One-dimensional topological quantum field theories with zero-dimensional defects and finite state automata

Mee Seong Im
United States Naval Academy
Annapolis, MD

Facets of Noncommutative Geometry:
Dedicated to the influence and memory of Tom Nevins
Department of Mathematics
University of Illinois, Urbana, IL

Joint with Mikhail Khovanov, Columbia University, New York, NY
arXiv:2202.13398

The night before our graduation at the University of Illinois.

Topological quantum field theory and geometry.

Tom worked, among many things, on moduli spaces, on holomorphic symplectic manifolds, on Springer resolutions and on categories of coherent sheaves on them, and on $\mathcal{D}$-modules. These structures play a fundamental role in TQFTs in dimensions 3 and 4.

In category $\mathcal{O}$, one assigns functors between blocks of category $\mathcal{O}$ to planar graphs with boundary. Then foams (cobordisms between planar graphs) induce natural transformations between these functors. Suitable cones of natural transformations between the functors lead to complexes for braids and tangles and to link homology. These categories are very close to $\mathcal{D}$-module categories studied by Tom. Category $\mathcal{O}$ can be rewritten as a category of $\mathcal{D}$-modules on a flag variety.

So $\mathcal{D}$-modules on flag varieties describe category $\mathcal{O}$, and the latter appears in the construction of link homology theories that categorify Jones and HOMFLY polynomials (Bernstein–I.Frenkel–Khovanov, Sussan).

Topological quantum field theory and geometry.

A similar structure is essential in the (derived) category of coherent sheaves on quiver varieties and Springer resolutions. These structures are again close to Tom's interests. Quiver varieties are holomorphic symplectic varieties, and Tom has been studying holomorphic symplectic varieties in great detail.

Categories of sheaves on (compact) holomorphic symplectic manifolds give rise to a Rozansky–Witten 3D TQFTs.

Derived categories of coherent sheaves on Springer resolutions and other quiver varieties appear in link homology theories (S.Cautis–J.Kamnitzer, R.Anno, V.Nandakumar, M.Aganagic).

Topological quantum field theory and geometry.

To relate to my recent work, I'll downsize from dimensions 3 and 4 to dimension 1 but add defects and change to a semifield from a field. Amazingly, that still gives a nontrivial structure, and the one heavily studied in computer science, that of regular languages and automata.

We're able to extend a regular language to a 1-dimensional topological theory with defects. To do that we need a "regularizer", a circular language. The resulting extension is a rigid (can U-turn morphisms and objects) symmetric monoidal category. It's a $\mathbb{B}$-valued TQFT (a stronger condition on state spaces) if a suitable state space is a *projective* semimodule (distributive lattice), and then there's a canonical circular language to build a TQFT from a regular language.

$\mathbb{B} = \{0, 1 : 1 + 1 = 1\}$ Boolean semiring.

Σ: alphabet (a finite set of letters). Σ^*: free monoid on the letters Σ.

A language L is a subset of Σ^*.

Example: alphabet $\Sigma = \{a, b\}$. Words $aaa, ababbba, bbaaab$, etc. Empty word $\varnothing$ is the unit element of the free monoid Σ^*.

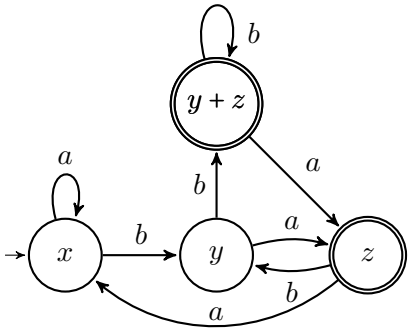
A language is *regular* if it's recognized by a *finite state automaton*, that is, by a machine with finitely many states and no memory.

Equivalently, a regular language is one given by a regular expression.

Example: $L = (a + b)^* b(a + b)$ consists of all words whose 2nd to last letter is b.

FSA (Finite State Automaton): words in letters Σ are inputs; finitely many states Q and transitions between the states $\Sigma \times Q \to Q$ according to the letters read. Has initial (starting) state q_{in} and terminating (accepting) states Q_t.

Example of FSA:



Language $L = (a+b)^*b(a+b)$.
Second from last letter is b. Four states.
Initial state given by the empty word $q_{\text{in}} = x$.
Accepting states $Q_t = \{z, y+z\}$.
The states z and $y+z$ correspond to the words $(a+b)^*ba$ and $(a+b)^*bb$, respectively.
Notation $y+z$ comes from relation to $\mathbb{B}$-modules.

A word can be viewed as an interval with dots (defects) labelled by letters in Σ. Reading a sequence of labels along an oriented interval gives a word $\omega = a_1 a_2 \cdots a_n \in \Sigma^*$.

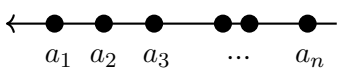
Idea: Interpret language L as a Boolean-valued function on words encoded as sequences of defects on an oriented interval.

Evaluation $\alpha_I : \Sigma^* \longrightarrow \mathbb{B}$ of decorated intervals is the same as an interval language L_I

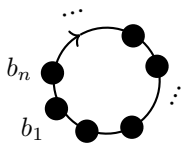
$$\omega \in L_I \Leftrightarrow \alpha_I(\omega) = 1 \ .$$

Add a circular language $L_\circ$ (for words on a circle $\omega_1\omega_2 \in L_\circ \Leftrightarrow \omega_2\omega_1 \in L_\circ$).

With pair $L = (L_I, L_\circ)$, associate a $\mathbb{B}$-valued multiplicative evaluation α of decorated 1-manifolds (defects labelled by letters in Σ).



$\omega = a_1 a_2 \cdots a_n$



$$\begin{aligned} \alpha : \text{closed } 1\text{-dimensional manifolds} &\longrightarrow \mathbb{B} \quad \text{which satisfies} \\ \alpha(M_1 \sqcup M_2) &= \alpha(M_1)\alpha(M_2), \\ \alpha(\varnothing_1) &= 1 \quad \text{since } m \text{ is multiplicative,} \\ \alpha(M_1) &= \alpha(M_2) \text{ if } M_1 \cong M_2. \end{aligned}$$

View interval as a “closed” 1-manifold.
$\alpha = (\alpha_I, \alpha_\circ)$ is determined by its values $\alpha_I(\omega)$ on decorated intervals and values $\alpha_\circ(\omega)$ on decorated circles:

$$\alpha_I(\omega) = 1 \Leftrightarrow \omega \in L_I \quad \text{and} \quad \alpha_\circ(\omega) = 1 \Leftrightarrow \omega \in L_\circ.$$

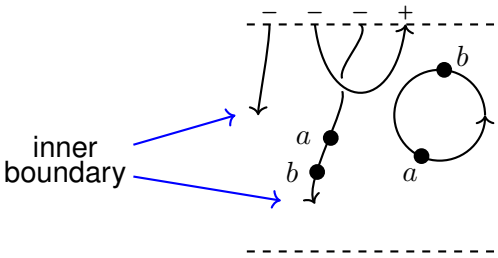
Universal construction starts with a (multiplicative) evaluation of closed n-dimensional objects and produces state spaces for $(n-1)$-dimensional objects and maps for n-cobordisms between these objects.

Use universal construction to define state spaces of oriented 0-dimensional manifolds (sign sequences $\varepsilon = (---+)$, for example). $n = 1$ in our case.

Sign sequence: $\varepsilon = (---+)$. Sign sequences are objects of our category of 1-dim cobordisms with 0-dim defects in Σ.

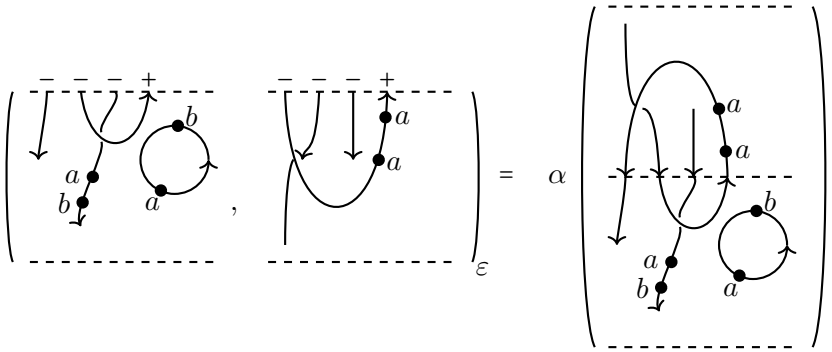
From α, one can define state spaces $A(\varepsilon)$ for 0-dimensional objects ε, by starting with a free $\mathbb{B}$-semimodule $\mathrm{Fr}(\varepsilon)$ with a basis $\{[M]\}_{\partial M \cong \varepsilon}$ given by formal symbols $[M]$ of all 1-dimensional objects M which have ε as outer boundary (with a fixed diffeomorphism $\partial M \cong \varepsilon$).

A state in the state space $A(\varepsilon)$:



On $\mathrm{Fr}(\varepsilon)$, introduce a bilinear pairing $(\ ,\ )_\varepsilon$ given on basis elements $[M_1], [M_2]$ with $\partial M_1 \cong \varepsilon \cong \partial M_2$ by coupling M_1, M_2 along the boundary and evaluating the resulting closed object $M_1 \cup_\varepsilon M_2$ via α:

$$([M_1], [M_2])_\varepsilon := \alpha(M_1 \cup_\varepsilon \overline{M_2}).$$



Note that $A(+) \cong A(-)^* = \mathrm{Hom}(A(-), \mathbb{B})$ via $\omega \mapsto (\omega' \mapsto \alpha(\omega'\omega)) \in \mathbb{B}$.

Now define the state space $A(\varepsilon)$ as the quotient of $\mathrm{Fr}(\varepsilon)$ by an equivalence relation,

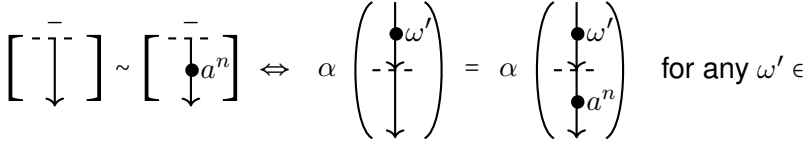
$$A(\varepsilon) \;:=\; \mathrm{Fr}(\varepsilon)/\sim,$$

where $\sum_i [M_i] \sim \sum_j [M'_j]$ if for any M with $\partial M = \varepsilon$,

$$\sum_i \alpha(M_i \cup_\varepsilon \overline{M}) = \sum_j \alpha(M'_j \cup_\varepsilon \overline{M}) \;\in\; \mathbb{B} = \{0, 1 : 1+1=1\}.$$

State space $A(\varepsilon)$ is spanned by $\mathbb{B}$-linear combinations of 1-manifolds M with $\partial M \cong \varepsilon$, modulo relations: two linear combinations are equal if for any way to close them up and evaluate using α, the result is the same.

One of the relations for the language $L_I = (a+b)^* b (a+b)$:



for any $\omega' \in \Sigma^*$.

If $\omega' = ba$, then

$$\alpha \left(\begin{array}{c} a \\ b \\ \cdot\!\!\times\!\!\cdot \\ \downarrow \end{array} \right) = \alpha \left(\begin{array}{c} a \\ b \\ \cdot\!\!\times\!\!\cdot \\ a^n \\ \downarrow \end{array} \right) = 1$$

If $\omega' = ab$, then

$$\alpha \left(\begin{array}{c} b \\ a \\ \cdot\!\!\times\!\!\cdot \\ \downarrow \end{array} \right) = \alpha \left(\begin{array}{c} b \\ a \\ \cdot\!\!\times\!\!\cdot \\ a^n \\ \downarrow \end{array} \right) = 0$$

State spaces $A(-), A(+)$ depend only on the interval language L_I, not on the circular language $L_\circ$ (spaces $A(+-)$, etc. depend on both).

An evaluation table of the language $L = (a+b)^*b(a+b)$ to compute the bilinear form on our spanning sets for $A(+)$ and $A(-)$ with values in $\mathbb{B}$. The matrix is not symmetric.

spanning elmt \ spanning elmt	x	x	y	x	z	y	$y+z$
x'	0	0	0	0	1	0	1
y'	0	0	1	0	0	1	1
y'	0	0	1	0	0	1	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
z'	1	1	1	1	1	1	1
z'	1	1	1	1	1	1	1

Defining relations:

$$x + y = y$$

$$x + z = z$$

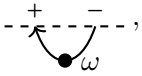
$$A(-) = \frac{\mathbb{B}x \oplus \mathbb{B}y \oplus \mathbb{B}z}{\langle x + y = y, x + z = z \rangle}$$

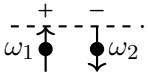
Consists of 5 elements:

$\{0, x, y, z, y+z\}$,

with x, y, z irreducible.

State space of $A(+-)$ is spanned by:



,


A 1-manifold M with $\partial M = \varepsilon' \sqcup -\varepsilon$ induces a map $A(\varepsilon) \longrightarrow A(\varepsilon')$ by concatenation.

Get a functor from category of Σ-decorated oriented 1-dim cobordisms to $\mathbb{B}$-semimodules. No subtraction in $\mathbb{B}$-semimodules; can add only.

A $\mathbb{B}$-semimodule V is a commutative idempotented monoid under addition:

$$a + a = a \text{ for } a \in V \text{ since } 1 + 1 = 1. \text{ Also } 0 + a = a,$$
$$a + b = b + a, \qquad (a + b) + c = a + (b + c).$$

Such V correspond to sup-semilattices, with join $a \vee b := a + b$, $a \leq b$ iff $a + b = b$.

0 is the minimal element, i.e., $0 \leq a$ for any a.

Any finite sup-semilattice is a finite lattice, with meet $a \wedge b := \sum_{c \leq a,b} c$ and $1 = \sum_{c \in V} c$.

In lattices:

Join: idempotent $a \vee a := a + a = a$ (since $1 + 1 = 1 \in \mathbb{B}$), commutative, associative.

Meet: idempotent $a \wedge a := \sum_{c \leq a} c = c + \ldots + a = a$, commutative, associative.

Meet and join: $a \vee (a \wedge b) = a, \qquad a \wedge (a \vee b) = a.$

Why?

$$a \vee (a \wedge b) = a + \left(\sum_{c \leq a,b} c \right) = a \qquad \text{and}$$
$$a \wedge (a \vee b) = a \wedge (a + b) = \sum_{c \leq a, a+b} c = a.$$

We mostly use $\mathbb{B}$-semimodule structure (join, not meet).

$$\mathbb{B}\text{-semimodules} \Leftrightarrow \text{comm. idemp. monoids} \Leftrightarrow \text{sup-semilattices (with } 0)$$

$$\text{finite (sup)-semilattices} \Leftrightarrow \text{finite lattices}$$

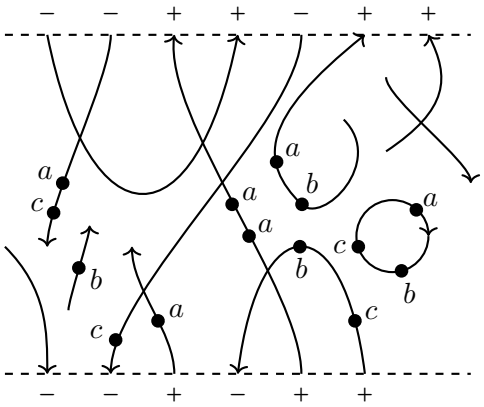
$\mathbb{B}$-semimodules constitute a category; morphisms are semimodule homomorphisms $f : V \longrightarrow W$, $f(0) = 0$, $f(a + b) = f(a) + f(b)$.

$\mathrm{Hom}(V, W)$ is a $\mathbb{B}$-semimodule (category $\mathbb{B}-\mathrm{mod}$ has internal homs). But $\mathbb{B}-\mathrm{mod}$ is not a rigid category (cannot "bend" objects and morphisms).

Subcategory of finite projective $\mathbb{B}$-semimodules (finite distributive (semi)lattices) is rigid.

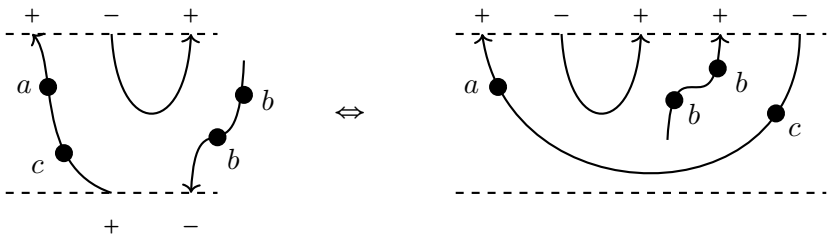
Categories of cobordisms in the universal construction that we build from evaluations are rigid.

Any cobordism C between $\varepsilon, \varepsilon'$ induces a semimodule homomorphism $A(\varepsilon) \to A(\varepsilon')$ of concatenation with C:



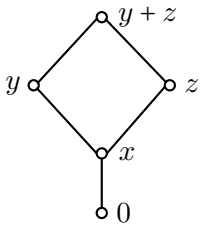
A cobordism from $(--+-++)$ to $(--++-++)$.

A cobordism from ε to ε' can be viewed as an element in the state space $A(\varepsilon' \sqcup -\varepsilon)$, i.e., a cobordism $C : \varepsilon = (+-) \to \varepsilon' = (+-+)$ corresponds to a state in the state space $A(+-+++-)$:

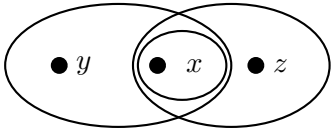


Recall the language $L = (a+b)^*b(a+b)$. The module $A(-)$ is spanned by x, y, z, and has relations $x + y = y$ and $x + z = z$.

The semimodule consists of 5 elements: $\{0, x, y, z, y+z\}$. The lattice corresponding to this language is:



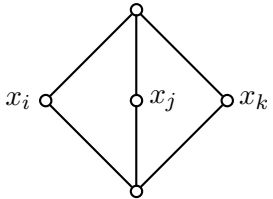
The finite topological space associated to this example:



Lattices that come from finite topological spaces are distributive.

If a lattice contains either as a sublattice,

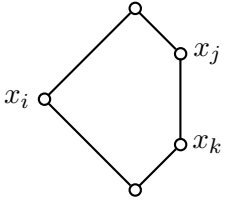
$$x_i + x_j = x_i + x_k = x_j + x_k$$



$$x_i \cap x_j = x_i \cap x_k = x_j \cap x_k$$

$$x_i \cap x_j + x_j = x_j$$

$$x_i \cap x_j < x_i, x_j, x_k$$



$$x_i + x_j = x_i + x_k \text{ for } j < k$$

$$x_j + x_k = x_j \text{ for } j < k$$

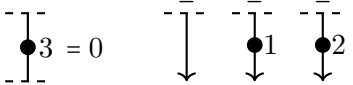
$$x_i \cap x_k < x_i, x_k$$

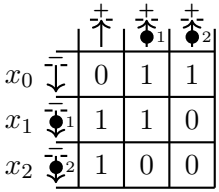
$$x_i, x_j < x_i + x_j$$

then the lattice is not distributive.

In such a case, there is no finite topological space associated to the language.

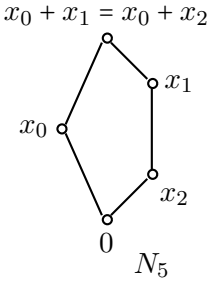
Example: for the language $L_I = \{a, a^2\}$, lattices $A(-), A(+)$ are not distributive.



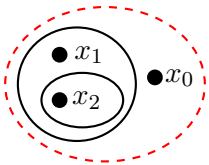


$x_1 + x_2 = x_1$

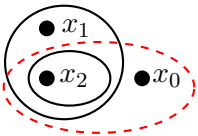
$x_0 + x_1 = x_0 + x_2$



For the language $L_I = \{a, a^2\}$, how should we draw the finite topological space associated to L_I?



But $x_0 \neq x_0 + x_1$. So the open set containing x_0 cannot be the entire space.



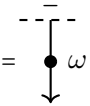
But since $x_0 \neq x_0 + x_2$, this finite topological space does not correspond to L_I as well.

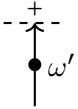
Theorem. Languages $L_I, L_\circ$ are regular iff the state space $A(\varepsilon)$ is a finite $\mathbb{B}$-semimodule for all sequences ε.

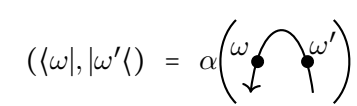
Get a $\mathbb{B}$-valued topological theory with finite hom spaces for any such pair of languages.

To recover minimal automaton for L_I, consider the state space $A(-)$. It consists of $\mathbb{B}$-linear combinations of diagrams below on the left, modulo equivalence relations coming from the pairing

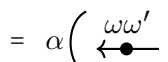
$$A(-) \times A(+) \longrightarrow \mathbb{B}.$$

$\langle\omega| :=$ 

$|\omega'\rangle :=$ 

$(\langle\omega|, |\omega'\rangle) = \alpha($ 

$)$

$= \alpha($ 

$) = \alpha_I(\omega\omega')$

How do we build the minimal deterministic FSA and nondeterministic FSA for L_I from $A(-)$?

Free monoid Σ^* generated by Σ (monoid of words) acts on $A(-)$, by composing with dots at the end of the strand.

State space $A(-)$ contains the subset $Q^- = \{\langle\omega|\}$ of *pure* states. Q^- is then the set of states of the minimal *deterministic* FSA for L_I. Action of Σ comes from restriction of its action on $A(-)$ (action by concatenation with dots at the top).

Initial state $q_{\text{in}} = \langle\varnothing|$. A state $\langle\omega|$ is accepting iff $\alpha_I(\omega) = 1$. *Nondeterministic* FSA for L_I come from coverings of $A(-)$ by free $\mathbb{B}$-modules with lifted action of Σ and unit, trace α maps.

$\widetilde{m}_a \circlearrowright \mathbb{B}^J$ free semimodule cover; minimal NFA for L_I, where $J = \mathrm{irr}(A(-))$ ($a \neq b + c$, where $b \neq a$, $c \neq a$)

$\downarrow$

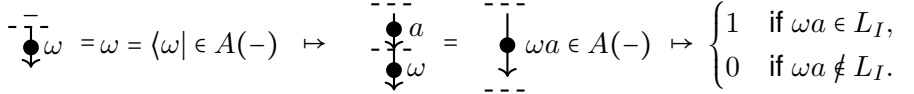
$m_a \circlearrowright A(-)$ state space of 0-manifold $-$

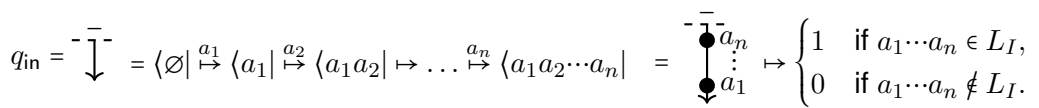
$\uparrow$

$m_a \circlearrowright Q_-$ minimal DFA for L_I

Every word gives a diagram in $A(-)$.

Start with a state and take images of all $\omega \in A(-)$ under the action by Σ^*, i.e.,

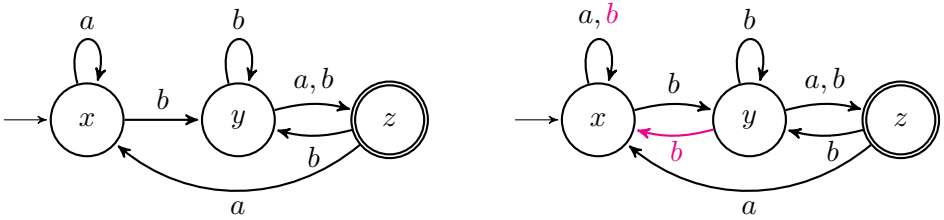




In general, there could be more than 1 minimal NFA.

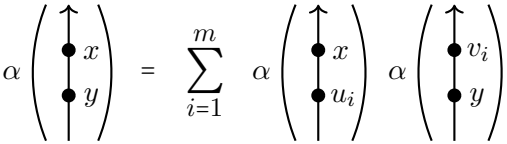
Two minimal nondeterministic automata on 3 states that accept the language $L = (a+b)^* b(a+b)$.

The second automaton has an additional b arrow from y to x and an additional b loop at x.



Multiple minimal NFA for L appear due to several ways of lifting action of Σ^* from $A(-)$ to $\mathbb{B}^J$.

Some regular languages allow decomposition of identity

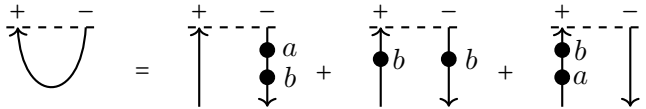


for some set of pairs of words (u_i, v_i), $1 \le i \le m$.

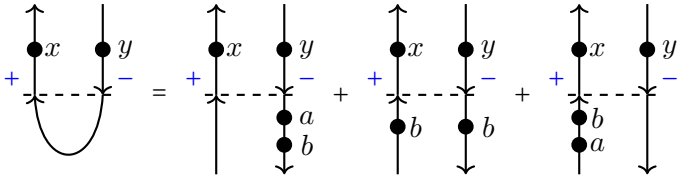
That is, for any $x, y \in \Sigma^*$,

$$\alpha_I(xy) = \sum_{i=1}^{m} \alpha_I(xu_i)\alpha_I(v_i y).$$

Returning to our example $L = (a+b)^*b(a+b)$,

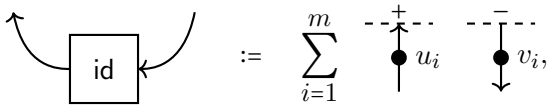


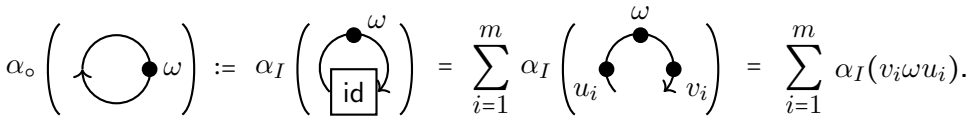
So



$$\alpha_I(xy) = \alpha_I(x)\alpha_I(bay) + \alpha_I(xb)\alpha_I(by) + \alpha_I(xba)\alpha_I(y).$$

For L_I with a decomposition of the identity, there is a unique associated circular language such that the decomposition still holds:





This gives a $\mathbb{B}$-valued TQFT: $A(\varepsilon)$ is the tensor product of $A(+), A(-)$ for the sequence of signs in ε.

For example, $A(++-) \cong A(+) \otimes A(+) \otimes A(-)$.

This is a TQFT for oriented 1-manifolds with 0-dimensional Σ-labelled defects, valued in the Boolean semiring $\mathbb{B}$.

Proposition. A regular language L has a decomposition of the identity if and only if $A(-)$ is a projective $\mathbb{B}$-semimodule (equivalently, a distributive lattice).

A finite semimodule P is projective if it's a retract of a free semimodule:

$$P \xrightarrow{\iota} \mathbb{B}^n \xrightarrow{p} P, \quad p\iota = \mathrm{id}_P.$$

Note that $\iota \circ p$ is an idempotent.

Such semimodules correspond to finite topological spaces X, with elements of the semimodule given by open subsets $U \subset X$ and $U + V := U \cup V$.

Thank you for your guidance, support, and trust.