Holographic treatment of noncommutative actions and a forgotten algorithm of MPS.

Gérard H. E. Duchamp♣
Collaborators : Jean-Léo Léonard♡, Marco Patriarca♢, Elsa Heinsalu♢.

♣ Galileo Institute (Paris XIII) & Institut Henri Poincaré (next door=ND), ♡ Université Paris-Sorbonne, STIH (Paris IV, ND) , ♢ National Institute of Chemical Physics and Biophysics, Tallinn, Harjumaa, Estonia.

Structural Complexity in Natural Language(s),
Maison de la Recherche (Université de Paris 3),
30-31 mai 2016, Paris.

Plan

Figure : The 4th International Conference on Complex Systems and Applications.

- **Salikoko S. Mufwene (University of Chicago)**

Title : What Makes Language Evolution So Complex and Difficult to Model Accurately?
Scheduled : Monday June 23, 2014, 16:00-17:30
Abstract: Like any other kind of evolutionary process, language evolution is unplanned. It is the outcome of how adaptively individuals interacting with each other both reproduce and innovate ways of communicating information successfully, a process marked by successive changes between time t1 and tn (t1, t2, t3... tn). While it is convenient and relatively easier to describe changing norms or communal patterns within a population, it is more difficult to account for the actuation of the changes. In the latter case, there are several ecological factors that bear on the evolutionary process, including: 1) the error-prone process of spreading innovations by learning and instantiating them (unlike in the inheritance of genetic traits); 2) variation in innovations, which generates competition between/among variants (during the convergence towards communal norms); 3) the periodization following from differences between the times of particular innovations (and their consequences); and 4) population structure (which constrains how innovations spread and how far). Perhaps more difficult is the fact that evolutionary theories have focused on documenting population-level outcomes of behaviors that are actually specific to and centered around individual agents, none of whom is aware of the full extant system at the time they innovate or borrow an innovation (and thereby contribute to spreading it). Often, they are not even aware of alternatives that already exist or are being innovated by others, though not exactly in the same way, at the time of their behavior. This is when the distinction between internally- and externally-motivated change (articulated inaccurately in traditional historical linguistics!) becomes significant to whoever in concerned with the actuation of change. Self-organization is of course part of the big picture, to the extent that it explains (or does it?) how everything comes together in the emergent system, but we do face the challenge of explaining the nature of the complexity inherent in language evolution. I submit an account that applies not only to structural changes but also to language vitality.

- **Léo Léonard (IUF & Paris 3, UMR 7018) & Vittorio dell'Aquila (CELE, Milan & Vaasa)**

Avec la collaboration da Antonella Gaillard-Corvaglia (Postdoc, Inalco)
Title : Algorithmic complexity applied to geolinguistic networks
Scheduled : Monday June 23, 2014, 17:30-18:00
Abstract: Dialectology has long been considered as a marginal field in linguistics, mainly concerned with the recollection of empirical facts, with low theoretical expectations. Nevertheless, thanks to quantification of dialectical data (dialectometry, see Goebl, 1981, 1982, 2002)), geolinguistics in particular turns out to be one of the most promising horizon for Complexity Theory (CT) - as much as CT opens new horizons for dialectology. We'll apply various methods to a Mazatec database from Paul L. Kirk (1966), providing 9000 tokens (750 cognates x 12 locolects): patristic distances (cladistics, see Hennig, 1950, 1966), Levenshtein algorithm (Beijering & al. 2008, Bolognesi & Heeringa, 2002) and dialect intelligibility tests (Kirk, 1970), in order to show multiplicity of prospects on a geolinguistic space. Mazatec as a case study for testing algorithmic complexity has been chosen on several grounds: i) it once provided the empirical base for a landmark study by Sarah Gudschinsky (1958) on the reconstruction of dialect diversification process (1958), ii) Kirk's data, with less than 10 000 tokens is easier to process than bigger data available on European languages, iii) we have thoroughly checked and revisited Kirk's data through fieldwork within the framework of an empirical research project (IUF, MAmP, 2009-14, see Léonard & al. 2012), iv) phonology and grammar of Mazatec dialects have been formalized within the same project (with declarative phonology and Paradigm Function Morphology). Conditions for a survey of algorithmic complexity are therefore met, allowing a multiplex modeling of Mazatec geolinguistics.

- **Stefan Balev (ISCN, Le Havre), Gérard H.E. Duchamp (LIPN, Paris 13) & Jean Léo Léonard (IUF & Paris 3)**

Title : Visualizing and Revisiting Dialect Intelligibility Networks: Mazatec as a case study
Scheduled : Monday June 23, 2014, 18:00-18:30
Abstract: Dialect Intelligibility Testing (DIT) has long been the focus of attention in Native American linguistics (Hockett, 1958: 321-30), as an alternative standpoint to taxonomies based on comparative or quantitative linguistics. It has contributed to enhance the heuristic value of continuous chain models, over discontinuous tree-like models in dialectology. DIT also highlights epigenetic trends over phylogenetic and ontogenetic assessments on dialect variation. In this talk, we'll revisit Kirk's and Casad's data on Mazatec mutual intelligibility patterns (Kirk, 1970; Casad, 1974: 46-51, 167-79) with visualizing tools such as GraphStream (http://graphstream-project.org/), and we'll question relevant thresholds in Mutual Intelligibility Networks (MIN). We'll suggest a finer-grained grid than the one initially used by Kirk & Casad, in the seventies, and we'll show how previously unobserved communal clusters may emerge from algorithmic complexity, out of raw MIN data, through visualization devices now currently used in processing complex systems.

- **Marco Patriarca (National Institute of Chemical Physics and Biophysics, Tallinn, Estonia)**

Title : Models of language competition
Scheduled : Tuesday June 24, 2014, 14:00-15:00
Abstract: This contribution presents an overview of language competition models that have been introduced and studied for many years by now in different fields such as mathematics, physics, and linguistics. The overview covers some relevant models which resemble the classical mathematical models of competition between biological species. A first aspect that we try to clarify is the statistical interpretation of the models and their corresponding meaning, in particular concerning the role of bilinguals in the competition between two languages. We begin with the Baggs-Freedman models, proceed through the Abrams-Strogatz and Minett-Wang models, and then turn to more recent models. Also the spatial side of language spreading is considered, by illustrating various ways in which the spreading of language (features) through space and time has been modeled so far in order to take into account geographical, cultural, and social factors.

- **Adam Lipowski (Faculty of Physics, Adam Mickiewicz University, Poznan, Poland)**

Title : Dynamics of Naming Games: Why is it slow and how to make it faster?
Scheduled : Tuesday June 24, 2014, 15:00-15:30
Abstract: Naming Game is an important model of agreement dynamics. It might be used to describe emergence of a common vocabulary but it was also used to describe an opinion formation in a large scale sensor network, simple grammar or leader formation mechanism. Since the time to reach consensus is an important characteristics of Naming Game, its dynamics has been intensively studied. We show that some previous results on the Naming Game needs to be modified. It turns out that due to the formation of stripe structures dynamics of this model is much slower than expected. We also show that one can construct a Naming Game that does not suffer from the formation of such structures and has a much faster dynamics. We suggest

Figure : A fragment of the abstracts.

MPS or Marco as we used to call him

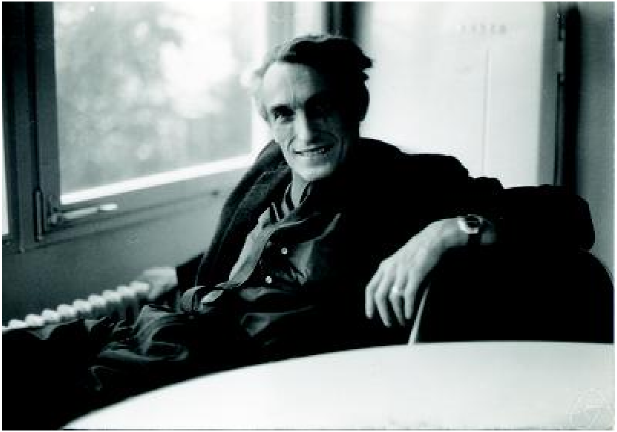
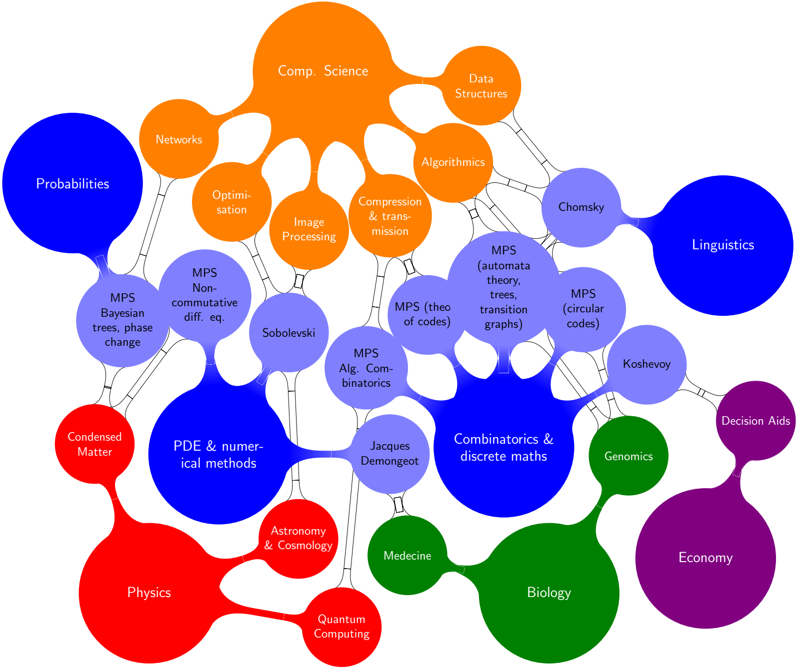


Figure : Marcel-Paul Schützenberger at Oberwolfach (1973)[1]

[1]Contrary to 1972 (Wikipedia)


Comp. Science
Data Structures
Networks
Optimi- sation
Image Processing
Compression & trans- mission
Algorithmics
Chomsky
Linguistics
Probabilities
MPS Bayesian trees, phase change
MPS Non- commutative diff. eq.
Sobolevski
MPS (theo of codes)
MPS (automata theory, trees, transition graphs)
MPS (circular codes)
Koshevoy
MPS Alg. Com- binatorics
Jacques Demongeot
Condensed Matter
PDE & numer- ical methods
Combinatorics & discrete maths
Decision Aids
Genomics
Astronomy & Cosmology
Medecine
Physics
Quantum Computing
Biology
Economy

About MPS (or Marco)

Then, MPS was a master in many things, in particular within the domain he created[2] : automata theory, transition systems, theory of codes, varieties of languages all domains which eventually revealed to be connected to representation theory, paths in categories, Hopf algebras, quantum groups and modern physics. He was regarded as an exceptionally creative combinatorialist[3], and the master of dynamic structures i.e. rigid structures with

- transformations (global, points)
- transitions (local, edges)
- evolution (global, points and edges)

And this in various domains and the two flavours : discrete (automata, trees) and continuous (non-linear automatics). Today, we will not touch the continuous realm as the *forgotten algorithm* is simpler in its discrete version.

[2] Theoretical Computer Science

[3] I.M. Gelfand, "who is regarded to be a prominent mathematician of the 20th century" said "Schützenberger is the best combinatorist of the world".

NC-MPS paper/1

THE ALGEBRAIC THEORY OF CONTEXT-FREE LANGUAGES*

N. CHOMSKY
Massachusetts Institute of Technology

AND

M. P. SCHÜTZENBERGER
Harvard University

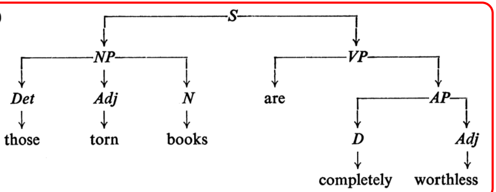
1. Linguistic Motivation

We will be concerned here with several classes of sentence-generating devices that are closely related, in various ways, to the grammars of both natural languages and artificial languages of various kinds. By a *language* we will mean simply a set of strings in some finite set V of symbols called the *vocabulary* of the language. By a *grammar* we mean a set of rules that give a recursive enumeration of the strings belonging to the language. We will say that the grammar *generates* these strings. (Thinking of natural languages, we would call the generated strings *sentences*; in algebraic

Figure : An algebraic model which can be useful for natural as well as for artificial languages.

ing subclassification. This information can be represented by a diagram such as (1):

(1)



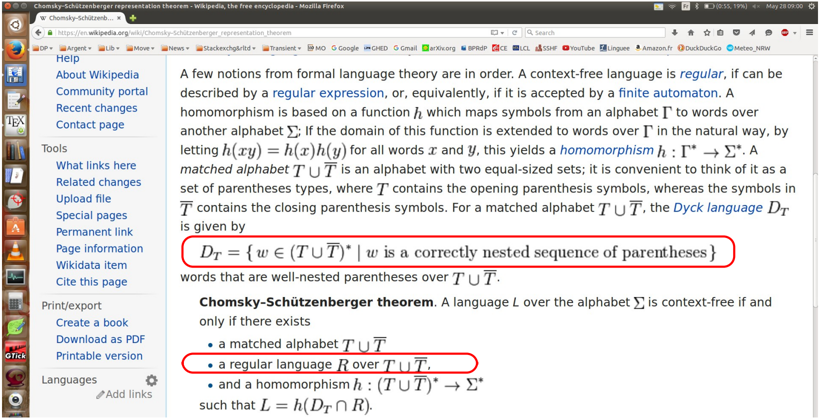
or, equivalently, by a labelled bracketing of the string, as in (2):

(2) $[_S\ [_{NP}\ [_{Det}\ \text{those}][_{Adj}\ \text{torn}][_N\ \text{books}]][_{VP}\ \text{are}\ [_{AP}\ [_D\ \text{completely}]\ [_{Adj}\ \text{worthless}]\ ]\ ]$.

A major concern of the general theory of natural languages is to define the class of possible strings (by fixing a universal phonetic alphabet); the class of possible grammars; the class of possible structural descriptions;

(Marked) trees are equivalent to words with brackets and, in turn, with words on an extended alphabet.

$$[_S \quad [_{NP}[_{Det}\ \text{those}\ ]_{Det}[_{Adj}\ \text{torn}\ ]_{Adj}[_N\ \text{books}\ ]_N]_{NP} \quad [_{VP}\ \text{are}\ [_{AP}[_D\ \text{completely}\ ]_D\ \text{worthless}\ ]_{AP}]_{VP} \quad]_S \qquad (1)$$



A few notions from formal language theory are in order. A context-free language is *regular*, if can be described by a regular expression, or, equivalently, if it is accepted by a finite automaton. A homomorphism is based on a function h which maps symbols from an alphabet Γ to words over another alphabet Σ; If the domain of this function is extended to words over Γ in the natural way, by letting $h(xy) = h(x)h(y)$ for all words x and y, this yields a *homomorphism* $h : \Gamma^* \to \Sigma^*$. A *matched alphabet* $T \cup \overline{T}$ is an alphabet with two equal-sized sets; it is convenient to think of it as a set of parentheses types, where T contains the opening parenthesis symbols, whereas the symbols in $\overline{T}$ contains the closing parenthesis symbols. For a matched alphabet $T \cup \overline{T}$, the *Dyck language* D_T is given by

$$\boxed{D_T = \{ w \in (T \cup \overline{T})^* \mid w \text{ is a correctly nested sequence of parentheses} \}}$$

words that are well-nested parentheses over $T \cup \overline{T}$.

Chomsky–Schützenberger theorem. A language *L* over the alphabet Σ is context-free if and only if there exists

- a matched alphabet $T \cup \overline{T}$
- $\boxed{\text{a regular language } R \text{ over } T \cup \overline{T},}$
- and a homomorphism $h : (T \cup \overline{T})^* \to \Sigma^*$

such that $L = h(D_T \cap R)$.

Figure : The two ingredients of CST : well-balanced (correctly nested) words and regular expressions. Dyck language D_T can be visualized by means of Dyck paths.

Nested structure

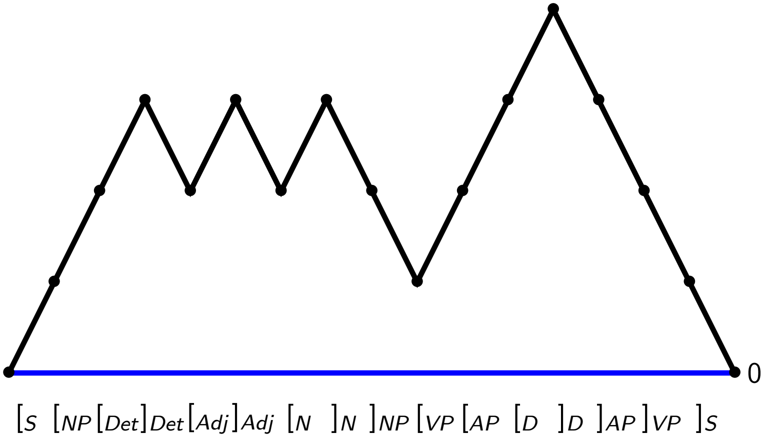


Figure : Bracket structure of Chomsky-Schützenberger's example

$[_S[_{NP}[_{Det}$ those $]_{Det}[_{Adj}$ torn $]_{Adj}[_N$ books $]_N]_{NP}[_{VP}$ are $[_{AP}[_D$ completely $]_D$ worthless $]_{AP}]_{VP}]_S$

The equational grammar of these correctly nested words is $D = \left(\sum_{b\in B} bD\bar{b}\right)D + \varepsilon$ (where B is the subalphabet of brackets).

Creation-Annihilation-processes/1

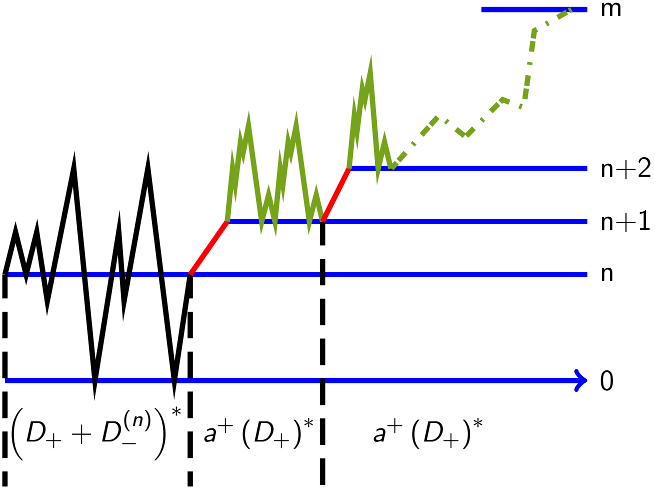
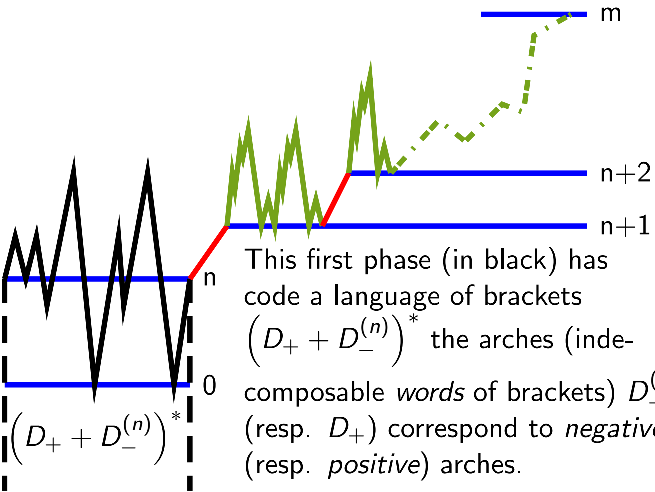


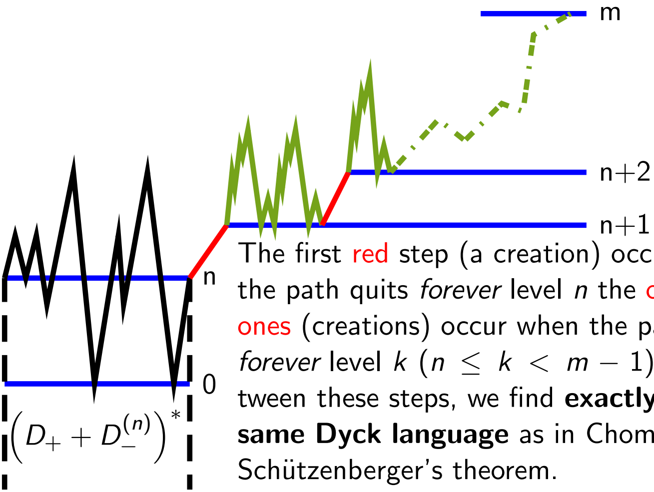
Figure : As a remarkable surprise, one finds these very two ingredients in combinatorial (quantum) physics (change of level from n to m).

Creation-Annihilation-processes/2



This first phase (in black) has code a language of brackets $\left(D_+ + D_-^{(n)}\right)^*$ the arches (indecomposable *words* of brackets) $D_-^{(n)}$ (resp. D_+) correspond to *negative* (resp. *positive*) arches.

Creation-Annihilation-processes/2



The first red step (a creation) occurs when the path quits *forever* level n the other ones (creations) occur when the path quits *forever* level k ($n \le k < m-1$). Between these steps, we find **exactly the same Dyck language** as in Chomsky-Schützenberger's theorem.

Regular languages and finite state machines

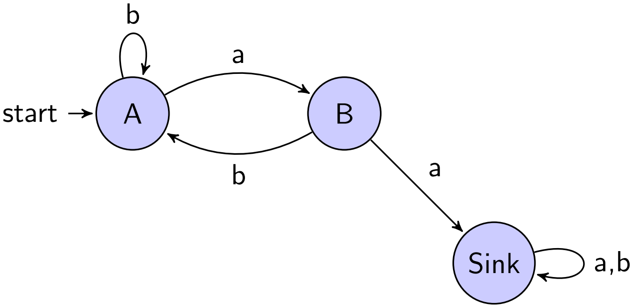


Figure : A simple finite-state automaton. Language recognized at state (A) is $(b + ab)^*$ i.e. words which factor into the blocks $\{b, ab\}$ (this is a "flower automaton"), the "Sink" is a special non-return state and the language recognized at (B) is $(b + ab)^* a$. The language recognized by $\{A, B\}$ is the set of words which do not contain the factor a^2. One has then $(a + b)^* - (a + b)^* a^2 (a + b)^* = (b + ab)^* (a + \varepsilon)$.

A simple transition system : weighted graphs

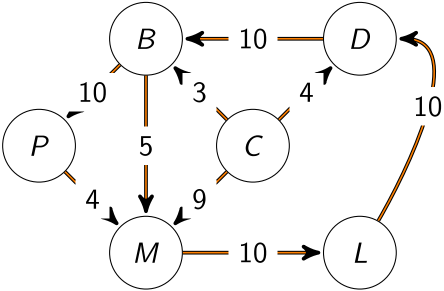


Figure : Directed graph weighted by numbers which can be lengths,time (durations),costs,fuel consumption, probabilities (with conditions). This graph is equivalent to a square matrix.

Holography begins : Layer automata

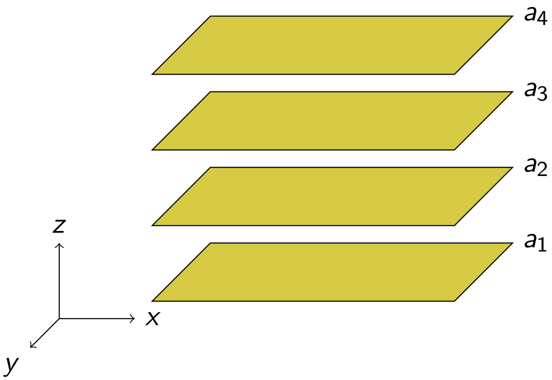


Figure : A four letter multiplicity automaton on the alphabet $\{a_1, a_2, a_3, a_4\}$: each layer contains a transition structure (weighted graph) on the same graph.

Multiplicity automata

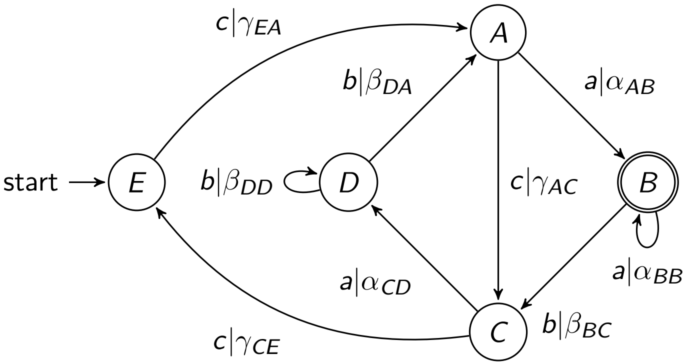


Figure : A multiplicity three-letter automaton. Here, the set of states is $\{A, B, C, D, E\}$, This automaton is equivalent to three layers of weighted graphs (with initial and final states). This automaton has three transfer (square 7x7) matrices (one per letter). As we know that weights multiply in series and add in parallel, we replace the element $a|\alpha_{AB}$ by $\alpha_{AB}a$ and get **one** (holographic) transfer matrix.

Words and paths/1

Let us take the simplest non-trivial example i.e. a two-state automaton with weights (labels multiply along a path in series and add along sets of paths with common source and target).

Remark : *This is theoretically sufficient as one can encapsulate many structures within the labels a_{ij} (as rectangular stochastic matrices or control automata).*

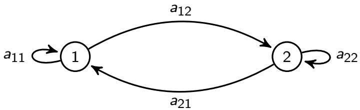


Figure : The generic transition structure over 2 states.

The transfer matrix is

$$T = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Words and paths/2

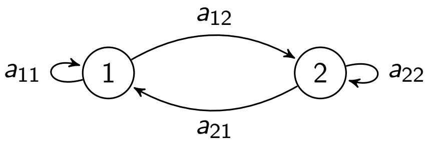
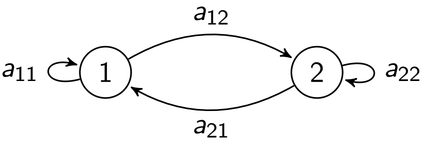


Figure : The generic transition structure over 2 states.

If you compute its powers, you get the magic formula (*ev* stands for "evaluation")

$$
\begin{aligned}
T^2 &= \begin{pmatrix} a_{11}a_{11} + a_{21}a_{12} & a_{11}a_{12} + a_{12}a_{22} \\ a_{21}a_{11} + a_{22}a_{21} & a_{21}a_{12} + a_{22}a_{22} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{|p|=2,p:1\to 1} ev(p) & \sum_{|p|=2,p:1\to 2} ev(p) \\ \sum_{|p|=2,p:2\to 1} ev(p) & \sum_{|p|=2,p:2\to 2} ev(p) \end{pmatrix} \\
\cdots\ T^m &= \begin{pmatrix} \sum_{|p|=m,ev(p):1\to 1} ev(p) & \sum_{|p|=m,p:1\to 2} ev(p) \\ \sum_{|p|=m,p:2\to 1} ev(p) & \sum_{|p|=m,p:2\to 2} ev(p) \end{pmatrix}
\end{aligned}
$$

Paths and loops lemma



Taking the coefficients as formal, one can prove that you can evaluate **all** the paths by the formula

$$
\begin{aligned}
T^* &= I + T + T^2 + \cdots T^k \cdots \\
&= \begin{pmatrix} (a_{11} + a_{12} a_{22}^* a_{21})^* & (a_{11} + a_{12} a_{22}^* a_{21})^* \, a_{12} \, a_{22}^* \\ (a_{22} + a_{21} a_{11}^* a_{12})^* \, a_{21} \, a_{11}^* & (a_{22} + a_{21} a_{11}^* a_{12})^* \end{pmatrix}
\end{aligned}
$$

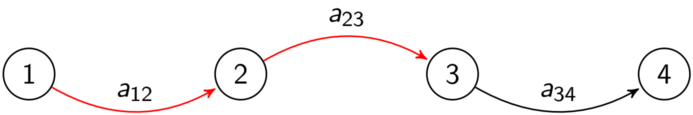
An this holds for all size of automata using the block-product.

What is required to multiply matrices (diagrammatic proofs)

So, if we want to be creative, we must question the limits of this model in order not to confine to conventional weights (probabilities, lengths etc.). The question is
Q) If the weights multiply in series and add in parallel what are the rules they must follow ?

What is required to multiply matrices (diagrammatic proofs)/products

Q) If the weights multiply in series and add in parallel what are the rules they must follow ?



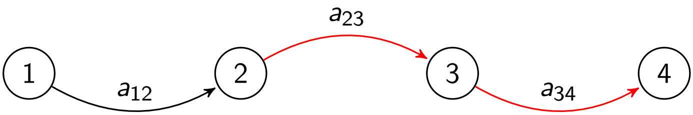


Figure : Associativity of products : Evaluating two ways the path $1 \to 4$ (products) yields $(a_{12}a_{23})a_{34} = a_{12}a_{23}a_{34} = a_{12}(a_{23}a_{34})$

What is required to multiply matrices (diagrammatic proofs)/sums

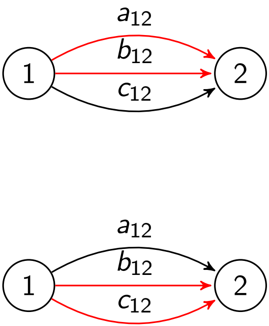
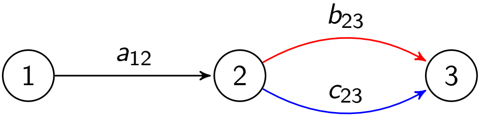


Figure : Associativity (and commutativity) of the sum : Evaluating two ways the paths $1 \to 2$ (products) yields $(a_{12} + b_{12}) + c_{12} = a_{12} + b_{12} + c_{12} = a_{12} + (b_{12} + c_{12})$ There is also a zero (neutral, disconnection) and the sum is commutative (the order of computation is not relevant).

What is required to multiply matrices (diagrammatic proofs)/distributivities



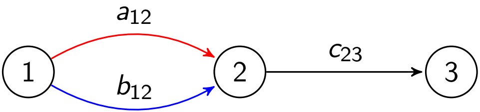
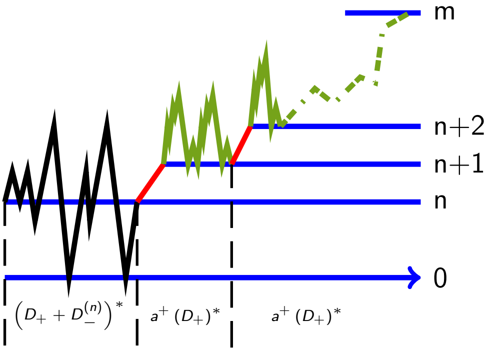


Figure : Distributivity of the product over the sum : Evaluating the set of paths $1 \to 3$, one gets (upper part) $a_{12}(b_{23} + c_{23}) = a_{12}b_{23} + a_{12}c_{23}$ and (lower part) $(a_{12} + b_{12})c_{23} = a_{12}c_{23} + b_{12}c_{23}$

Conclusion

A structure $(K, \oplus, \otimes)$ which fulfils these identities is called a *semiring*. Many of them were created for mathematics and for computer science classical rings of numbers and matrices, (max,$+$), (min,$+$). Moreover, one can **freely taylor new ones** as the one of shortest paths with memory. This last semiring (shortest path with memory) is the crucial part of an application of the "forgotten algorithm" (subject of a master thesis).
More generally, the formula of powers of the (holographic) transfer matrix allows to use the powerful (classic) algorithms for exponentiation and this paves the way towards a new non-commutative and very efficient (holographic) treatment of several layers of data.



Thank you for your attention !