

Minimization of register transducers

Gabriele Puppis Christian Bianchini

University of Udine

LICS, 19-23 July 2026

Outline

- ▶ provide a general proof for existence of *minimal models* of streaming transducers
- ▶ categorical framework to abstract and generalize to various models
- ▶ minimization as an algebraic property of *sub-objects/quotients* of the transducers

Transducer models

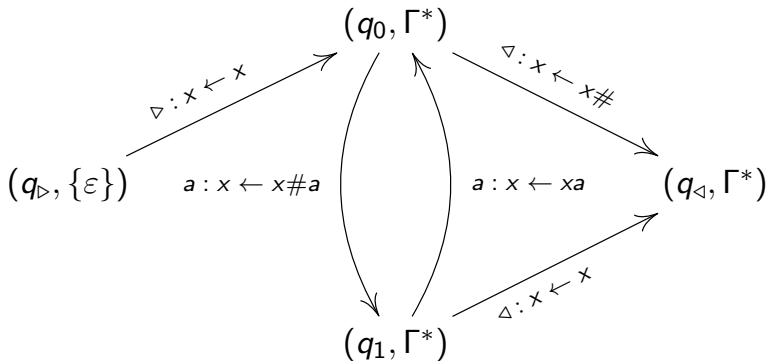
Streaming transducers:

- ▶ consume input from left to right
- ▶ manipulate data (e.g. registers) to construct output
- ▶ cannot test data to control flow

Examples of captured models:

- ▶ Subsequential string-to-string transducers
- ▶ Streaming string-to-string transducers (SSTs)
- ▶ Polynomial transducers
- ▶ Streaming string-to-term transducers (STTs), either in the upward or downward variant

Example of transducer



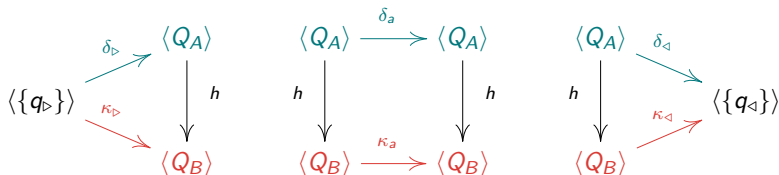
Transducer morphism

A morphism from a transducer $A = (Q_A, \delta_{\triangleright}, \{\delta_a\}_{a \in \Sigma}, \delta_{\triangleleft})$ to a transducer $B = (Q_B, \kappa_{\triangleright}, \{\kappa_a\}_{a \in \Sigma}, \kappa_{\triangleleft})$ is a total function $h : \langle Q_A \rangle \rightarrow \langle Q_B \rangle$ between the configuration spaces of A and B such that:

$$h \circ \delta_{\triangleright} = \kappa_{\triangleright}$$

$$h \circ \delta_a = \kappa_a \circ h, \\ \text{for all } a \in \Sigma$$

$$\delta_{\triangleleft} = \kappa_{\triangleleft} \circ h$$

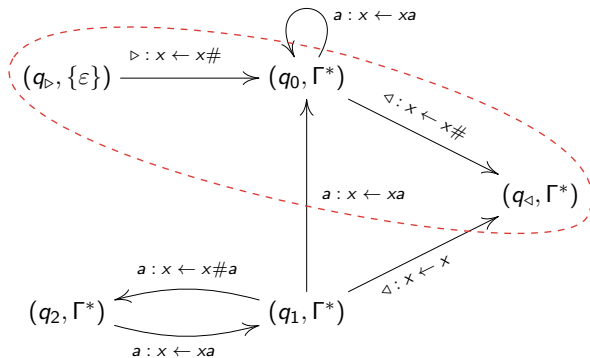


Sub-objects

A sub-object of B :

Intuitively: A obtained by “pruning” states from B

Categorically: there is a **strong mono** $A \twoheadrightarrow B$

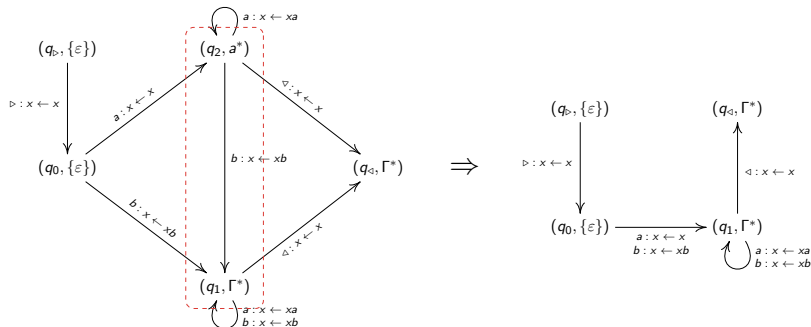


Quotients

Quotient:

Intuitively: A obtained by merging “equivalent” states of B

Categorically: there is an **epi** $B \twoheadrightarrow A$



Algebraically Minimal Object

Minimal object:

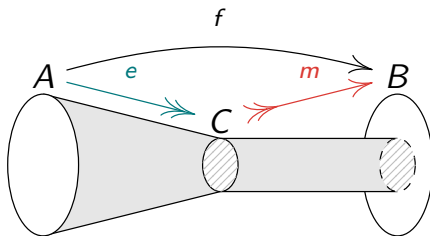
A transducer M is algebraically minimal if it is a *quotient of a sub-object* of every other equivalent transducer.

Sufficiency conditions from category theory

initial transducer I : every equivalent transducer A admits a unique morphism $f^A : I \rightarrow A$

final transducer F : every equivalent pruned transducer A admits a unique morphism $f_A : A \rightarrow F$

images of morphisms: every morphism f has a (unique) object C with an epi $e : A \rightarrow C$ and a strong mono $m : C \rightarrow B$ such that $f = m \circ e$



Minimal object as quotient of subobject

$\text{Reach}(A)$: image of initial morphism f^A

$\text{Obs}(A)$: image of final morphism f_A

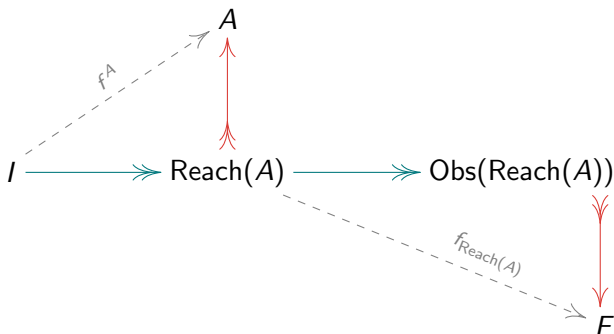


Figure: Overview of the category

Minimal object as quotient of subobject

$\text{Reach}(A)$: image of initial morphism f^A

$\text{Obs}(A)$: image of final morphism f_A

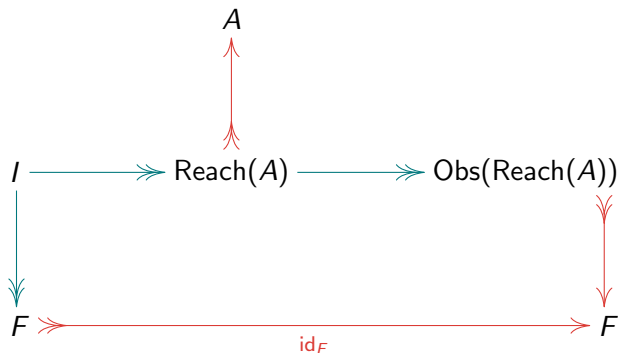


Figure: Overview of the category

Minimal object as quotient of subobject

$\text{Reach}(A)$: image of initial morphism f^A

$\text{Obs}(A)$: image of final morphism f_A

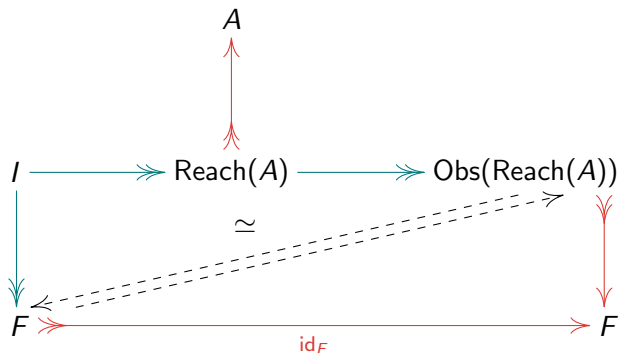
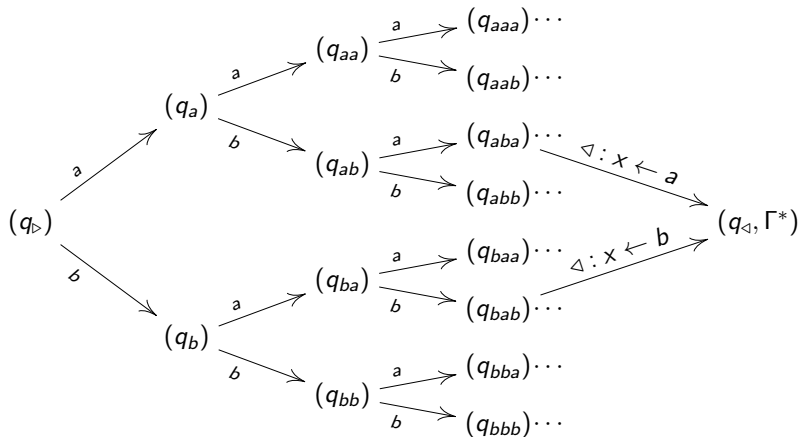


Figure: Overview of the category

Initial object

- ▶ infinite transducer with Σ^* as states, initial state is ε
- ▶ no internal registers
- ▶ $\phi(w)$ (when defined) as the final transition for the state w



Final object (Minimal object)

- ▶ final in the subcategory of initial images
- ▶ leverage an equivalence à la *Myhill-Nerode*
- ▶ need that every set of updates admits a **greatest common divisor**

Example

$$\text{GCD} \{x \leftarrow x \cdot abc, x \leftarrow x \cdot abd\} = x \leftarrow x \cdot ab$$

Images of morphisms

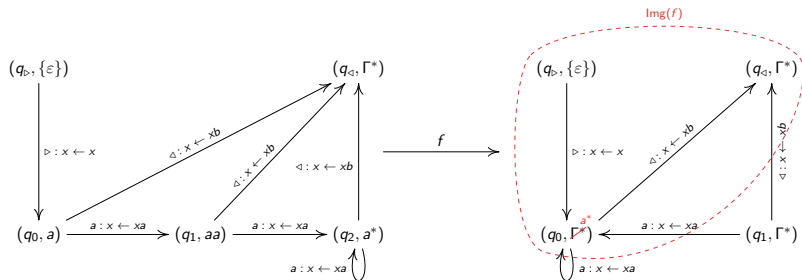


Figure: Example image of a transducer morphism

Minimization result

Theorem

*A class of transducers admits algebraically minimal models if and only if the underlying data structure admits **greatest common divisors** and **constrained domains** for every non-empty set of updates.*

In addition, if every constrained domain is also finitary, then algebraically minimal models also exist within the subclass of finite transducers.

Downward STTs

- ▶ data structure: terms containing variables occurrences
- ▶ domain $\mathbb{D}_n$: terms with exactly n variables
- ▶ update from $\mathbb{D}_n$ to $\mathbb{D}_m$: n terms with a total of m variables
- ▶ updates are copyless and non-erasing

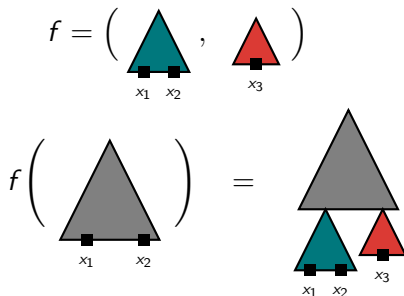


Figure: Example of STT update in $\mathbb{D}_{2 \rightarrow 3}$

Downward STTs

- ▶ constrained domains: trivial (either $\mathbb{D}_n$ or $\emptyset$)
- ▶ GCDs: anti-unifiers

$$\text{GCD} \left\{ \left(\begin{array}{c} \text{gray triangle} \\ \text{teal triangle} \quad \text{red triangle} \end{array}, \text{red triangle} \right), \left(\begin{array}{c} \text{gray triangle} \\ \text{teal triangle} \quad \text{green triangle} \end{array}, \text{teal triangle} \right) \right\} = \left(\begin{array}{c} \text{gray triangle} \\ \text{teal triangle} \quad \text{black square} \end{array}, \text{black square} \right)$$

Figure: Example of anti-unifier used as GCD

Conclusions

Results

- ▶ Necessary and sufficient conditions for the existence of the minimal transducer
- ▶ Application to downward and upward STT

Future work

- ▶ Minimization of SST
- ▶ Description of an algorithm for STT

Thank you for
your attention