COMPLEXITY AND DECOMPOSABILITY OF RELATIONS

ABSTRACT

IS THERE A GENERAL DEFINITION OF STRUCTURE?

WHAT DOES KNOWING STRUCTURAL COMPLEXITY GIVE YOU?

OPEN QUESTIONS, CURRENT RESEARCH, WHERE LEADING

Martin Zwick
Systems Science Ph.D. Program
Portland State University
Portland OR 97202-0751

zwick@sysc.pdx.edu

http://www.sysc.pdx.edu/Faculty/Zwick

prepared for:

Session on Measures of Structural Complexity
New England Complex Systems Institute
Sept. 21-26, 1997
ABSTRACT

A discrete multivariate relation, defined set-theoretically, is a subset of a cartesian product of sets which specify the possible values for two or more variables. Where three or more variables are involved, the highest order relation, namely the relation between all of the variables, may or may not be decomposable without loss into sets of lower order relations which involve subsets of the variables. In a completely parallel manner, a relation defined information-theoretically, namely a joint probability distribution involving all the variables, may or may not be decomposed without loss into lower-order distributions involving subsets of the variables. Decomposability analysis, also called “reconstructability analysis,” is the specification of the losses suffered by all possible decompositions.

The decomposability of relations, defined either set- or information-theoretically, offers a fundamental approach to the idea of “complexity” and bears on all of the themes prominent in both the new and the old “sciences of complexity. Decomposability analysis gives precise meaning to the idea of “structure,” i.e., to the interrelationship between a whole and its parts, where these are conceived either statically or dynamically. It specifies the structuring and distribution and the amount of information needed to describe complex systems. It sheds specific light on chaotic versus non-chaotic dynamics in discrete dynamic systems. It provides a framework for characterizing the dual processes of integration and differentiation which govern the diachronics of self-organization.
IS THERE A GENERAL DEFINITION OF STRUCTURE?

**ANSWER:** Yes, a **structure** is a set of **relations**.

1. **Relation** = a *constraint* linking entities, e.g., variables.

2. Variables can be *nominal* ⇒ discipline-general; can be *dynamic*.

3. Constraint defined, e.g., (a) set- or (b) info.-theoretically, i.e.,
   (a) subset of cartesian product or (b) multivariate probability distribution.

4. **Projections** of relation define **lattice of relations** (LOR).

5. **Structure** = *cut* through LOR = **decomposition** of a relation.

6. **Complexity**, *(structure) = # degrees of freedom (info.-theor.)*
   = # parameters needed to specify it

7. Represents *topology*, not *strength*, of constraints.

8. **Resolution**-dependent; data-independent.

9. **Lattice of structures** (LOS) = all possible decompositions.

*Other definitions possible in this framework.*
RELATION:

\[ R = \{ (a_i, b_j, c_k, d_l) \} \subseteq A \otimes B \otimes C \otimes D \]  

SET-THEOR.

\[ R = \{ p(a_i, b_j, c_k, d_l) \} \]  

INFO.-THEOR.

PROJECTION: \[ R \equiv R_{ABCD} \Rightarrow R_{ABC} \]  

call \( R_X \) simply \( X \)

LATTICE OF RELATIONS (LOR)

STRUCTURE = CUT thru LOR, e.g., \( ABC:ABD \)

COMPLEXITY\(_1\) (info.-theor.) = DEGREES OF FREEDOM

\[
\text{df}( ABC : ABD ) = \text{df}(ABC) + \text{df}(ABD) - \text{df}(AB) \\
= N_{ABC} - 1 + N_{ABD} - 1 - N_{AB} + 1
\]
WHAT DOES KNOWING STRUCTURAL COMPLEXITY GIVE YOU?

Descri. length/incompressibility/randomness of constraint related to dimension. INTEGRALITY.

CONTEXT: RECONSTRUCTABILITY (log-linear) MODELING

1. MAX ENTROPY distrib. given CONSTRAINTS (model).
   
   e.g., max $H_{\text{Shannon}}(q_{ABCD})$ given $p_{ABC}$, $p_{ABD}$
   
   (Equivalent max $H_{\text{Hartley}}$ for set-theor. reconstruction)

2. Descend LOS to simplest lossless structure.

3. Further decomposition $\Rightarrow$ loss. (Some might be acceptable.)

USE: TRADE OFF LOSS vs. COMPLEXITY

   e.g., LR-$\chi^2 + df \Rightarrow p(\text{error});$ \quad like RISSANEN descr. length

EX. of PREDICTIVE USE:

df (SIMPLEST LOSSLESS structure ), \quad LOSSES ( df )

improves CHAOS prediction in Elem. Cellular Automata.
PREDICTING CHAOS IN ELEM. CELLULAR AUTOMATA
(REDUCTION OF ATTRACTOR UNCERTAINTY)

q = rule attribute

| q          | H( a | q) | % ΔH | Pred.Power |
|------------|-------|------|-----------|
| -          | .679  |      |           |
| Walker-Ashby, Langton | λ   | .600 | 11.6      | .044       |
| COMPLEXITY | σ    | .553 | 18.6      | .069       |
| LOSS SPECTRUM | τ   | .263 | 61.3      | .102       |
| Wuensche   | Z    | .458 | 32.6      | .114       |

H (λ | τ ) = H (σ | τ ) = H (Z | τ ) = 0

CONCLUSIONS

τ vector is the BEST OVERALL (% ΔH) PREDICTOR

τ SUBSUMES ALL OTHER MEASURES

EVEN THOUGH:

SIMPLE LAWS ⇒ COMPLEX DYNAMICS

STILL:

MORE COMPLEX NON-DECOMPOSABLE LAWS ⇒

MORE COMPLEX CHAOTIC DYNAMICS
OPEN QUESTIONS, CURRENT RESEARCH, WHERE LEADING

THEORY/METHOD

1. Generalization of
   • complexity$_1$ to set-theoretic relations
   • reconstructability to fuzzy-distributions, non-max-H criteria

2. Improved algorithms for
   • searches through big lattices
   • optimal binning of quant. variables
   • optimal complexity/loss tradeoff

3. COMPLEXITY$_2$ : maximized between top and bottom of LOS
   • topological complexity
   • #structures(df)
   • sensitivity of complexity/loss tradeoff

SOME APPLICATIONS (reconstructability, complexity measure)

4. Applications of reconstructability to DATA-MINING.

5. Meta-dynamics of differentiation/integration (wholes $\leftrightarrow$ parts).

6. Extend CA work
   • more complex CAs
   • boolean nets
   • relate to continuous systems
SELECTED REFERENCES


International Journal for General Systems.