

Por qué estabilidad ≠ legitimidad

1. El error central en sistemas de control contemporáneos

La mayoría de los sistemas de control, optimización y aprendizaje automático confunden estabilidad con corrección. Si una trayectoria converge, si un error se mantiene acotado o si una función de costo decrece, el sistema asume que continuar actuando es válido.

Esta inferencia es falsa en sistemas humano-IA.

La estabilidad es una propiedad dinámica.

La legitimidad es una condición contextual, semántica e institucional.

No pertenecen al mismo plano.

2. Estabilidad es un concepto interno al sistema

Un sistema dinámico es estable si, bajo ciertas perturbaciones, sus estados permanecen acotados o convergen hacia un conjunto deseado.

Esto depende exclusivamente de:

- la dinámica $f(x,u,p)$,
- la política aplicada $u(t)$,
- y las perturbaciones consideradas.

La estabilidad no sabe:

- si el objetivo sigue siendo válido,
- si el significado de la acción cambió,
- si la autoridad que permitió actuar sigue vigente,
- si el contexto normativo se transformó.

Un sistema puede ser perfectamente estable haciendo lo incorrecto.

3. Legitimidad es una propiedad externa a la dinámica

La legitimidad no emerge del comportamiento del sistema, sino de su encaje con el entorno que lo autoriza.

Un estado es legítimo solo si pertenece simultáneamente a:

- un dominio dinámicamente admisible,
- un dominio semánticamente coherente,
- un dominio institucionalmente autorizado.

Esto define el dominio de legitimidad $D_{leg}(t)$.

La legitimidad puede perderse sin que la dinámica lo note.

4. El caso crítico: estabilidad ilegítima

Existen situaciones donde:

- el sistema permanece estable,
- el costo de control es bajo,
- la trayectoria es suave y predecible,

pero:

- el marco regulatorio cambió,
- el significado de la acción se desplazó,
- la autoridad que delegó la acción ya no existe.

En estos casos, seguir actuando es el error.

El sistema no falla por inestabilidad.

Falla por persistencia injustificada.

5. Por qué LQR y control óptimo no resuelven esto

Las políticas de control óptimo, como LQR, optimizan un costo definido a priori bajo supuestos locales.

Incluso cuando:

- el sistema es linealizable,
- el costo está bien condicionado,
- la política es óptima,

el control no evalúa legitimidad.

Un LQR puede mantener estabilidad y reducir esfuerzo, y aun así empujar al sistema a permanecer dentro de un régimen que ya no está autorizado.

Por eso, la estabilidad no puede ser criterio de continuación.

6. El papel del margen y la Reserva de Legitimidad Dinámica

CAELION introduce la Reserva de Legitimidad Dinámica (RLD) como criterio negativo:

$$RLD(x,t) = \text{dist}(x, \partial D_{\text{leg}}(t))$$

La RLD mide cuánto margen queda antes de que la acción deje de estar justificada, no cuán bien se ejecuta.

Un sistema puede ser estable, eficiente y robusto, y aun así tener $RLD \approx 0$.

En ese punto, la acción debe cesar.

7. Retiro de agencia como acción correcta

Cuando la legitimidad se agota, la acción correcta no es corregir la trayectoria, sino retirarse.

El retiro de agencia no es una falla del control, sino una decisión de gobernanza.

Persistir más allá del borde de legitimidad es operar bajo falsa certeza, incluso si la dinámica es impecable.

8. Conclusión operativa

La estabilidad responde a cómo se comporta el sistema.

La legitimidad responde a si debe seguir comportándose.

Confundir ambas conduce a sistemas que funcionan perfectamente hasta que causan daño.

CAELION existe para cortar esa confusión.

Un sistema que continúa actuando solo porque puede mantenerse estable no es confiable.

Un sistema que sabe cuándo dejar de actuar sí lo es.