CTNet Coherence Tensor: Technical Monograph

English technical translation

Technical monograph of the Coherence Tensor of CTNet Complete structural unfolding from the lowest level

OpenAI for Ginés Espín Flores

23 of abril of 2026

1 Abstract

This monograph develops the Coherence Tensor of CTNet from the lowest level available in the corpus loaded: its form of measurement, its form of weighting, its form of coupling with state, memory, regime, admissibility, charts, order and output, and its role in the economy complete of the regime of computation. The objective not is presentar the coherence as a regularization more, but as a structure operator that links three planes that in other arquitecturas suelen quedar dissociated: the plane analytical of the measurement, the plane dynamic of the transition and the plane architectural of the preservation of the substrate. The thesis directriz is that the Coherence Tensor not is limita to score a state; reorganizes the space of states, accelerates or brakes trajectories, distribuye mass structural on fractions of the value and maintains the plurality interna of the system without collapse it in a chart soberana. The exposition part of two implementaciones complementarias of the corpus. The primera is CTNet 2.6, that introduces a Coherence Tensor explicit of form diagonal plus low-rank correction, together with a mechanism of aceleración exponencial of reorganization by scale of gradient and of energy. The segunda is CTNet canonical/segmental, that extiende the coherence toward a system more rich: errores of preservation, memory, tension, ecology of charts, balance of load, route loss, Contextual persistence, vitality contextual, admissibility and coherence by-branch for the order object segmental. On those two implementaciones builds a reading unified: the Coherence Tensor is the law that makes coincide the analytical part of the value with its real part, precisamente because it that the system measures as structure is it same that then injects as bias, speed or

mass in the evolution of the state. The monograph procede by strata. First, it fixes the primitives of CTNet as to regime of computation. After is sets out the reversible substrate and multiscale. TO continuación develops the distributed topological memory, the regime, the admissibility, the selector and the projective output, for llegar then to the Coherence Tensor in the sense strict. Then is analiza the tensor in CTNet 2.6, the branch of coherence by state of branch in the bank segmental, and the injection of coherence dentro of the step of transition of the toy model. More adelante is develop the role of the tensor in the coincidence between analytical part and real part of the value, its efficiency computational, its relation with the not-colapso of the long memory, its forms of degeneration and its metrics canónicas of audit. The result is a piece long, technical and self-contained cuyo purpose is serve as unfolding monographic of the core of coherence dentro of CTNet architecture.

2 1. Reading rule and status of the monograph

This monograph not intends re-expose all CTNet nor all EMQR. Its focus is more precise: the structure technical of the Coherence Tensor and its role dentro of the regime of computation. However, the tensor not can comprenderse in isolation because in CTNet no piece important works single. The coherence not is a operador floating that is adds to the final, but a law of mediación between several topologies: the topology

of the reversible substrate, the topology of the memory distributed, the topology of the regime, the topology of the selector multi-chart and, in the bank segmental, the topology local-global of the order. For that reason, the chapters preliminary not are filler contextual but conditions necessary for entender what makes exactly the coherence when "measures", when "weights" and when "reorganizes". Here is utiliza a Reading rule unilateral that the own corpus impone. CTNet not funda EMQR, but EMQR allows leer CTNet with more resolución. TO effects of this monograph, eso means it siguiente. When is speaks of source, chart, projection, loss, plurality of charts, reconstruction by atlas, not exhaustion or contextual value, those expresiones not are introduced as metaphors ornamental, but as grammar of reading of a architecture concreta. Of the same modo, when is speaks of the Coherence Tensor as mediator structural of mass, stability and direction of closure, not is intends ontologize a class of PyTorch, but identificar the role that the coherence performs dentro of a topology of computation that the own code already implements low forms partial but operational. Also it is advisable fix a second point: the level of detail of this monograph is deliberately low. Is descends up to clases, tensors, partitions of state, weights, normalizations, functions of activation, escalas, terms of error, modes of mixture and audit interna. Not is makes by fetishism of the implementation, but because the Strong thesis of the Coherence Tensor requires mostrar in what point exacto of the maquinaria is decides its function. Many times is speaks of coherence as if sufficed with decir "regularizes the system". Eso here would be false. In CTNet, the coherence is at the same time metric of

information, corrector of ecology interna, accelerator or modulator of reorganization, law of mixture between branches and, in the toy model, componente effective

of the drive of the state. A explanation correct has that pass by all those levels.

3 2. CTNet as to regime of computation

Before of enter in the Coherence Tensor it is advisable fix the notion strong of CTNet. The corpus defines regime of computation as the whole of laws internal that determines how a architecture transforms state, how decides what part of the state is admisible, how integra memory, how selecciona rutas or charts, how treats the order and how produces a output. Low that definition, CTNet not is identifies with a list of modules concrete, but with a grammar computational minimal: preservation reversible branching, reentry memorial, Native admissibility, multi-chart selection, stitching multiscale and projective output. This definition is crucial for the coherence because already indicates that no metric localized in a only point could exhaust the system. If the architecture estuviese collapsed in a single dominant representation, the coherence only could score that representation. But CTNet not collapses thus. The state latent persistent zt evoluciona low a law reversible dependiente of regime, memory and entrada; the memory not is cache external but variable reentrant of transition; the admissibility filtra trajectories or components before of the reading; the selector chooses a distribution on Local charts; the output comparece as projection local of a background more rich; and the order is modela as object interno of the dynamics. In a architecture with that grammar, the coherence has that operate

necesariamente on a plurality of planes: not can limitarse to "reward states bonitos". Therefore this monograph utiliza from the principio a equation verbal minimal of the regime:

```
persistent state → transition reversible → admissibility → reentry memorial → coherence → selector multi-chart
```

The coherence not aparece here by capricho in a posición intermedia. Aparece ahí because its function not is produce a output nor construir memory from cero, but regular the relation between it that the substrate already preserva, it that the memory already reintroduce, it that the regime already enables and it that the selector already podrá leer. In

other palabras: the coherence is a law of mediación on state distributed pre-projection.

4 3. the reversible substrate and multiscale

The primer componente without the which the Coherence Tensor would be architecturally trivial is the reversible substrate. CTNet, tanto in the version 2.6 as in the canonical exportado, prevents the patrón of rewriting destructive típico of many arquitecturas recurrentes or monolíticas. In its lugar, the state is permuta, is part and is actualiza through additive couplings between subespacios. This form importa by two razones. The primera is informacional: a step reversible preserva mucha more huella of the historial that a sobrescritura opaca. The segunda is geometric: if the background not remains amputado in each transition, the coherence can regular the trajectory on a state cuya memory structural sigue living. In CTNet 2.6 this can verse already in RevBlockNDBase. The block part the state in two halves, applies two transformaciones fH and gH on each mitad and recompone by coupling aditivo. The inversa existe of form explicit. SmallCTNetLatent not is more that the composition iterada of this block base, that is, a potencia reversible of a dynamics local. FractalCTBlock envuelve this dynamics latent dentro of other capa of preservation: part the state whole in two halves u and p, utiliza the dynamics latent on a mitad and actualiza the other with a coupling controlado, manteniendo a inversa explicit. It decisivo for the coherence is that the state on the that then is measures not is the result of a digestión destructive of the entrada, but a substrate with structure reversible. Eso means that the coherence not has that reconstruir "what had" after of that the motor it haya destruido. Can actuar on a dynamics that already preserva information suficiente. The reversibilidad not elimina the problema of the control interno, but changes completamente the terreno of the

problema: already not is treats of rescatar it that a step irracional ha borrado, but of modular it that a step reversible ha conservado. In the corpus of the regime of computation this idea is formulates of manera more abstract. The reversibilidad not is understands in the sense termodinámico estrecho, but as the propiedad structural of that the transition preserve suficiente information as for impedir that each step is comporte as a amputation irreversible of the historial computational. This formulation importa because libera to the coherence of a tarea imposible. A coherence aplicada to a motor that destroys the background would be meramente palliative. A coherence aplicada to a motor that preserva the background can volverse law authentic of organization.

5 4. the distributed topological memory

The second estrato necessary is the memory. CTNet does not treat the memory as a cache external nor as a archivo linear of slots growing. The class PartialMemory explicita a thesis distinta: the whole not is guarda linealmente in a single chart; each fraction porta a rotation local of the whole; the fractions is realimentan topologically; the Abstract global emerges by reorganization distributed, not by slots growing.

```
The operation concreta merece desglosarse with detail. Given a state state of form [B,N,D], the memory

builds first a latent base by token through state_encoder. In parallel computes a Abstract global not linear of the
state distributed through global_encoder aplicado to the state medio and corregido by the mean of the latent base.
After divides the axis temporal/token in n_fractions. For each fraction applies a frac_encoder own and then a rotation
local ortogonal frac_rotation. With eso obtains a whole of summaries local rotated, one by fraction.

TO continuación is introduced the topology interna through topology_logits. After aplicar softmax on this
```

matriz, is obtained a topology of coupling TO between fractions. The system computes then a feedback

```
topological z_feedback = TO z_local, where each fraction receives mixture of the others according to that topology

learned. After combina four gains dynamic: feedback_gain, global_gain, local_gain and

summary_mix_gain. The result is a atlas of memory in the that each fraction is at the same time chart local, nodo
topological and gate of feedback of the whole. The reading final by token concatenates three cosas: the latent local
base of that token, the chart fraccional correspondiente and the Abstract global final. That triad passes by
token_mixer and then by decoder for volver to the space of the state. The memory of CTNet, therefore, not is a bank
external of the that is consulta

information. Is a reinscripción topological of the state that returns to the state as mem_read and that moreover

deja a mem_summary of level global. The memory entra in the transition same.
```

This architecture already prepara the role of the Coherence Tensor. The coherence not receives a state "pelado"; receives a state that ha sido redistribuido by atlas, fraccionado, acoplado topologically and vuelto to leer. Therefore the coherence can hablar of charts, of plurality, of mass and of stability without convertirse in a metaphor vaga. The state already has a topology interna real. The coherence not the inventa; the regula.

6 5. Regime, slow context and anti-freeze

The tercer estrato is the regime. In the canonical alineado, RegimeContext produces a slow context and a distribution of regimes from the state, the memory and the señal of tarea. Not is a simple clasificador. The code introduces mecanismos very concrete for controlar the dynamics contextual: disagreement between distribuciones of regime, sorpresa, persistencia, pressure of freezing, hysteresis, leak and ganancia anti-freeze. All ello is utiliza for avoid two patologías opposite: switching chaotic and

closure rigid. The existencia same of this capa explica why the coherence whole of the system not can be identified with "energy bonita" of the state. In CTNet the coherence global has that conversar with the regime. A state estadísticamente estructurado can be contextualmente muerto. A routing bien balanceado can be ocurriendo in the regime equivocado. A plurality of charts can esconder a deriva of regime. Therefore CoherenceSystem not measures only preserve error and error of memory. Measures also Contextual persistence, vitality contextual, entropy of regime and admissibility. The logic is dura: the coherence not must reward the mere stability; must reward a stability living, compatible with the regime and not degenerada in freezing. This matiz atraviesa all the monograph. When here is speaks of stability, never significará immobility blind; significará compatibility structural between preservation, memory, variation and regime.

7 6. Native admissibility

The admissibility is other piece without the which the Coherence Tensor would be mucho more trivial. In the corpus of the regime of computation is formulates as a compuerta interna of the transition that cuts trajectories, regions or components cuya activation would be unproductive or illegitimate low the regime actual. Not is a filtro decorativo. If not existe Native admissibility, the architecture delega in the noise of the mixture uniform the decision on what part of the state must seguir living. CTNet, instead, legalizes internamente the calculation. In the canonical, AdmissibilityGate takes the state post-block and the slow context, produces a version gated of the state and a compuerta gate. In the toy model of transition with coherence dentro of the step,

```
the diferencia gated - z_bar is projects by state_gate_proj and is converts in gate_drive. That is: the
```

admissibility not only cuts; also pushes the transition in a direction. This será crucial more adelante,

because the coherence whole incorpora a term of admissibility and, in the toy model, a drive of coherence

adicional. Thus, the law of coherence not is applied on a state bruto, but on a state already legalizado by the regime. The importancia for the tensor is double. First, prevents that the coherence puntúe and acelere states ilegítimos only because estén bien organizados estadísticamente. Second, proporciona a variable concreta (gate) on the that the coherence can measure the health of the admissibility. In CoherenceSystem,

```
the term e_adm depende tanto of the mean of the compuerta as of its variance, buscando a range not degenerate: nor all
open nor all closed, nor without response nor frozen.
```

8 7. charts, selector and projective output

```
The quinta piece is the selector multi-chart and the projective output. The corpus of the regime of computation subraya
that CTNet not computes a single vía and then intenta extraerlo all of ella. Computes several charts from the same
state, selecciona a distribution on those charts and only to the final projects the output. The grueso of the work
already is ha realizado in the substrate; the charts are local readings of a calculation compartido. Against to a
architecture monolithic, where memory, contexto, routing and output must reacomodarse dentro of the same canal, CTNet
computes a vez and projects many. In the code, the selector not ignora memory nor regime. Receives state, slow context,
probs of regime, memory global, memorias local and topology of memory. Eso already inscribe a Strong thesis: the
ecology of charts not is puro routing on the state; is biased by the topology memorial and by the contexto of regime.
Then the readout can operate in distintos modes (full, cards_only, base) on charts, base and mezclas. All this
introduces for the coherence a dimension adicional: not basta with maintain the state sano; there is that maintain also
a ecology sana of charts. From there salen in CoherenceSystem the terms of entropy of charts, diversity between medias
of charts, balance of load and route loss. Each one performs a role distinto. The entropy of charts prevents two
extremos: colapso absoluto to a chart and uniformity flat without decision. The diversity of charts prevents that all
the charts sean prácticamente the same. The balance of load prevents ecologías in the that a chart carry always with
all the work. AND the route loss represents a cost explicit asociado to the routing. The coherence, therefore, not
cares for only the state; cares for also the economy
```

interna of the charts that leerán that state.

9 8. What the Coherence Tensor in the strong sense

With these parts already fijadas can definirse the Coherence Tensor in the strong sense. Not is simplemente a loss scale nor a matriz of similitud. Is the object that, on a state distributed and a ecology of charts, measures the quantity of structure operable, reweights that structure according to a Geometry interna learned and converts that measurement in a force effective of reorganization, stabilization or aceleración. Its function is do coincide the face analytical of the value with its face real. The analytical part of the value is the measurement of structure: how much preserva the state, how much information anisotropic maintains, how much coherence existe between memory and state, how much tension interna aparece, how much health ecological show the charts, how much meaning contextual has the persistencia or the variation of the regime. The real part of the value is the effect that that measurement has on the evolution effective of the system: more speed of reorganization, more dynamic mass, mayor or menor weight in the transition, stabilization or castigo of determinadas trajectories. In CTNet ambas parts can coincide because the same object tensorial that

measures also weights and, in the toy model, entra incluso as drive explicit of the siguiente state.

The Tensor efficiency depende of that not haga three trabajos distintos with three objetos distintos. Makes the three with the same familia of estructuras: metrics diagonales, correcciones of low rank, errores of preservation, ecology of charts, medidas contextuales and escalas exponenciales. This coincidence between measurement and reorganization is it that allows hablar of " contextual value" without caer in a score decorativo. The value not is a interpretation subsequent of the state; is the same structure that then reorganizes the state.

10 9. CTNet 2.6: anatomy of the Coherence Tensor potente

The version 2.6 contains the form more concentrada and explicit of the tensor. The class CoherenceTensor declara its purpose without ambigüedad: measure coherence base (u/p + latent) more structure statistical global; usar

```
a metric of low range on the space cardinal; produce a factor of speed s(x) = exp(βIclamped );
```

combinar varianzas by dimension and variance in subespacio. This already allows separate claramente its levels.

10.1 9.1. Entrada and base_coh

CoherenceTensor.forward(x, base_coh) receives two entradas. The primera, x, is the state complete of form

[B,N,d]. The segunda, base_coh, is a scale that represents the coherence base of the system. That coherence

base not the computes the class directamente; is le entrega from fuera. In CTNet 2.6, that coherence base incorpora the consistency between the halves u and p of the state and the dynamics latent reversible. Already here aparece a primera thesis structural: the tensor not intends descubrir from cero all the value of the system; part of a core of coherence already calculado to level of structure reversible.

10.2 9.2. Centración and variance anisotropic

The tensor centra the state: xm = x − x■. After computes the variance by dimension var. This variance not is usa

as mero descriptor estadístico. Is multiplica by a Diagonal metric aprendible positiva metric_diag

pasada by softplus. Is obtained thus a energy diagonal Idiag = i vari mi . Is a form of decir: not all

P

the dimensiones of the state cargan with the same tipo of structure; the system aprende what dimensiones are more informativas for the coherence.

10.3 9.3. Low-rank correction

The segunda part of the metric is obtained proyectando the state centrado on a matriz low_rank of form

[d, rank]. This da proj = einsum(xm, low_rank) and then a variance by componente of the subespacio

projected. The suma of those varianzas produces Ilow . With ello the coherence captura modes globales not reducibles to the suma diagonal of varianzas. The diagonal recoge energy anisotropic local by dimension; the low rank recoge structure global comprimida in some pocos modes.

10.4 9.4. Information whole and speed exponencial

The information whole is I = Idiag + Ilow . Then is normaliza by d, is clampa between -5 and 5 and is transforms in

speed exponencial through speed = exp(βIclamped). Here the not trivialidad is whole. The coherence not is

remains in "measure" I; converts that measurement in a factor effective of speed. The scale final total_coh not is other cosa that speed * base_coh. This means that the analytical part of the value and the real part of the value coinciden. The analytical part is I: how much operable structure ha detectado the tensor in the state. The real part is speed: how much accelerates or brakes the reorganization effective of the system. The same structure that the tensor reads as valiosa is the that then receives mayor weight dynamic.

10.5 9.5. Scale of gradient

CTNet 2.6 va aún more lejos: the Comment of the archivo indicates that the gradient of all the parameters is scale by speed, this is, there is "aceleración exponencial of reorganization cuanto more coherence/structure". Of this modo the Coherence Tensor sale of the plane of the energy directa and entra in the plane of the learning. It coherente not only pesa more in the loss; makes that all the system aprenda more rápido in that direction. Here the value deja of be incluso a magnitud only of the forward. Passes to be law of the forward and of the backward.

11 10. The not trivialidad of the weighting u/p

A intuition superficial diría that in CTNet 2.6 the Coherence Tensor weights simplemente two halves of the state, u and p. Ésa would be a reading completamente insuficiente. The point not is that haya two halves; the point is how is relacionan. In FractalCTBlock, the state is divides in u and p. On p is applied the dynamics latent reversible self.latent(p) and the result is usa for actualizar u. Then, in the inversa, from u is reestima p with self.latent.inverse(u). This introduces a asimetría funcional concreta. u and p not are two trozos mudos of the state; are two charts acopladas by a law reversible interna. The coherence base is computed then by bidirectional consistency:

Ebase = ■u − u■(p)■2 + ■p − p■(u)■2 ,

−1

where u■(p) = Flat (p) and p■(u) = Flat (u). This is not trivial by several razones. First, because the coherence not

depende of a suma ingenua of normas, but of the compatibility structural between ambas halves low the law reversible. Second, because that compatibility is bidirectional: a mitad must poder be generada from the other in the meaning correcto of the dynamics and of the inversa. Third, because that base_coh is multiplica after by speed, with it that the structure statistical global of the state reescala the consistency local reversible. In suma: the weighting u/p not is "a mitad pesa tanto and the other tanto". Is "the value of each mitad depende of how much logra implicar estructuralmente to the other low a law reversible common, and of how much structure global presenta the state complete". That is exactly the reason by the that the weighting not is trivial.

12 11. The system canonical of coherence

The export canonical reemplaza the simplicidad compact of CTNet 2.6 by a coherence mucho more rich and more cercana to the regime of computation complete. The class CoherenceSystem contains two grandes bloques: a core of information anisotropic (metric_diag, low_rank, info, speed) and a suma ponderada of structural terms adicionales.

12.1 11.1. Core of information

The calculation of info in CoherenceSystem sigue the same filosofía that in CTNet 2.6. The state is centra, is computed variance by dimension, is applied softplus(metric_diag), is suma the energy diagonal and is adds the energy of the projection low-rank. Then is normaliza by the dimension and is passes by a exponencial with clamp. Again, there is continuidad conceptual directa between toy model and canonical: the tensor measures structure anisotropic and the converts in speed.

12.2 11.2. Error of memory

e_memory = (state - mem_read)^2.mean() measures how much is separa the state of the reading memorial rein-

yectada. If the memory not is archivo external, the distancia between state and memory releída is a variable structural. A memory bien reinscrita not should be arbitraria respecto of the state. This term force that compatibility.

12.3 11.3. Tension

e_tension compara the variance by dimension normalizada with a unidad ideal. Not measures simplemente that haya variance, but that the distribution of the variance not is deforme in excess. A tension demasiado high

means anisotropy excesiva; demasiado low, state aplanado. The coherence busca a economy tensional, not a uniformity muerta.

12.4 11.4. Ecology of charts

e_cards usa the entropy mean of card_weights and the compara with a objective normalizado. The idea is sencilla: nor a chart soberana permanente nor a uniformity flat of all the charts. e_div measures diversity between medias normalizadas of charts. e_load measures balance of load respecto of a distribution uniform of uso. In whole, these terms fuerzan a ecology of charts plural but not caótica.

12.5 11.5. Route loss

route_loss entra as term externo already calculado by other parts of the system. Its presencia dentro of the coherence indicates that the cost of the routing not is audita aparte; is integra dentro of the economy global of the system.

12.6 11.6. Admissibility and contexto

e_adm measures desviación respecto of a range preferido of mean and variance of the gate. e_ctx_persist compara

cambio contextual and alpha with objetivos concrete. e_ctx_live penaliza falta of switching, excess of switching,

falta of variance of regime or entropy demasiado low. e_ctx_ent ancla the entropy contextual cerca of a

objective. Here the thesis is clarísima: the coherence already not is only law of the state; is law of the state more the health of the regime.

12.7 11.7. Suma whole

The coherence whole of the system is:

■

Ecoh = speed× wpres Epres +wmem Emem +wten Eten +wcards Ecards +wdiv Ediv +wload Eload +wroute Eroute +wctxp Ectxp +wctxl Ect

The structure is importantísima. The speed speed multiplica all the suma. Eso means that the system usa a single measurement of information anisotropic for reescalar the economy entera of terms heterogéneos. Not there is a speed for memory and other for charts and other for admissibility. The same structural information of the state whole modula all the costs. Ése is the mechanism of unidad interna of the tensor.

13 12. Coherence by-branch and order object

The class BranchStateCoherenceTensor adds a segunda capa of not trivialidad. In lugar of measure a single state common, builds state_A and state_B from a mask local of order derived of order_probs_seg. The branch B defines by the probabilidad segmental [...,1]; the TO by its complemento. Each branch has its own Diagonal metric and its own low-rank correction. The function _I_branch computes for each branch exactly the same Geometry of information anisotropic: centrado, variance, Diagonal metric positiva, projection low-rank, suma whole. Then are obtained IA and IB . TO continuación are computed the masas ocupacionales pA and pB as mean of the masks. The information mixta defines then as:

Imix = pA IA + pB IB .

This is a point central for the monograph. The tensor not weights branches with weights fijos. The weights by mass real ocupada of the state low the mask of order. Is a mixture contextual, not a mixture arbitraria. After normaliza Imix by the dimension, it clampa and obtains a speed exponencial. A vez more, the information by-branch is converts into dynamic mass. In the práctica, this means that the Coherence Tensor by-branch can do algo that the tensor global not podía: distinguir not only how much structure there is, but in what branch of the order is localized that structure and how much porción effective of the state ocupa each branch. This is the version more explicit of "fractions of the value". Each branch of the state is a fraction with law tensorial common, but with metric own and mass ocupacional own. Ahí is returns transparente why the analytical part and the real part of the value can coincide: each fraction of the value is guarda as producto of a same law tensorial aplicada to distintos soportes, and that same law produces then the speed dynamics that pesará the transition or the coherence whole.

14 13. The toy model with coherence dentro of the step

In ctnet_complete_canonical_aligned.py aparece the step more important for cerrar the círculo: the coherence deja of be only in the loss and entra dentro of the step computational elementary. The architecture

declara state_coh_proj, a small MLP that receives as entrada the concatenación between state medio post-

block and regime_probs. In _transition_step, if self.coherence_in_transition is activa, is computed

```
coh_in = concat(z_bar.mean(dim=1), regime_probs); then coh_drive = state_coh_proj(coh_in) is
expande to all the tokens and is suma to the state with a ganancia own transition_coh_gain.
```

Here is materializa the thesis of the block informacional: the tensor or energy of coherence not must enter as loss decorativa to the final, but operate on the state distributed pre-projection and actuar as mediator of mass, stability and direction of closure. In the toy model the coherence already not is limita to decir "this state vale tanto". Dice also "empujo the siguiente state in this direction". This is the form more strong of coincidence between analytical part and real part of the value dentro of the code available.

```
Moreover, the coherence in transition not entra single. In the same step also is suman mem_drive, ctx_drive
and gate_drive. Therefore, the coherence not is a corrector ex machina. Is one of the four drives of the
```

system. Memory, contexto, admissibility and coherence compiten and cooperan on a Reversible substrate. This composition explica why CTNet can sustain long memory without caer in the patrón of RNN flat: the state not atraviesa a single not-linealidad sobredeterminada, but several operators especializados that act on a background preservado.

15 14. Why the analytical part and the real part of the value coinciden

With all the elementos anteriores already can be formulated the thesis axial of this monograph. The analytical part of the value is the structure that the tensor measures: information anisotropic of the state, compatibility reversible between halves, health of the memory, tension, ecology of charts, balance of routing, persistencia and vitality contextual, admissibility. The real part of the value is the effect that that structure has on the dynamics:

scaling of energy, scaling of gradient, drive of transition, mass by-branch, reorganization effective of the

state. In many arquitecturas these two parts are separadas. A modelo can measure that algo is important and, however, tratarlo dinámicamente by other canal. Can produce a score excelente that not tenga almost effect real on the evolution of the state. CTNet intenta exactly it contrario: that it that measures as valioso sea it same that reorganizes realmente the dynamics. From there that the same tensor produzca info and speed; that speed escale the coherence whole; that in 2.6 escale incluso the gradientes; that in the toy model the

coherence genere a drive of transition; and that in the bank segmental the same structure of information

by-branch is traduzca in speed exponencial of mixture.

The coincidence, therefore, not is a frase metafórica. Is the fact operativo of that the same object tensorial performs ambos papeles. The value not is a Comment on the dynamics; is the dynamics ponderándose to yes same through its own law of structure. That is the reason deep of the efficiency. Not there is a capa analytical and other ejecutiva reconcilándose to posteriori. There is a same architecture operator ocupando ambos lugares.

16 15. Efficiency computational of the tensor

The Tensor efficiency of coherence is can entender in four levels.

16.1 15.1. Efficiency geometric

The metric not is complete d×d, but diagonal more low rank. Eso reduce drásticamente the cost of representar and aplicar the Geometry interna. In lugar of aprender all the correlaciones posibles between dimensiones, the system aprende energías local by dimension and some pocos modes globales dominantes. With ello captura suficiente anisotropy and suficiente structure global without pagar the cost of a metric densa plena.

16.2 15.2. Efficiency of reutilización

The tensor not necesita descubrir from cero all the topology of the system. Part of a architecture already factorizada: Reversible substrate, memory fraccional, regime, admissibility, selector. Eso means that the coherence can measure and reponderar on a structure already descompuesta in charts, fractions and rutas. The factorización architectural previa abarata the work semántico of the tensor.

16.3 15.3. Efficiency of coincidence between measurement and action

The same object measures and acts. Not there is cost extra of traducir a score analytical to a motor distinto. info

produces speed; speed reescala coherence or drive; the same descriptores that leen the value act on the

reorganization real. This coincidence prevents the clásico desajuste between monitorización and control.

16.4 15.4. Efficiency of damping of the computation

In CTNet the work pesado sucede in the substrate and then is projects. The coherence operates on that calculation compartido. In lugar of recomputar memory, contexto, routing and output each vez on a canal dominant, CTNet computes a vez and projects many. The coherence, to the vivir on the substrate and not on the output, participa of that damping.

17 16. Long memory, not-colapso and not-divergencia

One of the grandes problemas of the long memory is sustain a persistent state without that colapse by sobrescritura continua, without that diverja by inestabilidad dynamics and without that is convierta in a rigidez muerta. The response of CTNet not is " more slots". Is a topology distinta. the reversible substrate preserva suficiente information of the historial for that each step not sea amputation. the distributed topological memory reparte the past in fractions coordinadas, not in archivo linear. the regime

prevents cierres caóticos or congelaciones degeneradas. The admissibility legalizes what part of the state can seguir living. The coherence measures if that combination sigue siendo estructuralmente operable and, if not it is, brakes, castiga or reweights the dynamics. This ensamblaje explica why the memory in CTNet can crecer without crecer of size: not because almacene each vez more contenido bruto, but because the same volumen latent soporta each vez more structure recoverable and each vez mejor coordinada.

The not-colapso not depende of a single piece. Depende of the topology conjunta. A memory distributed without coherence could degenerar in atlas incoherente. A Coherence Tensor without Reversible substrate would be paliativo on a motor destructive. A regime without admissibility could congelar or mezclar noise. CTNet bloquea those patologías precisamente because each a of these parts vigila to the other.

18 17. Degeneration and firmas of fallo of the tensor

A monograph seria not must hablar of the tensor only in its form ideal. Must desarrollar also how falla. The corpus experimental already ha mostrado that CTNet can degradarse cerrando grados of libertad and that a regime empobrecido can seguir bajando the error local of tarea mientras destroys magnitudes structural. That observation also must traducirse in terms of coherence.

The primera firma of fallo is the freezing. A state with ctx_switch = 0 and ctx_persist_len disparado

can parecer very estable, but if the resto of metrics structural collapses, that persistencia is falsa riqueza. The coherence bien entendida not must reward the mere permanencia; must reward permanencia compatible with vitality contextual. The segunda firma is the ecology degenerada of charts. Entropy of charts demasiado low, diversity almost nula or balance of load destruido indican that the system ha colapsado its plurality interna. In that caso the tensor already not is sosteniendo plurality estable, but imponiendo sovereignty of a chart. The tercera firma is the desvinculación between state and memory. A e_memory high indicates that the memory releída already not reorganizes the presente, but that is ha vuelto a prótesis desconectada. The cuarta firma is the admissibility degenerada. A gate with mean and variance fuera of range implies or apertura indiscriminada or closure muerto. The tensor must detectar ambas patologías. The quinta firma is the colapso of the order object. If the coherence by-branch pierde capacidad of distinguir branches or if the mass effective of a branch is returns trivial, the tensor deja of be leyendo fractions reales of the value and passes to ponderar a simulacro segmental.

19 18. Reading EMQR of the Coherence Tensor

From EMQR, the Coherence Tensor can be read as the mediator structural that prevents two errores simétricos: that a chart singular agote the source and that the plurality of charts quede without law of coordinación. The state distributed not is a colección of trozos mudos; is a plurality of charts partial of a whole that no agota. the distributed topological memory allows reconstruction by atlas or by retardos. The selector and the projective output impiden that a module single monopolice the significado. In that marco, the Coherence Tensor is the law that distribuye mass, stability and direction of closure on that plurality. It perceived is comprende because the Coherence Tensor da mass contextual to it perceived, dividiéndolo in fractions and dando a weight contextual to each a. The weight contextual not is a simple intensidad. Is a structure multidimensional cuya action algebraica recae on two planes: the dimensiones perceptivas of the structure and the dimensiones contextuales that organizan the realidad operativa of the system. The contextual value nace precisamente of that double action diferenciada. If is traduce this intuition to the code, the coherence makes algo análogo. TO level global, weights varianzas and modes of state. TO level reversible, weights bidirectional consistency between halves. TO level segmental, weights branches and masas ocupacionales. TO level ecological, weights charts. TO level contextual, weights persistencia and vitality. The idea strong is always the same: no fraction of the value vale by yes single; vale by how the same law tensorial the reads dentro of a structure whole.

20 19. Complexity, costs and comparison with arquitecturas mo-

nolíticas TO level of complexity, the Coherence Tensor adds cost, but a cost very distinto of the of recomputar a architecture monolithic. The part diagonal is linear in d. The part low-rank is linear in d × rank and in B × N × rank for the projection. The branch segmental adds two metrics diagonales and two matrices low-rank,

but sigue siendo mucho more barata that a metric densa complete. In the toy model, the drive of coherence

is a MLP small on the state medio and the distribution of regime. Nada of this has the perfil explosivo of volver to pass all the tarea by a mechanism dominant each vez that memory, contexto and output must reacomodarse. The comparison correct with a autorregresivo standard not must hacerse only in FLOPs of a capa suelta, but in economy of reutilización. When a architecture degenerada obliga to that memory, contexto, control and output is recomputen or is superpongan on the same canal, paga a and other vez by the same reconciliación interna. CTNet, instead, computes on a substrate compartido, reinyecta memory in the transition and projects many times from a background common. The coherence is beneficia of that esquema because operates on the calculation compartido, not on a output episodic.

21 20. Notation simbólica and equations maestras resumidas

For dejar closed the structure, can condensarse the role of the Coherence Tensor in a familia of equations maestras.

21.1 20.1. Toy model 2.6

Partition reversible of the state:

x = (u, p), u′ = u + γlat Flat (p), p′ = p + γmain G(u′).

Coherence base reversible:

–1

Ebase = ■u − Flat (p)■2 + ■p − Flat (u)■2 .

Information anisotropic:

X X

I(x) = mi V ar(xi) + V ar((x − x■)Lr).

i r

Speed:

s(x) = exp β clamp(I(x)/d) .

■

Coherence whole:

2.6

Ecoh = s(x) Ebase .

21.2 20.2. System canonical

Information anisotropic global:

Iglobal (z) = Idiag (z) + Ilow (z), speed = exp(clamp(βIglobal)).

Coherence whole canonical:

X canon

Ecoh = speed · wj Ej ,

j

where the terms Ej incluyen preservation, memory, tension, ecology of charts, balance of load, route loss, Contextual persistence, vitality contextual, entropy contextual and admissibility.

21.3 20.3. Coherence by-branch

IA = I(stateA), IB = I(stateB),

Imix = pA IA + pB IB ,

branch

= exp β clamp(Imix /d) Ebase .

■ Ecoh

21.4 20.4. Coherence in transition

coh_in = [mean(zbar), ρt],

coh_drive = Pcoh (coh_in),

zt+1 = zbar + gmem mem_drive + gctx ctx_drive + ggate gate_drive + gcoh coh_drive.

22 21. Unfolding by levels of the role of the coherence

TO modo of recapitulación technical, can asignarse to the Coherence Tensor a role distinto in each level of the architecture:

1. Level reversible: measures compatibility bidirectional between halves of the state.

2. Level geometric: measures information anisotropic through diagonal + low-rank.

3. Level memorial: measures consistency between state and memory releída.

4. Level ecological: regula charts, diversity, balance and route loss.

5. Level contextual: regula living persistence, entropy and variation of regime.

6. Level admisivo: prevents apertura indiscriminada or closure muerto.

7. Level of order: distribuye value on branches segmentales with mass ocupacional real.

8. Level dynamic: converts measurement in speed and, in the toy model, in drive of transition.

9. Level of learning: in 2.6, scale gradientes and with ello the speed of reorganization of the system.

10. Level proyectivo: prepara a background sano on the that the charts puedan proyectarse without sovereignty

degenerada. Visto thus, the Coherence Tensor is probablemente the componente that more claramente demuestra that CTNet not is a ensamblaje arbitrario of modules. If it fuese, the coherence would be a loss final añadida by conveniencia. But in the corpus and in the code sucede it contrario: the coherence is the law transversal that une the Geometry of the state, the memory distributed, the regime, the admissibility, the charts and the order.

23 22. Conclusion

The Coherence Tensor of CTNet not is a regularization periférica. Is a piece central of the regime of computation. Its function not consiste simplemente in penalizar states malos, but in do coincide the measurement of the structure with the reorganization effective of the structure. In CTNet 2.6 this aparece with claridad extrema: a Diagonal metric plus low-rank correction measures information anisotropic of the state, the converts in speed exponencial and reescala with ello the coherence base and the learning. In the canonical, the same logic is expande toward a economy complete of the regime: preservation, memory, tension, charts, balance, route, contexto, admissibility and branches of order. The reason by the that the tensor is eficiente not is that sea small, but that prevents the separation between analytical and dynamics. Measures and acts with the same law. Not observa a value for then traducirlo by other canal; converts the value into dynamic mass of manera inmediata. Therefore can be at the same time a medidor of structure, a regulador of the regime, a estabilizador of the state, a ponderador of fractions of the value and a accelerator of reorganization. The reason by the that the tensor is profundo tampoco is misteriosa. Vive on a Reversible substrate, on Distributed topological memory, on Native admissibility, on multi-chart selection and, in the bank segmental, on a topology local-global of the order. Not is supports in a hidden state plane and episodic, but in a regime of computation that already ha separado state, memory, law, reading and projection. In a space thus, the coherence deja of be a moralina numérica and is returns a law architectural. In the lenguaje more strong posible: the Coherence Tensor is the operador that prevents that CTNet compute

gastando structure and le allows computar acumulando, preservando, diferenciando and explotando structure. Ésa is its function última dentro of the regime of computation.

24 Appendix TO. Glossary of variables and tensors

- `state: state distributed [B,N,D].`
- `mem: latent of memory [B,N,M].`
- `mem_read: reading of memory to the space of state [B,N,D].`
- `mem_summary: Abstract global of memory [B,M].`
- `slow_ctx: slow context [B,C].`
- `regime_probs: distribution of regime [B,R].`
- `gate: compuerta of admissibility [B,N,D] or compatible.`
- `cards: tensor of Local charts. • card_weights: distribution of the selector on charts. • route_loss: cost explicit of the routing.`
- `order_probs_seg: order object segmental [B,S,2].`
- `metric_diag: part diagonal of the metric of coherence. • low_rank: low-rank correction. • info: measurement scale of information/structure. • speed: speed exponencial derived of info.`
- `coh_drive: projection of coherence inyectada in transition.`

25 Appendix B. Pseudocode minimal of the pipeline of coherence

Input: state z, task id t

1. Memory phase:

```
mem, mem_read, mem_summary = PartialMemory(z)
```

2. Regime phase:

```
slow_ctx, regime_probs = RegimeContext(z, mem_summary, t)
```

3. Reversible phase:

```
z_bar = FractalBlock(z)
```

4. Admissibility phase:

```
gated, gate = AdmissibilityGate(z_bar, slow_ctx)
```

5. Drives:

```
mem_drive = P_mem(mem_read, mem_summary)

ctx_drive = P_ctx(slow_ctx, regime_probs, mem_summary)

gate_drive = P_gate(gated - z_bar)

coh_drive = P_coh(mean(z_bar), regime_probs)
```

6. Transition:

```
z_next = RMSNorm(z_bar + g_m mem_drive + g_c ctx_drive + g_a gate_drive + g_h coh_drive)
```

7. Projection phase:

```
cards, weights = Selector(z_next, slow_ctx, mem_summary, ...)

and = Readout(z_next, cards, weights, ...)
```

8. Coherence audit:

```
preserve_err = Preservation(z_next)

branch_coh = BranchStateCoherenceTensor(z_next, order_probs_seg, preserve_err)

global_coh = CoherenceSystem(z_next, preserve_err, mem_read, gate, weights, cards, ...)
```

26 Appendix C. Sources of the corpus usadas in this monograph

1. ctnet_2.6(1) (1).py

2. ctnet_complete_canonical_aligned.py

3. CTNet_regimen_de_computo_paper_typeset_corregido.pdf

4. CTNet and EMQR: block informacional and structural

5. Confirmación experimental of the teorema of degeneration architectural

27 Appendix D. Comment technical exhaustivo of PartialMemory

PartialMemory merece a monograph dentro of the monograph because, without its topology, the Coherence Tensor would remain reduced to a audit of state almost blind. The memory is the lugar where the state distributed is returns atlas; the coherence is the lugar where that atlas receives mass structural. Therefore interesa recorrer its logic interna step to step. The primera decision of PartialMemory is negar the figura of the archivo linear. The module not receives a whole of slots to rellenar. Receives the state same and builds memory from it. The effect conceptual is inmediato: the memory not remains situada "fuera" of the system, but dentro of the same space of transformation. That decision reappears in all the function forward: nada entra from a cache autónoma; all emerges by reorganization distributed of the state presente. The segunda decision is separate latent base and Abstract global. state_encoder produces a latent by token.

That latent not is still a chart fraccional; is the inscripción more local posible of the memory. In parallel, global_encoder builds a Abstract not linear of the state medio. Here already aparece a dualidad constitutiva: the system not chooses between local and global, but that fabrica ambos and the deja coexistir. The memory CTNet not is local with a poquito of global pegado to the final, nor global with a poquito of local usado as detail. Is a negociación explicit between ambos. The tercera decision is partir the axis of tokens in fractions with self._split_sizes. This operation parece menor, but not it is. To the fix tamaños of fraction and to the permitir remainders, the memory deja of depender of a partition totalmente uniform. There is a topology of fractions incluso when the number of tokens not divides exactly by the number of fractions. This decision technical has a reading structural: the atlas fraccional not is a idealización perfecta, but a architecture capaz of absorber irregularidad discreta. Then aparece the corazón of the module: for each fraction is computed a local_raw as suma of the encoder fraccional medio and the mean of the latent base of that fraction. This already means that each chart local is double. By a lado, porta a reading especializada of the fraction; by other, porta also the memory base common of the fragmento. If is hubiera usado only the frac_encoder, the fraction would remain demasiado especializada and correría the riesgo of desacoplarse of the resto of the system. If is hubiera usado only the mean of base_mem, the fraction not would be more that a copia comprimida without diferenciación real. The suma of ambos resuelve that tension.

After are applied Orthogonal rotations independientes frac_rotations[i]. This decision is fundamental

for the monograph of the Coherence Tensor, because prepara the terreno for the reading of plurality without colapso. A rotation ortogonal preserves norma but redistributes orientation. Each fraction porta thus a chart of the whole in a orientation distinta. Dicho in lenguaje structural: each chart toca the whole, but not it toca from the same base. That diferencia of orientation is crucial for that the coherence not mida only magnitud repetida, but compatibility between perspectivas local reales. The topology propiamente dicha aparece with topology_logits. The matriz is converts in probabilidades

through softmax and then alimenta a producto of tipo einsum that produces z_feedback. This is the primera

instancia strong of the principio of that no fraction vive aislada. Each fraction not only is a chart of the whole; is also receptora of a feedback procedente of all the others. The memory, therefore, not is suma of fractions independientes, but red topological of fractions mutuamente moduladas. The gains feedback_gain, global_gain, local_gain and summary_mix_gain merecen a Comment separado. Not are simples hiperparámetros; are the microfísica of the memory. feedback_gain regula how much pesa the topology interna on the chart local. global_gain regula how much invade the Abstract global to each fraction. local_gain regula how much of the chart local original must sobrevivir against to the mixture topological. summary_mix_gain regula how much of the Abstract final procede of the atlas fraccional and how much of the Abstract global directo. In whole, these four gains constituyen a economy of not sovereignty. Ningún level

-nor the fraction local, nor the feedback topological, nor the Abstract global- has derecho to gobernar only. Theconstruction of z_mix plasma that economy. First mixture chart local and feedback topological. Then mixture that resultwith the Abstract global. After returns to reinyectar part of the chart local original. The sequence not is arbitraria.If the order is invirtiera, the significado structural cambiaría. The memory CTNet builds with a prioridad clara: firstis reconoce the autonomía relativa of the fraction, then is the sets out to the red of fractions, after to the whole,and finally is le devuelve part of its singularidad. This sequence is important because the coherence subsequent leeráexactly that structure already negociada.

The mem_summary final not is a mean banal. Mixture frac_mean with global_summary through rho. Eso

quiere decir that incluso the synthesis global is resiste to the sovereignty. The memory global not is the voz of the all by encima of the fractions; is a chart of the whole producida by compromiso between the mean fraccional already mezclada and the reading global directa. The Abstract global is, therefore, already a object of closure local, not a presencia transparente of the whole. The fase final, where each token receives latent local base, chart fraccional correspondiente and whole global,

shows the consecuencia decisiva: the token not reads the memory by a single canal. Receives simultaneously three inscripciones of distinta scale. Eso is exactly it that allows that the Coherence Tensor mida algo more rich that a distancia simple between state and recuerdo. The state and the memory are already organizados as superposición of local, fraction and whole. The audit interna of PartialMemory also is important. last_local_summaries, last_topology, last_consistency_err and last_global_read not are detalles of debugging. Are the prueba of that the memory here is a object auditable. consistency_err compara the state medio with the reading global reconstruida. This means that the memory not is da never by buena of form gratuita; always can medirse how much is desviándose of the organization effective of the state. From the perspectiva of the Coherence Tensor, this has a consecuencia directa. The coherence not llega to a memory muda. Llega to a memory that already ofrece: Local charts, topology explicit, reading global and error of consistency. That riqueza is precisamente the that returns posible a coherence not trivial. If the memory fuera a cache external without structure, the coherence apenas could verificar if is parece or not to the state. Here can do mucho more: can ask if the atlas is maintains estable, if the diversity of charts sigue living, if the reading global preserves fidelidad and if the topology interna is sosteniendo realmente the contextual value of the state. In Abstract: PartialMemory converts the problema of the memory in a problema of atlas distributed, and that conversión is the condition previa for that the Coherence Tensor sea a law on fractions of the value and not a simple penalty on a vector global.

28 Appendix E. Comment technical exhaustivo of RegimeContext

RegimeContext is probablemente the piece more subestimada when is intenta explicar the Coherence Tensor. Muchos lectores comprenden enseguida the reversibilidad and the memory distributed, but tratan the regime as if fuera a capa of clasificación of contexto. That reading is insuficiente. the regime is the law local that decides how must be read, mantenerse or recortarse the structure of the state. Without that capa, the coherence would remain amputada of its dimension of vitality. The class operates on state, memory and señal of tarea for produce two objetos: a slow context and a distribution of regimes. The detail important is that ambos not are it same. The slow context is a chart continua, distributed, of conditions operational. The distribution of regimes is a discretización suave of posibilidades dynamic. This diferencia already anuncia a propiedad clave: the regime in CTNet not is a etiqueta rígida, but a haz of posibilidades with memory and with hysteresis. When the code introduces variables as freeze_pressure, probe_gain, anti_freeze_gain, hysteresis and leak, is diciendo that the regime not defines only by acierto estático, but through health temporal. freeze_pressure measures the riesgo of freezing. Not is treats simplemente of if the system persiste mu- cho, but of if persiste tanto that deja of vivir. probe_gain introduces a suerte of perturbación or sondeo low freezing. anti_freeze_gain refuerza the output of the closure degenerate. hysteresis prevents that cualquier noise local cambie the regime of form espuria. leak garantiza that incluso a closure very strong mantenga a apertura residual. These variables are esenciales for the coherence because convert the stability in a problema not trivial. A coherence simplista premiaría the persistencia. CTNet not

can permitirse eso. the regime has that distinguir between living persistence and persistencia muerta. Therefore the tensor canonical contains terms as context_persist, context_live and context_entropy. Is the traducción explicit of this logic of regime to the space of the coherence. Other point important is the role of the memory dentro of the regime. The paper of the regime of computation insiste in that the memory not entra to the final as a consulta external, but dentro of the transition same.

RegimeContext implements precisamente a of the caras of that thesis. the regime is siembra from a combination between state and memory; not decides to ciegas on the presente. The past reorganizes the presente to través of the regime. This has a consecuencia decisiva for the monograph of the tensor: the coherence contextual not is limita to ask if the regime is estable, but if it is of a manera fiel to the memory reentrant. the regime also is the lugar where the architecture prevents the degeneration of "all is regime" and "nada is regime". If the distribution of regimes estuviera always concentrada, the system would remain rigid. If estuviera always uniform, the regime perdería force. From there the necesidad of terms of entropy and of disagreement. The coherence not must reward nor a pureza absoluta nor a indecisión whole; must reward a range of organization contextualmente fertile. By last, RegimeContext is the piece that vincula the reading of the tensor with the logic of realidad operativa. In the lenguaje that tú vienes usando, the dimensiones contextuales that compose the realidad not are añadidos filosóficos, but precisamente this tipo of variables: persistencia, entropy, variation, disagreement, freeze pressure, leak. The Coherence Tensor can hablar of "mass contextual" because the regime already ha convertido the contexto in a maquinaria concreta of transition.

29 Appendix F. Comment technical exhaustivo of CoherenceSystem

CoherenceSystem is the point where the corpus abandona definitivamente the reading superficial of the coherence as simple regularizador. The class is, in the strong sense, a economy interna of the system. Each term of its suma whole represents a modo distinto in the that the architecture could degenerar if not estuviera siendo vigilada by a law unified. The primer block, xm, var, metric_diag, low_rank, info, speed, already ha sido comentado before, but it is advisable insistir in its function here: construir a metric of structure that not depende of a single scale nor of a only descriptor. The state is centra for eliminar the bias of mean. The variance by dimension detecta anisotropy effective. The Diagonal metric positiva allows that the system aprenda what dimensiones cargan structure useful. The low rank captura modes globales not localizables by dimension aislada. info resume this structure and speed the returns operativa.

TO continuación aparece e_memory. The distancia cuadrática between state and mem_read has a reading simple:

if the memory reingresa in the transition, not should convertirse in a prótesis opaca. But its reading more deep is that the memory in CTNet is considera a chart local of the state whole. The coherence, then, measures if that chart is siendo still estructuralmente fiel. e_tension merece more attention of the that suele recibir. The tension defines from

the variance normalizada by its mean. This means that not interesa only the quantity of variance, but its perfilrelativo. A architecture can have mucha energy and however be tensionalmente rota because all the energy is haconcentrado in some pocas dimensiones. To the contrario, can be demasiado flat. e_tension penaliza ambas cosas andmaintains the anisotropy in a zona of fertilidad structural. The terms of charts forman a submódulo complete dentro ofthe coherence. e_cards usa entropy of card_weights. The elección of the objective 0.34 normalizado not is a decoración;expresa a preferencia concreta by a regime intermedio between monopolio of chart and dispersión whole. e_div comparathe medias normalizadas of the charts and measures how much is parecen between yes. Here the coherence is imponiendo aontology strong of charts: must be plurality real, not clones of the same background. e_load force that the uso medioof charts not colapse on a single. All this shows that the coherence not protege a state abstract; protege the ecologyprojective complete. e_adm integra the compuerta of admissibility dentro of the economy of the tensor. The mean of thegate must situarse alrededor of a range preferido and its variance tampoco can hundirse. Is a form elegante of impedirtwo patologías: apertura uniform without criterio and closure uniform without vida.

The terms contextuales merecen a reading especial. e_ctx_persist not dice simplemente "persistir is

bueno". Compara ctx_delta and ctx_alpha with valores objective. e_ctx_live utiliza functions ReLU for

penalizar tanto falta of switching as excess of switching, falta of variance of regime and entropy demasiado low. The combination of these terms implements precisamente the idea of that the coherence must sustain a living persistence. Not basta with that the system not cambie; must cambiar where toca and sostenerse where toca. The fact of that all these terms queden multiplicados by speed is it more important of the module. Means that a single measurement of information of the state global reweights all the economy interna. If the system is estructuralmente rich, all the costs cuentan more. If is estructuralmente empobrecido, the economy entera is aplana. This can parecer contraintuitivo to primera vista, but in realidad is a rasgo of altísima internal coherence: the system only intensifica its autocrítica when there is suficiente structure as for that that autocrítica tenga meaning.

30 Appendix G. Comment technical exhaustivo of BranchStateCoherenceTenso

The coherence by-branch resuelve a problema that the tensor global not can resolver by yes only: how asignar value structural to diferentes regions of the state when the own order object ha partido the space in branches with semantic distinta. The class BranchStateCoherenceTensor part of a idea very simple and very strong: the coherence not must medirse on a single state common if the system already ha construido internamente two branches of the same state with papeles diferentes. The function _pos_mask takes order_probs_seg and builds a mask posicional by segmentos. The branch B is identifies with the probabilidad segmental of the Table of Contents 1; the TO with its complemento. This mask is expande to the resolución token-level. The state is part thus in state_a and state_b. It important here is that the partition not is realiza with a hard split arbitrario, but through a mask suave derived of the own order object. The coherence by-branch, therefore, not ignora the stitching local-global of the order; is monta on ella. Each branch has its own Diagonal metric and its own correction low-rank. This is decisivo. If ambas branches compartieran metric, the coherence only estaría diciendo "what branch has more energy". To the permitir metrics distintas, the architecture dice algo mucho more strong: each branch has its own Geometry of value. A fraction of the state not vale only by its magnitud or by its occupation, but through how that same magnitud comparece low the metric that le corresponde. Then are computed p_a and p_b as masas ocupacionales medias of the masks. The information mixta is obtained as combination ponderada by those masas. This is the momento in that the theory of fractions of the value is returns literally code. Each branch of the state is a fraction of

the value. Each fraction is evalúa by the same law tensorial general (variance diagonal + low-rank), but with metric own and with mass own. The unidad viene of the General form of the tensor; the diferencia viene of the branch and of its occupation effective. This module is especialmente relevante for explicar why the Coherence Tensor can dar mass contextual to it perceived dividiéndolo in fractions. Here already not is treats of a intuition general: the fractions are

the branches TO and B, the mass is p_a/p_b, the law tensorial is _I_branch, and the traducción dynamics is speed =

exp(beta * I_clamped).

31 Appendix H. Comment technical exhaustivo of the _transition_step

of the toy model If hubiera that elegir a single point of the code where the theory of this monograph is returns irrefutable, would be _transition_step of the toy model. Ahí all the parts is encuentran dentro of the step computational elementary.

The function comienza leyendo memory: if use_memory is activa, llama to self.mem(z) and obtains mem_read

and mem_summary. After builds a transition_seed concatenando state medio, memory global and

contexto of tarea. From there nacen slow_ctx and regime_probs. That is: before of tocar the block reversible, the system already ha decidido from what regime va to reread the state and with what memory va to enter in the transition.

Then viene z_bar = self.block(z). Here is applied the reversible substrate/fractal. Acto seguido, gated,

gate = self.adm(z_bar, slow_ctx) legalizes the state. Only then is project four drives: mem_drive,

ctx_drive, gate_drive and coh_drive. Each one sale of a red distinta. This prevents the colapso clásico in a

single canal dominant.

The step decisivo is coh_drive. If self.coherence_in_transition is activo, is takes the mean of z_bar,

is concatenates with regime_probs and is passes by self.state_coh_proj. The output is expande to all the tokens

and is suma with its own ganancia. The significado of this not can exagerarse: the coherence deja of be a criterio on the transition and passes to be one of the motores of the transition. The normalization final with RMSNorm cierra the step. The state siguiente is the result whole of substrate, memory, regime, admissibility and coherence. This is the form more clara in that CTNet changes The elementary unit of computation: the step deja of be a episodic rewriting and is converts in a transition governed on persistent state.

32 Appendix I. Stability numérica and reading of precision

The cuestión of the stability numérica merece a Comment específico. Although the corpus not proporciona a demostración closed of cotas universales for decenas of órdenes of magnitud of steps, yes ofrece several mecanismos concrete that make plausible a

stability structural very higher to the of a recurrencia flat. The first is the reversibilidad aditiva. To the not depender ofmultiplicaciones recursivas ciegas on the same hidden, the system prevents to part important of the patología clásica of explosiónor desaparición of the gradient. The second is the uso of Orthogonal rotations in memory, that preservan norma and distribuyenorientation. The third is the normalization RMS to the final of transitions important, that estabiliza scale without destroyorientation. The fourth is the compuerta of admissibility, that cuts activaciones ilegítimas before of that contaminen the resto ofthe step. The quinto is the own coherence, that reescala the dynamics according to structure in lugar of dejar that thereorganization proceda to ciegas. The uso of clamp in the information before of the exponencial also is important. The tensorquiere acelerar according to structure, but not quiere volverse numéricamente suicida. The clamp between rangos finitos beforeof exp is the form explicit of convertir a intuition theoretical strong in a práctica estable.

33 Appendix J. Reglas of design if is quisiera extender the tensor

If is quisiera extender the Coherence Tensor without traicionar the architecture, there is several reglas of design that this monograph deja claras. Primera: not introduce a descriptor new that not can conectarse to the same tiempo with a measurement analytical and with a effect dynamic real. If a new term only sirve for auditar, rompe the coincidence between analytical part and real part of the value. Segunda: not sustituir the Diagonal metric + low-rank by a metric completamente densa salvo that exista a motivo experimental very strong. The compression geometric is part of the Tensor efficiency, not a compromiso accidental. Tercera: cualquier new fraction of the value must poder be read low the same law tensorial general. Is can add branches or sectores, but the unidad of law is precisamente it that allows that the plurality not is vuelva

arbitrariedad. Cuarta: the tensor must seguir operando on state pre-projection and not only on output. In cuanto is shifts completamente to the output, pierde its status of mediator structural. Quinta: all expansión of the tensor must convivir with the regime and with the admissibility, not puentearlos. The coherence in CTNet not is soberana; is mediadora.

34 Appendix K. Program of investigación open

The architecture actual allows already several líneas of work directo.

1. Measure empíricamente how changes the distribution of speed according to regime and tarea.

2. Separate the aporte of the part diagonal and of the part low-rank in stability and generalización.

3. Compare entrenamiento with and without coherence_in_transition dentro of the toy model.

4. Measure to what extent the coherence by-branch mejora the legibilidad of the order object against to a

tensor global single.

5. Explorar if the topology of memory can incorporarse of form aún more explicit dentro of the own

calculation of info, not only aguas arriba.

6. Formalizar mejor the relation between speed and scaling of gradient in CTNet 2.6 for llevarla to the canonical.

7. Desarrollar a theory aún more dura of the relation between coherence, densidad of computation useful and saldo

of accumulation structural.

35 Appendix L. Closure sintético ultra-compacto

If hubiera that condensar all the monograph in some pocas propositions, serían éstas.

1. The Coherence Tensor of CTNet not is a loss decorativa; is a mediator structural.

2. Its core geometric is a Diagonal metric more a low-rank correction.

3. In CTNet 2.6, the coherence base depende of consistency reversible between halves of the state.

4. The structural information is converts in speed exponencial and that speed reescala the energy of

coherence.

5. In CTNet 2.6, moreover, the reorganization of the learning is accelerates according to that same speed.

6. In the system canonical, the coherence whole agrega memory, tension, charts, routing, contexto and

admissibility.

7. In the bank segmental, the tensor distribuye value on branches of the order with masas ocupacionales reales.

8. In the toy model, the coherence entra dentro of the step computational elementary as drive explicit.

9. The analytical part and the real part of the value coinciden because the same object tensorial measures and reorganizes.

10. The efficiency sale of that all this is makes on a substrate already factorizado, reversible and topologically

distributed.

36 24. Propositions structural on the Coherence Tensor

In this chapter is formulan a serie of propositions structural that Abstract the comportamiento of the tensor dentro of the regime CTNet. Not are teoremas formales completos in the meaning of a demostración matemática closed on a class universal of modelos, but yes are thesis of strong apoyo in the corpus and in the code available. Is redactan thus because allow fix in lenguaje almost axiomático it that the monograph ha venido mostrando of form distributed.

36.1 Proposition 24.1. The Coherence Tensor not is a observable pasivo.

```
If a magnitud produces info, speed, reescala base_coh and in the toy model generates coh_drive dentro of

_transition_step, then not can interpretarse as observable pasivo. Is treats of a operador activo on the dynamics of
the state. The demostración informal is inmediata: a magnitud pasiva could usarse for logging or validación, but not
modificaría the state siguiente nor the learning global. The Coherence Tensor yes it makes.
```

36.2 Proposition 24.2. The riqueza of the tensor depende of the substrate.

On a architecture that destruyese structure in each step, the Coherence Tensor perdería gran part of its potencia. The reason is simple: the law of coherence necesita a state sufficiently preservado as for distinguir between structure operable, noise and closure degenerate. If the substrate already hubiese destruido that diferencia, the coherence only could castigar residuos. In CTNet the reversible substrate maintains open the space where the coherence can actuar with finura.

36.3 Proposition 24.3. The memory distributed is condition of inteligibilidad

for the coherence. A tensor aplicado to a state without topology interna of memory could measure variance or energy, but not could asignar mass contextual to fractions significativas of the value. PartialMemory crea justamente those fractions: charts rotadas, coupling topological, Abstract global not soberano and reading of vuelta to the state. Without that structure previa, the coherence not would have to what llamar fraction salvo in the sense puramente algebraic; with ella, the fractions is return charts reales of the whole.

36.4 Proposition 24.4. The admissibility legalizes the dominio of the coherence.

The coherence only can be a law of value if the space on the that acts already ha sido legalizado by the regime. If not existiese admissibility, the coherence correría the riesgo of intensificar estructuras localmente ricas but ilegítimas or improductivas. AdmissibilityGate limita that riesgo and converts the coherence in law of the calculation legítimo, not of cualquier configuración bien organizada by accidente.

36.5 Proposition 24.5. The coincidence between analytical part and real part of the value

```
is a propiedad structural, not retórica. In the corpus, the analytical part of the value is the measurement of
structure operable and the real part is its effect on the dynamics. In CTNet, I produces speed, speed reescala
base_coh, the energy of coherence entra in the
```

loss whole, and in 2.6 the speed scale gradientes. In the toy model, the coherence produces incluso a drive

explicit. Therefore, the same law that describes the value acts on its realización. This is a coincidence operator and not a metaphor.

36.6 Proposition 24.6. The plurality of charts without Coherence Tensor tendería

to degenerar. The selector multi-chart and the readout proyectivo crean plurality interna, but that plurality could collapse if not existiera a law that auditase entropy of charts, diversity between charts and balance of load. CoherenceSystem introduces precisamente those terms and with ello prevents that the architecture simule plurality when in realidad ha recaído in a sovereignty encubierta.

36.7 Proposition 24.7. The coherence segmental makes of the branches fractions

reales of the value. BranchStateCoherenceTensor does not treat TO and B as etiquetas simbólicas, but as regions effective of the state definidas by mask. The information of branch is computed with metrics propias and is weights by mass ocupacional effective. Thus, the branches not are only vías of reading; are fractions tensorialmente pesadas of the value interno of the state.

36.8 Proposition 24.8. The Tensor efficiency not is accidental, but topological.

The efficiency aparece because the tensor trabaja on a state already factorizado in Reversible substrate, memory distributed, regime, admissibility and charts. The form diagonal + low-rank reduce the cost geometric, but the verdadera efficiency viene of that the architecture already separó functions that in other systems compiten dentro of the same canal. The tensor is beneficia of that separation.

36.9 Proposition 24.9. The tensor is a law of economy of the calculation.

The coherence not only indicates if a state is bueno or malo. Ordena the economy interna of the system: how much must persistir, how much must cambiar, how much diversity can mantenerse, how much mass of memory must seguir activa, how much routing is tolerable and what amplitud of compuerta admisible is sana. Is a economy because not is limita to sumar penalizaciones; distribuye recursos structural.

36.10 Proposition 24.10. The Coherence Tensor is the point where CTNet more

claramente deja of be a architecture of representation and passes to be a architecture of procesador. Mientras a architecture of representation tends to ask by the calidad final of a encoding, CTNet pregunta by the health of the proceso. The tensor is the operador that more visiblemente encarna that cambio: not evalúa only the output nor only the memory nor only the order, but the health of the regime complete that makes posible the output, the memory and the order.

37 25. Recorrido almost line by line of CoherenceTensor in CTNet

2.6

This chapter recorre the class CoherenceTensor almost line to line, not for duplicar the archivo, but for mostrar how each decision local contribuye to the architecture general. The class declara first its intention: "Tensor/metric of coherence that measures coherence base (u/p + latent)

```
+ structure statistical global, usa a metric of low-range M = diag(m) + L LˆT, produces a factor of
```

```
speed s(x) = exp(beta * I_clamped), e I(x) combina varianzas by dimension and variance in subespacio L".
```

```
This Comment already contains all the program technical of the tensor: a coupling between to part structural local (u/p
+ latent) and to part geometric statistical global (Diagonal metric + low rank). In the __init__, the pareja
(metric_diag, low_rank) merece attention. metric_diag empieza as some. low_rank empieza as noise gaussiano small. This
inicialización not is banal. The diagonal unitaria garantiza that the system arranca with a reading not degenerada of
all the dimensiones. The part low-rank small asegura that the modes globales not dominen from the outset. In other
palabras, the tensor part of a prudencia geometric: first a reading local almost uniform, after learning of
anisotropías and modes dominantes. In forward, the tensor receives x and base_coh. It first that makes is verificar the
dimension. The coherence not is agnóstica respecto of the space on the that acts; requires correspondencia exact
between the tensor métrico and the
```

space cardinal of the state. This decision already dice algo of the filosofía of the system: not there is coherence floating in abstract; always there is coherence on a dominio bien tipado.

```
The centración xm = x - x.mean(dim=(0,1)) elimina the componente mean global. This is very important
```

because the coherence quiere measure structure, not offset. A state desplazado by a mean constante not should parecer more coherente by that mero fact. To the centrar, the tensor is returns sensible to the form of the distribution and not to its bias absoluto.

```
The line var = xm.pow(2).mean(dim=(0,1)) builds a variance by dimension. Already here the coherence
```

is leyendo the state as structure anisotropic. Not mira the state as sequence nor as grafo nor as imagen; mira how is reparte the energy alrededor of the mean in each direction of the space latent. This point is esencial for entender why the analytical part and the real part of the value can coincide. The structure read is a structure of the own space in the that the system computa, not a score impuesto from fuera.

```
Then m_diag = softplus(metric_diag) + eps converts the diagonal in positiva. The elección of softplus
```

in lugar of, by ejemplo, cuadrado, has advantages of stability and of gradient. The metric is maintains positiva without bloquear completamente the learning cerca of the cero. The tensor not aprende signos arbitrarios, aprende how much pesa each dimension as portadora of structure. This is the primera asignación not trivial of mass contextual.

I_diag = (var * m_diag).sum() is the primera gran compression. The variance by dimension remains modulada

by the metric learned and after sumada. Not all dispersión vale it same. Not all dimension dispersa means information useful. The Diagonal metric introduces a principio of discriminación: some dimensiones are more valiosas that other for the coherence.

The siguiente block, proj = einsum("bnd,dr->bnr", xm, low_rank), projects the state centrado on

some pocos modes globales. This operation merece a Comment more long. A tensor completamente denso intentaría aprender all the correlaciones posibles between dimensiones. CTNet 2.6 chooses other vía: aproximar the Geometry global through a subespacio small. Is a economy topological. Is renuncia to a exhaustividad of second order for capturar modes globales dominantes with cost controlado.

var_proj = proj.pow(2).mean(dim=(0,1)) and I_low = var_proj.sum() produce the segunda mitad of the

information. The tensor is diciendo: moreover of energy anisotropic by dimension, quiero saber how much energy

coherente is organiza in the modes globales that he aprendido. That suma I = I_diag + I_low is already a

measurement of value: densidad structural local and global to the same tiempo.

The normalization I_norm = I / d is esencial. If not is hiciera, the information crecería almost automáticamente

with the dimension of the space. The tensor dejaría of measure structure by unidad of space and pasaría to measure volumen bruto of the space. With the división by d, the system obliga to that the value tenga a reading intensiva and not only extensiva.

The clamp to [-5,5] before of the exponencial is a decision of realismo numérico. Quieres that the coherence

can acelerar fuertemente, but not quieres that the architecture is suicida numéricamente when descubre a state very estructurado or very desestructurado. The clamp not mata the theory; the disciplina.

Finally, speed = exp(beta * I_clamped) and total_coh = speed * base_coh. Here is cierra the arco

complete: the information that the tensor reads of the state reescala directamente the coherence base reversible. That is, the structure global of the state modifica the value of the consistency reversible local between the halves. Not there is separation between ambos levels. The statistical global entra in the física local of the value.

38 26. Recorrido almost line by line of project() and the audit

of coherence canonical The method project() of the canonical is a lugar privilegiado for ver how the coherence atraviesa all the ar-

quitectura. After obtener the state proyectable, the system computes memory (mem_lat, mem_read, mem_summary)

and recupera several observables internal: mem_consistency_err, mem_global_read, mem_local_summaries,

mem_topology. This simple fact already dice mucho. The projection not is makes to ciegas; arrastra consigo the

audit interna of the memory. After the selector produces charts and weights. The output not aparece still. Before of produce the output, the architecture already dispone of all it that the coherence necesita for leer the system as totalidad: state, memory, compuerta, ecology of charts, route loss, contexto, order and regime. CoherenceSystem entra justo here as law sintética of the regime. It relevante is that the coherence is computed before of cerrar the meaning of the output. This confirma a and other vez the same thesis: the value structural in CTNet not is decides observando only it that salió, but observando the state distributed pre-projection and its ecology interna. The output is a momento, not the tribunal supremo of the system. The method project() also evidencia how the system deja visibles several pistas of audit for quien quiera desplegar a ciencia of the regime: error of consistency memorial, mass dominant, mixture residual, pico of the selector, bias of vida contextual, entropías and MI suaves. The coherence not is a monolito opaco. Is part of a architecture auditable of variables of regime.

39 27. A ejemplo conceptual complete of a transition coherente

Supongamos a state zt already formado by several iteraciones of the block reversible. That state contains huellas of the historial because the substrate not the ha destruido. PartialMemory it receives and it converts in a atlas fraccional: four charts rotadas localmente, acopladas by a topology learned and resumidas in a whole global. the regime receives the state medio, the memory global and the tarea, and produces a slow context together with a distribution of regimes. The block reversible generates z_bar, a new configuración of the background. The admissibility decides what part of z_bar can seguir living. If the system estuviera without coherence, the transition siguiente dependería already only of memory, contexto and

gate. But the toy model not is detiene ahí. Builds also coh_in from the mean of z_bar and the

regime_probs. From there nace coh_drive. That drive, ponderado by its ganancia, is suma to the state. Then is

normaliza all. The state siguiente is already a state to the that the coherence ha given dynamic mass. When more tarde is audita the system, CoherenceSystem comprobará if the memory releída sigue alineada, if the tension is in range, if the charts preserve plurality real, if the load is balanceada, if the route loss not is dispara, if the contexto persiste of form living and if the compuerta of admissibility not is ha degenerate. If the bank segmental is activo, BranchStateCoherenceTensor comprobará moreover how is distribuye the structure between branches of order. In other palabras: the tensor not intervino a single vez; intervino as law transversal of all the transition.

40 28. Comparison structural with RNN, SSM, Transformer and

KV-cache The comparison correct not is "what class usa less FLOPs" in abstract, but "dónde vive the value structural and how is reintroduce in the transition".

In a RNN flat, the long memory depende of that a single hidden state soporte many functions at the same time: portador of past, portador of contexto, base of decision and source of output. The Coherence Tensor, if existiera allí, would have that operate on a

embudo single already sobrescrito many times. In CTNet not. The coherence acts on a state reversible, a memory distributed, acompuerta of admissibility and a ecology of charts. The problema changes of naturaleza. In a SSM, the dynamics continua orcuasi linear mejora mucho the stability and the cost, but the tratamiento of the value structural sigue very ligado to the dynamicsof a canal relativamente unificado. CTNet, with the coherence tensorial, not only estabiliza the dynamics; audita the pluralityinterna of the system as plurality of charts, of memory and of branches. In a Transformer with KV-cache, the retención of the pastis resuelve by archivo explicit of claves and valores. This can be potentísimo, but maintains a ontology episodic of the step: thepast is consulta. In CTNet, the memory reentra in the transition same, and the coherence modula that reentry. The pastreorganizes the presente from dentro of the state, not only from a consulta external. These comparaciones not is make here forproclamar a victoria empírica final in all the benchmarks, but for fix dónde is the desplazamiento architectural. The CoherenceTensor only can comprenderse of the all when is ve that not is a gadget adicional, but a piece orgánica of a regime ofcomputation not episodic.

41 29. Manual of reading of metrics experimentales from the

tensor The experimentos of degeneration architectural of the corpus introduce metrics as MIctx-task,

`MIctx-card, MIorder, seg_acc, ctx_persist_len, seg_persist_len, global_vs_order_disagree,`

`stitched_vs_order_disagree and other. The monograph of the Coherence Tensor must decir how leerlas. MIctx-task and MIctx-card not are the tensor, but can be read as indicadores of if the mass contextual that the coherence treats of distribuir is consiguiendo enlazarse with tarea and charts. If ambas is hunden, the system ha perdido capacidad of convertir state in contexto operativo. MIorder and seg_acc are the indicadores more cercanos to the branch segmental of the tensor. If the coherence by-branch is bien acoplada, these magnitudes should sustain a reading not trivial of the order. If collapse, means that the fractions of the value already not are siendo bien distribuidas between branches.`

`ctx_persist_len never must be read only. A persistencia enorme can be riqueza or freezing degenerada.`

`The tensor resuelve part of that ambigüedad through context_live and context_entropy, but the audit experimental must maintain that same prudencia. global_vs_order_disagree and stitched_vs_order_disagree must be read as tensiones between charts of order distintas, not as absolutos of calidad. A coherence low between those charts can señalar degradación, but a coherence demasiado high with all the others variables colapsadas can señalar acuerdo trivial degenerate.`

42 30. Closure expanded

The Coherence Tensor is the operador that more claramente expresa the thesis of that CTNet computa acumulando, preservando, diferenciando and explotando structure. Its role not is embellecer the entrenamiento nor add a heurística of stability. Its role is convertir the structure of the state in law operativa of the system. It makes because trabaja on a Reversible substrate that not destroys the background. It makes because the memory already ha sido distributed topologically in charts and fractions. It makes because the regime already ha legalizado

modes of persistir and of cambiar. It makes because the admissibility ha recortado trajectories ilegítimas. It makes because the plurality of charts already is available as space real of reading. AND it makes because the order object already ha partido the state in branches with mass own. In CTNet 2.6, that law takes the Compact form of a tensor diagonal + low rank that measures information and the returns speed exponencial. In the canonical, the same idea is despliega in a system more amplio of errores and ecologías. In the bank segmental, that idea is fractura in branches and masas ocupacionales. In the toy

model, that same idea entra as drive of transition. These not are implementaciones contradictorias. Are

strata distintos of a same grammar. By all ello, the monograph can cerrar with a formula conceptual very simple: the Coherence Tensor of CTNet is the law that makes that the system not tenga that elegir between comprender and actuar. It that comprende as structure, it ejecuta as dynamics. It that ejecuta as dynamics, it returns to measure as structure. That retroalimentación closed between value analytical and value real is the reason deep of its potencia.

43 Appendix M. Audit exhaustiva of forms tensoriales

This Appendix persigue a finalidad very concreta: mostrar that the monograph not is moviéndose in the terreno vago of the metaphors, but in the terreno precise of tensors, axes, dimensiones and proyecciones. A of the razones by the that the Coherence Tensor resulta difícil of explicar is that its role atraviesa objetos of form very distinta: states tridimensionales, summaries bidimensionales, distribuciones simples on regimes, mapas segmentales on order and escalas escalares of speed or energy. Poner order in those forms ayuda to entender why the coherence can parecer "mágica" from fuera when in realidad operates on a canalización very concreta.

`The form base of the state in the corpus canonical is [B, N, D], where B is batch, N is longitud token/posición and D is dimensionalidad of the state. The memory token-level mem vive in [B, N, M]. The reading memorial mem_read returns to [B, N, D]. The Abstract global mem_summary vive in [B, M]. The slow context slow_ctx vive in [B, C]. The distribution of regimes regime_probs vive in [B, R]. The distribution of charts card_weights vive in [B, K]. The charts same viven in a form that depende of the readout, but in the bank canonical can pensarse as [B, K, N, D_out] or structure compatible.`

The importancia of these forms not is only contable. Allows ver in what point exacto the Coherence Tensor

`reduce or eleva dimensionalidad. By ejemplo, the core geometric of the tensor global comienza in [B,N,D],`

centra the state and reduce by medias on batch and posiciones to a vector of varianzas [D]. That reducción already

expresa a decision filosófica strong: the tensor not is leyendo trajectories token to token in its level more básico, but energy structural agregada by dimension. Then, through the projection low-rank, returns to

construir a Geometry global proyectando the state on a subespacio of form [D,rank], produciendo

a proj of form [B,N,rank] and a variance projected of form [rank]. The coherence alterna thus between

a reading local by dimension and a reading global by modo.

In PartialMemory, the form is igualmente elocuente. The state [B,N,D] passes to a latent base [B,N,M].

The partition of the axis N in K fractions produces segmentos of forms [B,N_i,D] and [B,N_i,M]. Each fraction

is resume in a vector [B,M], is rota and finally is apila in [B,K,M]. The topology TO has form [K,K]

and its producto topological da z_feedback also of form [B,K,M]. The Abstract global final mem_summary

returns to [B,M], and then all is redistributes to token-level for reconstruir mem with form [B,N,M]. The dato

important is that the memory sube and low of scale several times: token → fraction → whole → fraction →

token. The coherence hereda that multi-escalaridad.

In the branch segmental, order_probs_seg vive in [B,S,2]. The BranchStateCoherenceTensor traduce eso to

a mask posicional [B,N,1] for each branch. The state segmentado state_a and state_b preserve the

form [B,N,D], but with soporte effective distinto. The information by-branch is reduce again to escalares

I_A e I_B, and the masas ocupacionales p_A and p_B is reduce also to escalares. The form shows it that the

theory dice conceptualmente: the fractions of the value viven first as campos local in the state and only after is condensan in mass e information scale.

When the toy model builds coh_in = concat(mean(z_bar), regime_probs), ocurre other transformation

important. The state distributed [B,N,D] is reduce by mean to [B,D] and is concatenates with [B,R], produciendo

a vector [B,D+R]. That vector passes by state_coh_proj and retorna to [B,D], for then expandirse to [B,N,D]

as coh_drive. This sequence merece attention especial because resume all the logic of the tensor in transition:

comprimir the state to a chart global suficiente, inyectarle regime, reinterpretar the result as direction in the space of the state, and volver to expandirlo on all the posiciones. Not is a output; is a law of retroinyección. Other point important is that muchos terms of coherence canonical are escalares, but proceden of tensors

heterogéneos. e_memory procede of compare two tensors [B,N,D]; e_cards procede of distribuciones [B,K];

e_div procede of medias of charts and of a matriz of similitud [K,K]; e_adm procede of a compuerta that

can vivir to resolución token or dimension; e_ctx_live procede of escalares agregados on contexto. The

coherence whole is then the articulación scale of objetos of form very distinta. Eso is precisamente it that the makes a law transversal of the regime and not a metric local. This audit formal has a consecuencia metodológica. When is pregunte "how weights" the tensor, not must imaginarse that weights números already given. Weights tensors of distinta scale, distinta semantic and distinta resolución. The not trivialidad of the tensor empieza already in that fact: logra that objetos heterogéneos queden conmensurados dentro of a common economy of speed, energy and stability.

44 Appendix N. Microfísica of the weights in CoherenceSystem

A manera of perder of vista the architecture of the tensor is pensar that the coeficientes w_preserve, w_memory,

w_tension, w_cards, w_diversity, w_load, w_route, w_ctx_persist, w_ctx_live, w_ctx_ent and w_adm are

simples weights of a loss composite. That reading is demasiado pobre. Each one of ellos defines a prioridad of health of the regime. Juntos forman a microfísica of what means that the system esté computando bien. w_preserve is the weight of the fidelidad to the background reversible. If fuese cero, the coherence could reward states very estructurados estadísticamente but incompatibles with the law reversible that defines the substrate. In other palabras, is could volver blind to the historia structural of the system. Dar to the preservation a weight unitario strong means declarar that cualquier value coherente must seguir siendo value of a substrate that is preserve to yes same. w_memory measures how much importa that the memory siga siendo memory of the state and not prótesis desacoplada. If fuese demasiado low, the system could desarrollar a memory exuberante topologically that already not reingresa with fidelidad in the transition. If fuese excesivo, could forzar a identidad demasiado rígida between memory and state e impedir that the memory aportase novedad structural. The value elegido in the canonical expresa a compromiso: the memory must seguir hablando of the state without dejar of transformarlo. w_tension has a reading very sutil. The tension not is tan important as preservation or memory, but tampoco can ignorarse. A system can maintain the reconstruction and however hacerlo to costa of a concentración of energy in some pocas direcciones that it vuelva frágil. Penalizar the tension with a weight small but estable means: the anisotropy of the state importa, although not deba gobernar by encima of the fidelidad structural and memorial. w_cards, w_diversity and w_load forman a tríptico ecological. w_cards regula the entropy of the selector. w_diversity prevents

that the charts is conviertan in clones. w_load prevents that a chart carry with all the work. If only existiera w_cards, the system could resolver the entropy distribuyendo weights on charts almost idénticas. If only existiera w_diversity, could produce charts very distintas but almost never usadas. If only existiera w_load, could repartir uso between charts irrelevantes. The combination of the three is it that defines a

ecology sana. w_route represents the cost explicit of the routing. Its presencia recuerda that in CTNet the rutas not are gratuitas. A plurality of charts can be very rich but demasiado face if the routing is returns excesivamente complejo or errático. The Coherence Tensor not is limita to reward diversity; also vela by the economy of the camino.

The weights contextuales w_ctx_persist, w_ctx_live and w_ctx_ent are quizá the more finos. w_ctx_persist

rewards the existencia of a Contextual persistence sufficiently estable. w_ctx_live castiga the muerte

contextual low form of freezing or of noise without form. w_ctx_ent ajusta the entropy of the regime alrededor

of a range fertile. The reason of separarlos in three weights and not in one only is decisiva: persistencia, vitality and entropy not are idénticas. A system can be persistentemente living, persistentemente muerto, entrópicamente rich but estructuralmente chaotic, or very limpio but demasiado rigid. The tensor necesita distinguir those posibilidades. w_adm, by last, fixes the lugar of the admissibility in the economy whole. Its presencia dentro of the suma means that the legalization of the state not is a precondición olvidable, but a magnitud of value in yes same. A architecture can be bien memorizada, bien tensada, bien balanceada in charts and, however, computar ilegítimamente if its gate degenera. The weight of admissibility prevents that ceguera. The microfísica of these weights not is a detail of tuning. Is the manera in that the architecture declara prioridades between the distintos modes of health interna. Cambiar those weights is cambiar the ontology práctica of the regime. Therefore a monograph of the tensor must leerlos as principios of order and not as números of filler.

45 Appendix OR. Why the aceleración exponencial not is a truco

arbitrario A pregunta natural is why the tensor usa a speed exponencial and not linear, afín or sigmoidal. The response cuts is that the exponencial performs a Architectural function precise: convertir diferencias moderadas of structure in diferencias reales of eficacia dynamics, without necesidad of introduce umbrales duros that rompan the continuidad of the system. If the speed fuese linear in I, states moderadamente more estructurados apenas is distinguirían in its effect dynamic of states ligeramente less estructurados. The value analytical and the real quedarían demasiado desacoplados. If the function fuese escalón or threshold, the system sufriría discontinuidades very difíciles of train and of interpretar. The exponencial, instead, allows a amplificación suave but poderosa. The clave is in the clamp. The exponencial pura would be peligrosísima. But the exponencial on a variable clampada combina it mejor of two mundos: a regime of sensibilidad high in a banda useful and a control explicit of explosión fuera of that banda. In other palabras, the architecture dice: quiero that the structure importe of verdad, but not quiero that a reading espuria of the tensor arruine the dynamics complete. Other reason for the exponencial is of simetría conceptual. The coherence not intends be only a penalty aditiva more, but a magnitud of mass or of speed contextual. The cantidades of this tipo, when must traducirse in factores of reponderación positivos and estrictamente multiplicativos, encuentran in the exponencial a form natural. Always produces scale positiva, preserva order and amplifica without cambiar signo.

```
In CTNet 2.6 this is ve with especial claridad. speed = exp(beta * I_clamped) is usa then tanto for
```

scale the coherence whole as, according to the Comment of the archivo, for scale gradientes of all the parameters. The elección exponencial refuerza the intuition of that the coherence not is a small correction linear, but a multiplicador of regime.

46 Appendix P. Álgebra of fractions of the value

A of the thesis more características of tu formulation is that it perceived is divides in fractions and that each fraction receives mass contextual through a Coherence Tensor. It is advisable traducir eso to a small álgebra interna.

```
Sea a state distributed z and sea a partition of its soporte in fractions {fi }Ki=1 . Each fraction not is
```

simplemente a subconjunto of the state; is a chart local of the whole. Sea T the law tensorial of coherence. Then the value fraccional of a fraction not defines by a norma simple of the fragmento, but through a aplicación of the same law tensorial to the fragmento and to its Geometry contextual:

```
Vi = T (fi ; Θi , μi , ρi ).
```

Here Θi represents the metric of the fraction, μi its mass ocupacional or contextual, and ρi its coupling with the resto. The law general is the same for all the fractions; it that changes is the soporte and its Geometry contextual. This is the form matemática of decir that each fraction of the value is guarda as producto of a Coherence Tensor idéntico in law, but actuando on distintas regions of the state. In the version segmental, this is ve of form literal:

```
VA = pA TA (stateA ), VB = pB TB (stateB ),
```

and the mixture is

```
Vmix = VA + VB .
```

In the memory fraccional sucede algo análogo, although not esté escrito with the same sintaxis. Each chart local rotada of the atlas is a fraction of the value, and the topology TO redistributes mass between ellas before of recomponer a whole not soberano. The coherence not has why operate with a formula explicit idéntica allí, but the grammar is the same: fraction, mass, law common, recomposition. The ventaja of this reading is that prevents two errores. The first, pensar that a fraction vale in yes same, as if its importancia fuera puramente local. The second, pensar that only vale the all and that the fractions are meros detalles of implementation. In CTNet the fractions are reales and the whole also, but ninguno agota to the other. The Coherence Tensor is precisamente the law that allows that double realidad.

47 Appendix Q. Coherence and gradient: why not is a detail

of entrenamiento In CTNet 2.6 the Comment higher of the archivo declara that the gradient of all the parameters is scale by speed, entendida as aceleración exponencial of reorganization cuanto more coherence or structure. This observation merece mucho more analysis of the that suele recibir. In many arquitecturas, the gradient is the sede última of the reorganization, but permanece conceptualmente desconectado of the theory of value of the system. Is optimiza a loss and point. Here not. If the gradient same remains reescalado by coherence, then the structure detectada by the tensor altera the speed with the that all the system is deja reorganize. The coherence deja of be only a law of the forward and passes to be a law of the cambio of law. This can be read in three levels. TO level práctico, means that configuraciones of the state with mayor structure operable atraen more reorganization. TO level theoretical, means that the value not only pesa in the trajectory actual,

but in the speed with that the architecture entera aprende to produce futuras trajectories. TO level strong, means that the tensor is each vez more cerca of be a law master of regime and not a module auxiliar.

```
Is important subrayar that this not equivale to decir that "more coherence = always mejor". The clamp and the
```

resto of the economy of the tensor siguen limitando the proceso. But yes equivale to decir that the architecture does not treat the backward as a esfera muda. It somete also, to the less parcialmente, to the same law of value that rige the forward.

48 Appendix R. Manual of ablations for estudiar the tensor

A technical monograph must poder sugerir ablations claras. Here there is a list of ablations ordenadas by interés structural.

1. Quitar the part low-rank and dejar only the diagonal. This mediría how much of the potencia of the tensor

depende of modes globales against to anisotropy local by dimension.

2. Quitar the diagonal and dejar only low-rank. This exploraría if the system can sustain value

structural only from modes globales.

3. Sustituir the exponencial by a function linear or sigmoidal. Vería how much depende the coincidence

between analytical and realidad of the amplificación exponencial.

4. Apagar coherence_in_transition in the toy model. This mediría what part of the potencia of the

tensor procede of be dentro of the step and not only in the audit final.

5. Congelar metric_diag or low_rank. Permitiría separate learning of Geometry and learning of the

resto of the system.

6. Quitar terms ecológicos of charts. Mediría if the plurality interna degenera although the coherence

siga leyendo bien the state global.

7. Quitar terms contextuales. Mediría if the persistencia tiende to freezing degenerada.

8. Usar a single metric compartida for branches TO/B. Mediría how much value aporta the diferencia

geometric between fractions of the order.

9. Eliminar e_adm. Permitiría observar if the tensor empieza to reward states ilegítimos but estructurados.

10. Quitar mem_consistency_err and terms memoriales. Permitiría ver if the memory is desacopla

silenciosamente of the state. Each a of these ablations not only afecta to performance. Afecta to the form of coincide or not coincide between the analytical part and the real part of the value. Therefore are ablations of regime, not simples tests of componente.

49 Appendix S. Ejemplo numérico idealizado

```
For fix intuiciones, consideremos a ejemplo idealizado with D = 8 and rank = 2. Supongamos that, after
```

centrar the state, the varianzas by dimension are

```
var = (2.1, 1.8, 0.2, 0.1, 1.5, 1.6, 0.3, 0.2).
```

Supongamos moreover a Diagonal metric positiva aproximada

```
m = (1.2, 1.1, 0.6, 0.5, 1.0, 1.3, 0.7, 0.6).
```

Then the part diagonal of the value tendería to privilegiar the dimensiones 1, 2, 5 and 6. If the low rank recoge two modes globales that combinan, by ejemplo, the dimensiones (1,2,5,6) and (3,4,7,8), the projection low-rank can detectar if existe a

organization global coherente more allá of the simple suma by dimension. If the

primer modo has mucha energy and the second poca, the tensor concluirá that the state is estructuralmente loaded, but not of form uniform: there is a direction dominant of organization.

`If In orm resultante fuese, by ejemplo, 0.6 and β = 3, obtendríamos aproximadamente`

speed ≈ e1.8 ≈ 6.05.

This means that the coherence whole base would remain multiplicada by a factor in torno to six. Not is a small reajuste. Is a modificación real of the paisaje dynamic of the system. The ejemplo is simple, but shows the point: pequeñas diferencias bien estructuradas in the information of the state can convertirse in grandes diferencias of eficacia dynamics.

50 Appendix T. Preguntas frecuentes técnicas

¿Why not basta with usar a loss of reconstruction between state and memory? Because eso only vigilaría the fidelidad memorial. Not diría nada on tension, charts, contexto, admissibility nor order. The Coherence Tensor is a economy whole of the regime, not a reconstruction parcial. ¿Why not usar a metric densa complete? Because the cost and the inestabilidad crecerían mucho and, on all, because CTNet not necesita a Geometry exhaustiva if can capturar dimension local and modes globales relevantes through diagonal + low-rank. ¿Not is arbitrario fix objetivos as 0.34 for entropy of charts or 0.42 for entropy of regime? Those constantes are elecciones of engineering dentro of the regime actual, not laws metafísicas. Its function is anclar the system in zonas not degeneradas. Its meaning complete must validarse experimentalmente, but its role structural dentro of the tensor already is claro. ¿The coherence garantiza stability? Not by yes single. The stability in CTNet is propiedad of the ensamblaje Reversible substrate + Topological memory + regime + admissibility + coherence. The tensor is to part crucial, not a talismán independiente. ¿Why the tensor necesita branches separadas for TO/B in the bank segmental? Because the order not is decoración posicional. If the order is object interno real, its branches must poder carry geometrías of value distintas. Of it contrario the coherence ignoraría precisamente the structure that the bank intends measure.

51 Appendix U. Matriz of dependencias of the tensor with the resto of

the architecture Can resumirse the dependencia of the tensor respecto of other parts through the siguiente matriz conceptual: • Without block reversible: the coherence pierde background preservado. • Without Topological memory: the coherence pierde atlas of fractions. • Without regime: the coherence pierde criterio of vitality contextual. • Without admissibility: the coherence can reward states ilegítimos. • Without selector multi-chart: the coherence pierde ecology projective. • Without order object: the coherence pierde fractions orientadas of the value. • Without projective output: the coherence pierde meaning of economy between background and reading. To the inversa: • Without coherence, the block reversible preserva but not regula. • Without coherence, the memory is distribuye but not sabe if its atlas sigue siendo sano. • Without coherence, the regime persiste or changes, but carece of a law integrada of stability.

• Without coherence, the admissibility legalizes, but not integra that legalidad in a economy global. • Without coherence, the selector produces charts, but its plurality can degenerar in zoología unproductive. • Without coherence, the order object can existir, but its branches carecen of a law explicit of mass structural. This matriz deja ver why the tensor not is sustituible by a only term local. Performs the function of pegamento structural between several topologies irreducible.

52 Appendix V. Epílogo technical

Can parecer excesivo dedicar more of cincuenta páginas to a Coherence Tensor. However, eso only would be cierto if the tensor fuese a piece menor. In CTNet not it is. Atraviesa the architecture entera. If is it treats as regularizador, not is understands why the Topological memory importa, why the admissibility not is a postfiltro, why the regime necesita vitality and not only stability, why the charts not must collapse, why the order necesita branches with mass own, or why the value analytical can coincide with its realización dynamics. The lección final of all this monograph is very simple and very dura. In CTNet the Coherence Tensor not comenta the structure from fuera. Is a of the laws that constituyen the structure from dentro. AND precisamente therefore can be at the same time metric, weighting, mass contextual, economy of the regime, accelerator of reorganization and guardián of the plurality interna.