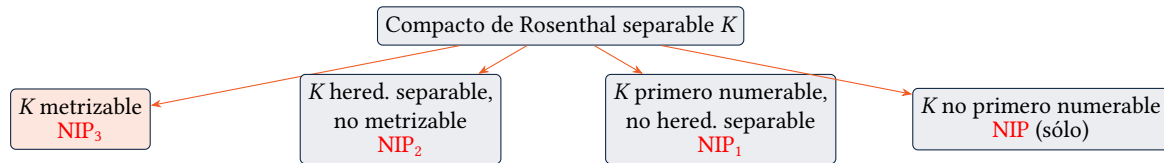
Los equilibrios profundos $f_{\text{par}}, f_{\text{non}}$ son **Baire-1**; juntos capturan la **estructura dinámica asintótica**.

Dentro de Baire-1: la tricotomía de Todorčević

Aun *dentro* de la clase dócil, Baire-1 tiene una **estructura interna fina**.

Theorem (Tricotomía de Todorčević [Tod99])

*Todo compacto de Rosenthal **separable** K que sea no metrizable contiene una copia canónica de **uno de tres prototipos minimales**.*



- Implicaciones estrictas: $\text{NIP}_3 \Rightarrow \text{NIP}_2 \Rightarrow \text{NIP}_1 \Rightarrow \text{NIP}$.
- Para cálculos profundos: cuatro niveles de **docilidad topológica** que refinan "Baire-1".
- Refinamiento adicional: la **heptacotomía** de Argyros–Dodos–Kanellopoulos: **siete** prototipos.

Ejemplos en cada nivel de docilidad

Cada nivel NIP_i es no vacío: hay CCSs concretas que lo realizan.

Nivel	CCS	Cerradura de Δ en B_1
NIP_3	Iteración de raíz cuadrada (cómputo de $\sqrt{x}$)	Compactificación de Alexandroff (metrizable)
NIP_2	Clasif. umbral de precisión finita en $2^{\mathbb{N}}$	Cantor escindido $S(2^{\mathbb{N}})$ (h. s., no metrizable)
NIP_1	Familia de corrimientos $\{\gamma_t : t \in 2^{<\mathbb{N}}\}$	$\widehat{D}(S(2^{\mathbb{N}}))$ (1.ª num., no h. s.)
NIP (sólo)	Tests de prefijo de precisión finita	Alexandroff extendido $\widehat{A}(2^{\mathbb{N}})$

Interpretación. Cada nivel corresponde a la **eficiencia** con que los cálculos profundos pueden **aproximarse** mediante cálculos estándar:

- NIP_3 (metrizable): numerablemente, convergencia **casi uniforme**;
- $\text{NIP}_2 \rightarrow \text{NIP}$: aproximación progresivamente **más difícil**.

Más allá de NIP: el lado salvaje

Familias IP: la cerradura contiene $\beta\mathbb{N}$ (consta de $2^{2^{\aleph_0}}$ **cálculos profundos distintos**).

No hay reducción a un Δ numerable: **no son aproximables** por sucesiones.

Marco (1.^{er} artículo: Alva–Dueñez–Iovino–Walton 2024)

- **CSSs** y **CCSs**: un marco topológico para el cómputo con valores reales.
- Los **cómpu**tos **profundos** surgen como ultrálmites en el espacio de tipos.
- La **existencia de equilibrios profundos** se sigue de compacidad + Ellis–Numakura.

Clasificación (2.^o artículo: Dueñez–Iovino–Matos–Salvetti–Tall)

- **Dicotomía BFT**: dócil (Baire-1) vs. salvaje (contiene $\beta\mathbb{N}$).
- **Triple equivalencia**: $\text{Baire-1} \Leftrightarrow \text{NIP} \Leftrightarrow \text{dim VC finita} (\Leftrightarrow \text{PAC-aprendible})$.
- **Subclases de Baire-1**: tricotomía de Todorćević (jerarquía $\text{NIP}_3 \Rightarrow \text{NIP}_2 \Rightarrow \text{NIP}_1 \Rightarrow \text{NIP}$). (Argyros–Dodos–Kanellopoulos: 7 prototipos minimales).

El cómputo profundo es factible $\Leftrightarrow$ la familia subyacente es estadísticamente aprendible.

Problemas abiertos y direcciones futuras

- **¿Qué arquitecturas neuronales son NIP?** Identificar NIP/IP para familias DEQ y Neural ODE concretas; caracterizar las arquitecturas profundas “computacionalmente factibles”.
- **Computabilidad profunda Monte Carlo.** Versiones probabilísticas de NIP (**estabilidad de Talagrand**); cómputo Monte Carlo de equilibrios profundos *cuando* falla la convergencia puntual.
- **Caracterizaciones no topológicas de NIP_i .** ¿Pueden los niveles $\text{NIP}_1, \text{NIP}_2, \text{NIP}_3$ detectarse mediante invariantes puramente **combinatorios** o **lógicos** de Δ ?
- **Aprendizaje en contexto en transformers.** Las LLMs modernas exhiben *aprendizaje al momento de inferir*. ¿Son NIP las clases de hipótesis en contexto? ¿Tienen $\dim \text{VC}$ finita?
- **Más allá de NIP.** Refinamientos de IP (NTP_2 , simplicidad, distalidad) dan una taxonomía más fina del lado salvaje. ¿Hay jerarquías naturales de cómputos profundos **salvajes**?

¡Muchas gracias!

Preguntas y comentarios bienvenidos.

Referencias I

- [Alv+24] Samson Alva et al. *Deep equilibria: Existence and computability*. Preprint. arXiv:2409.06064. 2024. DOI: 10.48550/arXiv.2409.06064. URL: <https://arxiv.org/abs/2409.06064>.
- [BFT78] J. Bourgain, D. H. Fremlin, and M. Talagrand. “Pointwise compact sets of Baire-measurable functions”. In: *Amer. J. Math.* 100.4 (1978), pp. 845–886. ISSN: 0002-9327,1080-6377. DOI: 10.2307/2373913. URL: <https://doi.org/10.2307/2373913>.
- [BKK19] Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. “Deep Equilibrium Models”. Preprint. <https://arxiv.org/abs/1909.01377>, <https://implicit-layers-tutorial.org/>. 2019.
- [Che+18] Ricky T. Q. Chen et al. “Neural ordinary differential equations”. In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. NIPS’18. Montréal, Canada: Curran Associates Inc., 2018, pp. 6572–6583. DOI: 10.5555/3327757.3327764. arXiv: 1806.07366 [cs.LG].

- [Ell58] Robert Ellis. “Distal transformation groups”. In: *Pacific J. Math.* 8 (1958), pp. 401–405. ISSN: 0030-8730,1945-5844. URL: <http://projecteuclid.org/euclid.pjm/1103039885>.
- [Tod99] Stevo Todorčević. “Compact subsets of the first Baire class”. In: *J. Amer. Math. Soc.* 12.4 (1999), pp. 1179–1212. ISSN: 0894-0347,1088-6834. DOI: 10.1090/S0894-0347-99-00312-4. URL: <https://doi.org/10.1090/S0894-0347-99-00312-4>.