Categorical information flow

Tahiry Rabehaja[1], Annabelle McIver[2], Carroll Morgan[3], and Georg Struth[4]

[1] Optus Macquarie University Cyber Security Hub, Australia
[2] Dept. of Computing, Macquarie University, Australia
[3] School of Comp. Sci. and Eng., UNSW, and Data61, Australia
[4] Dept. of Comp. Sci., The University of Sheffield, United Kingdom

Abstract. We propose a categorical model for information flows of correlated secrets in programs. We show how programs act as transformers of such correlations, and that they can be seen as natural transformations between probabilistic constructors. We also study some basic properties of the construction.

1 Introduction

The foundations for Quantitative Information Flow (*QIF*) [2–5, 14, 15] has had great successes in explaining the impact of information flows in security systems. A basic assumption in *QIF* is that there is a single secret, it does not change over time, and that one risk of observing the behaviour of the system is that partial information about the secret leaks out. A *QIF* analysis is then able to answer questions such as "can adversaries use the information leaked to their advantage?"

In contexts where the secret does change over time however, the analysis becomes considerably more difficult. Even more challenging are contexts where there are multiple secrets, some correlated with others, some that change, whilst others remain unchanged. Here we are concerned with this most egregious case.

Consider the scenario of A who is burdened by her company's policy of forcing employees to change passwords every month. Moreover, since the company wants its employees to put serious effort into picking a strong password, they measure the time it takes for the new password to be selected, to avoid the situation that the employees will simply add another digit to their current password. A decides to use a workaround to this draconian practice: she writes a short but effective password selection program, which first allows time to pass for some random interval determined by the current value of the password, and then selects independently a completely random password which is ultimately stored automatically in A's keychain. Her code looks something like this, where `pw` stands for her current password.

```
while (pw>0) { // Wait for some time
  pw--;
}
pw:= uniform(0, N) // Select uniformly at random a new password
```

Over lunch, A tells her friend G about her scheme; G is a little concerned because of a known timing attack on the loop which counts down from the initial value of pw. However, since it only reveals the *initial* value of pw, which is then immediately reset to a completely random value, there seems little to worry about.

A few months later the IT department introduces a new computer system to supplement the existing one. Whilst some aspects of her work environment improve, unfortunately now A has to reset *two* passwords every month — one for the old system, and an additional one for the supplementary system. She decides on an easy extension of her password updating program, setting her new password npw to her old password, and then updating pw as before:

```
npw:= pw;
while (pw>0) { // Wait for some time
  pw--;
}
pw:= uniform(0, N) // Select uniformly at random a new password
```

Unfortunately, since it has been months since she started using the original scheme, she has forgotten about G's brief analysis concerning the leaking of the initial value of pw, and indeed there was nothing to worry about since there was only one secret. But now, in this extended context, there is a serious vulnerability, not for pw, but for npw — its value is correlated with the initial value of pw which is completely revealed through the timing attack during the iteration of the loop. Thus G's original analysis, which assumed a single secret, although effective under that assumption, turns out to be unhelpful in general. [5]

In this paper we consider the general situation illustrated by that story. We find that any general analysis should take into account possible *future correlations* with other secrets. Our general approach suggests that secrets should never be considered in isolation, and that instead the basic currency of *QIF* should be *correlations of secrets*, and that any analysis should study leaks about correlations. It turns out that a categorical approach can be used to give a description of this situation. In particular we show how information flow can be summarised by a natural transformation between two constructors for structuring prior and posterior knowledge.

This paper is organised as follows. In Sec. 2, we revise the use of *HMM*'s in modelling information flow pertaining to static and dynamic secrets. We also introduce the notion of correlated secrets. In Sec. 3, we construct a category within which correlated secrets live. This form the fundamental basis for our new compositional semantics where security programs modelled as *HMM*'s are in turn interpreted as natural transformations over this category. We also explore some characteristic properties of this category and the new semantic maps. We conclude in Sec 4.

[5] Recall the Ariane disaster, which occurred when the software was executed within an environment for which it was not originally designed.

2 Channels, *HMM*'s and secrets that change

Hidden Markov Models or *HMM*'s are commonly used to solve inference problems such as to infer a sequence of hidden states that could have produced a sequence of observations. The most well known practical application is speech recognition, where a *HMM* is trained to recognise and translate speech waveforms into texts. *HMM*'s have also been applied to cryptanalysis where a cryptographic algorithm can emit a sequence of observations if given a sequence of states. In [7], Karlof and Wagner use *HMM*'s to model noisy side channels. Their work provides a generalised framework through which a large class of side channel attacks can be modelled. More recently, Smith provided a general foundation upon which the notion of information leakage is precisely defined, and then computed [14]. Even though Smith works mainly with static secrets, his notions of channel and their leakages are central to the subsequent advancement of Quantitative Information Flow (*QIF*).

Fundamental to *QIF* is the idea of an information flow channel. It models the process by which a secret with values drawn from a basic type $\mathcal{X}$ can be partially revealed to an adversary. We say that a channel C is an $\mathcal{X}{\times}\mathcal{Y}$ *stochastic matrix* denoted by $\mathcal{X}{\rightarrowtriangle}\mathcal{Y}$. Here $\mathcal{Y}$ denotes the set of observables available to a passive but curious adversary. [6] An entry C_{xy} is the (conditional) probability that, given the secret has some value x in $\mathcal{X}$, then the observation is $y \in \mathcal{Y}$. Stochastic means that, for each x in $\mathcal{X}$, we have that $\sum_{y\in\mathcal{Y}} C_{xy} = 1$. What can an adversary do with this information? We assume first that the adversary has some prior knowledge about the possible values of $\mathcal{X}$ and that this is modelled as a probability distribution $\pi \in \mathbb{D}\mathcal{X}$. Once we have a channel C and a prior π we can form the joint distribution $\pi\rangle C$ in $\mathbb{D}(\mathcal{X}{\times}\mathcal{Y})$ defined

$$(\pi\rangle C)_{xy} \quad := \quad \pi_x C_{xy} \ , \tag{1}$$

which describes the probability that the secret is x and the observation is y. We focus on two interesting distributions that provide some insight into what exactly has been leaked by the channel, and what the adversary can do with that leak.

The first is the marginal distribution over the observation set $\mathcal{Y}$ — this is the probability that a specific y occurs when we don't know which particular x occurred. It is

$$(\pi\rangle C)_{-y} \quad := \quad \sum_{x\in\mathcal{X}} (\pi\rangle C)_{xy} \ .$$

Next we look at the adversary's revised knowledge of the secret given that observation y has occurred. This is the *posterior* distribution wrt y and is the conditional distribution Eqn. (1) relative to y:

$$(\delta^y)_x \quad := \quad (\pi\rangle C)_{xy}/(\pi\rangle C)_{-y} \ .$$

[6] We do not assume that $\mathcal{X}$ and $\mathcal{Y}$ are finite. We do however assume that they are discrete and countable.

It turns out that the marginal and the posterior disributions can be combined to form a hyper-distribution –a distribution over distributions– of type $\mathbb{D}\mathbb{D}\mathcal{X} = \mathbb{D}^2\mathcal{X}$, where the outer distribution corresponds to the marginal (at some y) and the inner distribution corresponds to the posterior. We write $[\pi\rangle C]$ for the hyper-distribution corresponding to the joint distribution $\pi\rangle C$, where for $\delta{:}\,\mathbb{D}\mathcal{X}$ the posterior corresponding to observation y, we have

$$[\pi\rangle C]_\delta \quad := \quad \sum_{y:\mathcal{Y}_\delta} (\pi\rangle C)_{-y} \ ,$$

where $\mathcal{Y}_\delta \subseteq \mathcal{Y}$ is the subset of $\mathcal{Y}$ consisting of observations y such that δ and δ^y are the same distribution, i.e. $\delta_x = (\delta^y)_x$ for all x. Note that the hyper-distribution formulation represents the collection of posteriors and their respective marginal probabilities related to the joint distribution Eqn. (1) *except* that the outer probability $[\pi\rangle C]_\delta$ is the *sum* of the marginal probabilities associated with any other posterior that is the same as δ. The reason is that separating those posteriors does not affect the computation of information flow [12]. Using this fact allows us to give an abstract information flow semantics as *abstract channels* of type priors to hyper-distributions, that is $[\![C]\!]{:}\,\mathbb{D}\mathcal{X}{\rightarrow}\mathbb{D}^2\mathcal{X}$ defined by

$$[\![C]\!].\pi \quad := \quad [\pi\rangle C]\ ^7 \ .$$

In our example in Sec. 1 above, we see a more complicated case of information flow, where the original secret `pw` was updated. In other work [9] we showed how a generalisation of abstract channels called abstract Hidden Markov Models (*HMM*'s) can be used to model this situation. *HMM*'s combine the effect of information flow (as defined above) for channels together with "Markov updates" to describe how secrets can be changed. Intuitively, a single *HMM* step describes the effect of first leaking some information about a secret through a channel matrix C followed by an immediate update of the secret via a Markov matrix M. This single step *HMM* is denoted by a matrix $H = (C{:}M){:}\,\mathcal{X}{\multimap}\mathcal{Y}{\times}\mathcal{X}$ such that

$$H_{xyx'} \quad := \quad C_{xy}M_{xx'} \ .$$

More complicated *HMM*'s can be formed by sequentially composing smaller *HMM* matrices to give a pattern of leaking information about the *current* value of the secret, followed by a possible update. For example the sequential composition of two *HMM*'s $H{:}\,\mathcal{X}{\multimap}\mathcal{Y}_1{\times}\mathcal{X}$ and $K{:}\,\mathcal{X}{\multimap}\mathcal{Y}_2{\times}\mathcal{X}$ gives a *HMM* $H{;}K{:}\,\mathcal{X}{\multimap}(\mathcal{Y}_1{\times}\mathcal{Y}_2){\times}\mathcal{X}$ such that

$$(H;K)_{x(y_1,y_2)x'} \quad := \quad \sum_z H_{xy_1z}K_{zy_2x'} \ . \tag{2}$$

We note that the "observations" now record a pair of elements from the observable set $\mathcal{Y}_1{\times}\mathcal{Y}_2$, and in general the observations for n sequential compositions will be a trace of length n.

[7] $f.x$ denotes the application of a function f to the argument x.

Just as for channels, *HMM*'s also transform initial prior information into a hyper-distribution through the construction of a joint distribution. But now consider the situation of executing an *HMM* over basic secret of type $\mathcal{X}$ which is correlated with some other secret of type $\mathcal{Z}$, by which we mean that there is a dependency between $\mathcal{X}$ and $\mathcal{Z}$ described by a joint distribution. This was the situation described in Sec.1, where the *explicit* leak about `pw` caused a *collateral* leak about a different dependent secret `npw`. To account for this "collateral leak", we start with a correlation between the secret and some arbitrary secret ranging over $\mathcal{Z}$ modelled as a joint distribution $\Pi{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$. We can now track what happens to the correlation of the two secrets when *HMM* H executes. It results in a joint distribution in $\mathbb{D}(\mathcal{Z}{\times}\mathcal{X}{\times}\mathcal{Y})$:

$$(\Pi\rangle H)_{zx'y} \quad := \quad \sum_x \Pi_{xz} H_{xyx'} \ .$$

Next, as we did for abstract channels, we abstract from the observation name y to obtain a hyper-distribution in $[\Pi\rangle H]{:}\,\mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z})$ such that $[\Pi\rangle H]_\Delta$ is the probability that $\Delta{:}\,\mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z})$ occurs as a (posterior) distribution. This implies that the abstract type of our *HMM* H, given some other secret $\mathcal{Z}$ that is not changed by H, but could be correlated with $\mathcal{X}$ is now

$$[\![H]\!]^{\mathcal{Z}}{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}){\rightarrow}\mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z}) \ , \tag{3}$$

where correlation $\Pi{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$ is mapped to hyperdistribution

$$[\![H]\!]^{\mathcal{Z}}.\Pi = [\Pi\rangle H] \ . \tag{4}$$

Here, we see that this type captures clearly the role of H — that it is a mechanism for leaking and changing secrets, but in a way that it has an effect on other, possibly correlated secrets, which might or might not become relevant in a wider context. In fact the abstract semantics is now a set of transformations (parameterised by $\mathcal{Z}$). We show below in Sec. (3) that in fact H can be seen as a natural transformation.

Before moving further, let us see how the program used by A in the previous section is expressed in our framework. It is made up of 3 parts: the first assignment `npw:= pw`, followed by the while loop, and then the final resetting of `pw` to a random value. The first assignment establishes a correlation Π between `pw` and `npw`, which describes a distribution with the property that the only non-zero entries occur when `pw` and `npw` are the same. The while loop leaks the exact value of `pw` and then sets it to 0; recall that a curious eavesdropper would be able to count how many iterations the loop performs. If we assume that $\mathcal{X} = \{0, 1\}$, the while loop is then represented by the following *HMM* matrix:

$$\begin{array}{cc} & \begin{array}{cccc} \multicolumn{2}{c}{\overbrace{\hphantom{0\quad 1}}^{\circ\texttt{S}}} & \multicolumn{2}{c}{\overbrace{\hphantom{0\quad 1}}^{\circ\texttt{B}}} \\ \texttt{0} & \texttt{1} & \texttt{0} & \texttt{1} \end{array} \\ \begin{array}{c}\texttt{0}\\ \texttt{1}\end{array} & \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \end{array}$$

The labels $\circ\texttt{B}$ and $\circ\texttt{S}$ denote the observations that the body of the loop was executed or not, respectively. The other column labels come from the column labels in the execution of the body of the loop. Thus each column is labelled by a pair in $\{\circ\texttt{B}, \circ\texttt{S}\}\times\mathcal{X}$. The first row (labelled 0) describes the scenario where the initial value of `pw` is 0 so that the loop terminates immediately (i.e. there is a single 1 in the column $(\circ\texttt{S}, \texttt{0})$).

The last program that randomly sets `pw` is simply a Markov matrix with 0.5 in all positions. The two programs are composed using the sequential composition defined at Eqn. (2) above to model to model the effect of information leakage on the secret stored in `npw`.

3 A category of correlations

In order to study the collateral consequences of information flow of a program operating over a single secret, we must first find a mathematical space that treats correlations as “first class citizens”. In this section we describe the space of collateral types, show that it forms a category and we study some of its properties.

We begin by considering a basic relationship between two secrets. In general this relationship is not necessarily symmetric and can be expressed in the form of a channel, which means that observation of the value of one collateral secret may leak information regarding the value of the other collateral secret.

Definition 1. *The category* Corr *of collateral types contains finite sets $\mathcal{Z}_1, \mathcal{Z}_2$ as objects and channel matrices $Z{:}\,\mathcal{Z}_1 \rightarrowtriangle \mathcal{Z}_2$ between them as arrows. The composition of channel in this category is the matrix multiplication of channel matrices. That is, given $Z_1{:}\,\mathcal{Z}_1 \rightarrowtriangle \mathcal{Z}_2$ and $Z_2{:}\,\mathcal{Z}_2 \rightarrowtriangle \mathcal{Z}_3$, we have $Z_1{\cdot}Z_2{:}\,\mathcal{Z}_1 \rightarrowtriangle \mathcal{Z}_3$.*

Intuitively, objects of Corr correspond to spaces in which all of A’s passwords live. This includes `pw` and `npw`.

Lemma 1. Corr *is a category.*

Proof. This is clear since matrix multiplication behaves exactly as a functional composition and identity arrows are given by identity matrices.

Recall that Eqn. (3) gives the type of a *HMM* on the secret type $\mathcal{X}$ and taking into account a collateral type $\mathcal{Z}$. The main intuition here is that, instead of simply transforming information regarding the *mutable secret type* $\mathcal{X}$, the program transforms joint information, which is an object in $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$. When the mutable secret type $\mathcal{X}$ is fixed, the set $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$ is parametrised by $\mathcal{Z}$.

More precisely, for a fixed $\mathcal{X}$ we can construct a functor $\mathbb{F}_{\mathcal{X}}$ from Corr to a category containing $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$ as an object instance. Since $\mathcal{X}\times\mathcal{Z}$ is a finite set, the set $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$ is a compact metric space with respect to the Kantorovich metric [9, 13] [8]. We know that the compact metric spaces and the continuous

[8] In $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$, the Kantorovich metric is equivalent to the standard Euclidian distance.

functions between them form a category which we refer to as Comp [1,16]. The functor $\mathbb{F}_{\mathcal{X}}\colon \textsc{Corr}\to\textsc{Comp}$ constructs the set of correlations $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$, for each type $\mathcal{X}$.

Definition 2. *Let $\mathcal{X}$ be a fixed mutable type. We define the* collateral functor $\mathbb{F}_{\mathcal{X}}\colon \textsc{Corr}\to\textsc{Comp}$ *as follows:*

- *For objects, $\mathbb{F}_{\mathcal{X}}(\mathcal{Z}) := \mathbb{D}(\mathcal{X}\times\mathcal{Z})$ is the set of all joint distributions over $\mathcal{X}$ and $\mathcal{Z}$ for any collateral type $\mathcal{Z}$.*
- *For an arrow $Z\colon \mathcal{Z}_1 \rightarrowtriangle \mathcal{Z}_2$, we have $\mathbb{F}_{\mathcal{X}}(Z)\colon \mathbb{D}(\mathcal{X}\times\mathcal{Z}_1)\to\mathbb{D}(\mathcal{X}\times\mathcal{Z}_2)$ such that, for every $\mathit{\Pi}$ in $\mathbb{D}(\mathcal{X}\times\mathcal{Z}_1)$ we have*

$$\mathbb{F}_{\mathcal{X}}(Z)(\mathit{\Pi}) \quad := \quad \mathit{\Pi}\cdot Z \ , \tag{5}$$

  which is the matrix multiplication of $\mathit{\Pi}$ and Z.

The extra line `npw` := `pw` in the second program above creates a diagonal matrix in $\mathbb{F}_{\mathcal{X}}(\mathcal{X}')$ where $\mathcal{X}'$ is a copy of type $\mathcal{X}$. This produces a diagonal correlation between the values of these two variables.

Lemma 2. *$\mathbb{F}_{\mathcal{X}}$ is a functor.*

Proof. *Let $\mathcal{Z}$ be a collateral type. It is clear that $\mathbb{F}_{\mathcal{X}}(\mathrm{id})\colon \mathbb{D}(\mathcal{X}\times\mathcal{Z})\to\mathbb{D}(\mathcal{X}\times\mathcal{Z})$ is the identity map when $\mathrm{id}\colon \mathcal{Z}\rightarrowtriangle\mathcal{Z}$ is the identity channel matrix on $\mathcal{Z}$.*

The functor $\mathbb{F}_{\mathcal{X}}$ preserves composition follows from the associativity of matrix multiplication.

Finally, for well-definedness, as long as we start with finite (or compact) $\mathcal{X}$ and $\mathcal{Z}$, the set $\mathbb{D}(\mathcal{X}\times\mathcal{Z})$ is a compact metric space with the Kantorovich distance and the function $\mathbb{F}_{\mathcal{X}}(Z)$ is continuous [13].

We can rewrite (3) using the newly defined functors $\mathbb{F}_{\mathcal{X}}$. That is, the denotation of a program H with mutable type $\mathcal{X}$ and collateral type $\mathcal{Z}$ has type

$$[\![H]\!]^{\mathcal{Z}}\colon \mathbb{F}_{\mathcal{X}}(\mathcal{Z}) \to \mathbb{D}\mathbb{F}_{\mathcal{X}}(\mathcal{Z}) \tag{6}$$

Note however that the denotation of H should only depend on $\mathcal{X}$ explicitly. In other words, $[\![H]\!]^{\mathcal{Z}}$ is parametric in $\mathcal{Z}$. This means that $[\![H]\!]^{\mathcal{Z}}$ in (6) is actually a component of the denotation of the *HMM* H given that we know the collateral type is $\mathcal{Z}$. This is an important observation because it implies that the "unparametrised" denotation of H should have type

$$\mathbb{F}_{\mathcal{X}} \Rightarrow \mathbb{D}\mathbb{F}_{\mathcal{X}} \ ,$$

that is, $[\![H]\!]$ should transform the functor $\mathbb{F}_{\mathcal{X}}$ into the composite functor $\mathbb{D}\mathbb{F}_{\mathcal{X}}$. This suggests that the denotation of a program computing with secrets is a natural transformation because the parameter $\mathcal{Z}$ should be manipulated in a syntactic manner. More precisely, $[\![H]\!]$ should apply to the collateral type $\mathcal{Z}$ polymorphically. This is exactly the rationale behind our next definition.

Definition 3. *An* abstract program *with mutable type $\mathcal{X}$ is a natural transformation $h{:}\,\mathbb{F}_{\mathcal{X}}{\Rightarrow}\mathbb{D}\mathbb{F}_{\mathcal{X}}$. That is, for every correlated type $\mathcal{Z}_1, \mathcal{Z}_2$ and channel arrow $Z{:}\,\mathcal{Z}_1{\mathrel{-\!\!\triangleright}}\mathcal{Z}_2$, we have that the following diagram commutes:*

$$\begin{tikzcd}
\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_1) \arrow[r, "h^{\mathcal{Z}_1}"] \arrow[d, "\mathbb{F}_{\mathcal{X}}(Z)"'] & \mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z}_1) \arrow[d, "\mathbb{D}\mathbb{F}_{\mathcal{X}}(Z)"] \\
\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_2) \arrow[r, "h^{\mathcal{Z}_2}"'] & \mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z}_2)
\end{tikzcd}$$

This definition implies that an abstract program h operates on the category Corr directly. Its denotation is independent of the collateral types, that is, any relationship (channel) between $\mathcal{Z}_1$ and $\mathcal{Z}_2$ can be accounted for before or after the execution of h while preserving information flow.

The following result shows that if H is a *HMM* matrix then $[\![H]\!]$ is indeed a natural transformation.

Theorem 1. *For any HMM matrix H, $[\![H]\!]$ is a natural transformation.*

Proof. Let $\mathit{\Pi}{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_1)$ and $Z{:}\,\mathcal{Z}_1{\mathrel{-\!\!\triangleright}}\mathcal{Z}_2$ be any channel arrow. We need to show that

$$\mathbb{D}\mathbb{F}_{\mathcal{X}}(Z){\circ}[\![H]\!]^{\mathcal{Z}_1}.\mathit{\Pi} = [\![H]\!]^{\mathcal{Z}_2}.\mathbb{F}_{\mathcal{X}}(Z)(\mathit{\Pi}) \ . \tag{7}$$

We use Eqns. (4) *and* (5) *to instantiate the inners of the left and right hand side hyper distributions.*

On the one hand, an inner γ^y of the left-hand side of Eqn. (7)*, associated to observation y, takes the form*

$$\gamma^y_{x'z_2} = \frac{\sum_{z_1}\left(\sum_x \mathit{\Pi}_{xz_1}H_{xyx'}\right)Z_{z_1z_2}}{\overline{\gamma}^y}$$

where $\overline{\gamma}^y = \sum_{xx'z_1z_2}\mathit{\Pi}_{xz_1}H_{xyx'}Z_{z_1z_2}$.

On the other hand, an inner δ^y of the right hand side, associated with the observation y, takes the form

$$\delta^y_{x'z_2} = \frac{\sum_x\left(\sum_{z_1}\mathit{\Pi}_{xz_1}Z_{z_1z_2}\right)H_{xyx'}}{\overline{\delta}^y}$$

where $\overline{\delta}^y = \sum_{xx'z_1z_2}\mathit{\Pi}_{xz_1}Z_{z_1z_2}H_{xyx'}$. It follows, after some arithmetic, that $\delta^y = \gamma^y$.

This theorem shows that there are indeed natural transformations from $\mathbb{F}_{\mathcal{X}}$ to $\mathbb{D}\mathbb{F}_{\mathcal{X}}$, of which the denotations of *HMM* matrices form a subset. Thus, the first program in the introduction is also a natural transformation where $\mathcal{X}$ is the set of all possible values that A can choose for variable `pw`.

To compose arrows of type $\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}){\to}\mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z})$, we need to define the Giry monad $(\mathbb{D}, \eta, \mathsf{avg})$. The natural transformation η is the identity with $\eta{:}\,\mathcal{X}{\Rightarrow}\mathbb{D}$ and

$\eta_{\mathcal{X}}(x)$ is the point distribution at x. The natural transformation avg "squashes" higher level distributions (such as hyper-distributions) by averaging. That is, $\mathsf{avg}\colon\mathbb{D}^2{\Rightarrow}\mathbb{D}$ where, for each $\mathit{\Delta}{\in}\mathbb{D}^2\mathcal{X}$, we intuitively have

$$(\mathsf{avg}_{\mathcal{X}}(\mathit{\Delta}))_x = \sum_{\delta\in\mathbb{D}\mathcal{X}} \delta_x \mathit{\Delta}_\delta \ .$$[9]

It is shown in [10] that

$$\llbracket H;K\rrbracket^{\mathcal{Z}} \quad = \quad \llbracket H\rrbracket^{\mathcal{Z}};\llbracket K\rrbracket^{\mathcal{Z}} \quad := \quad \mathsf{avg}_{\mathbb{D}(\mathcal{X}\times\mathcal{Z})}{\circ}\mathbb{D}\llbracket K\rrbracket^{\mathcal{Z}}{\circ}\llbracket H\rrbracket^{\mathcal{Z}} \tag{8}$$

where the right hand side is the definition of sequential composition in the Giry monad [6].

Eqn. (8) is central for the compositionality result we show in [10]. It assumes that the mutable secrets of H and K are represented by the exact same variables ranging over $\mathcal{X}$. A slightly more general version is proven in the Appendix (Lem. 4).

Next, we show that the sequential composition of *HMM*'s in (2) is defined using the vertical composition of natural transformations $(k{\bullet}h)^{\mathcal{Z}}{:=}\,k^{\mathcal{Z}}{\circ}h^{\mathcal{Z}}$.

Theorem 2. *Let H and K be HMM matrices on the mutable secret type $\mathcal{X}$. Then $\llbracket H;K\rrbracket$ is also a natural transformation where*

$$\llbracket H;K\rrbracket = \mathsf{avg}_{\mathbb{F}_{\mathcal{X}}} \bullet \mathbb{D}\llbracket K\rrbracket \bullet \llbracket H\rrbracket \quad ,$$

in which $\bullet$ is the vertical composition of natural transformations.

Proof. Let $\mathcal{Z}$ be a collateral type. By definition of the vertical composition, we have

$$(\mathsf{avg}_{\mathbb{F}_{\mathcal{X}}}{\bullet}\mathbb{D}\llbracket K\rrbracket{\bullet}\llbracket H\rrbracket)^{\mathcal{Z}} = \mathsf{avg}_{\mathbb{F}_{\mathcal{X}}(\mathcal{Z})}{\circ}(\mathbb{D}\llbracket K\rrbracket)^{\mathcal{Z}}{\circ}\llbracket H\rrbracket^{\mathcal{Z}} = \mathsf{avg}_{\mathbb{D}(\mathcal{X}\times\mathcal{Z})}{\circ}\mathbb{D}\llbracket K\rrbracket^{\mathcal{Z}}{\circ}\llbracket H\rrbracket^{\mathcal{Z}}$$

The second equality follows from the definition of $\mathbb{F}_{\mathcal{X}}(\mathcal{Z})$ and the composition of a functor with natural transformation, namely, $(\mathbb{D}\llbracket H\rrbracket)^{\mathcal{Z}} = \mathbb{D}\llbracket H\rrbracket^{\mathcal{Z}}$.

The result now follows from Eqn. (8).

The denotation of a *HMM* is therefore polymorphic with respect to the object of the underlying category Corr. In particular, the mutable type $\mathcal{X}$ of the *HMM* is also an object of Corr and we will show that the arrow $h^{\mathcal{X}}$ completely characterises the natural transformation.

Let $\mathcal{Z}$ be an arbitrary collateral type and $J{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$ be any prior joint distribution. We can decompose J into a joint distribution $\mathit{\Pi}{:}\,\mathbb{D}\mathcal{X}^2$ and channel matrix $Z{:}\,\mathcal{X}{\rightarrowtriangle}\mathcal{Z}$ such that $\mathit{\Pi}\rangle Z = J$. This decomposition can be achieved within the context of the collateral functor $\mathbb{F}_{\mathcal{X}}$ and the construction is given in the following lemma.

[9] Rigorously, the sum in right hand side is an integral when $\mathit{\Delta}$ is not a discrete distribution. In this case, the left hand side should be applied to a measurable subset rather than the singleton.

Lemma 3. *Let $\mathcal{X}$ be a finite mutable type and $\mathcal{Z}$ be a non-empty object in* Corr. *For every joint-distribution $J{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$, there exists $\mathit{\Pi}{:}\,\mathbb{D}\mathcal{X}^2$ and $Z{:}\,\mathcal{X}{\rightarrow\!\!\triangleright}\mathcal{Z}$ such that*

$$\mathbb{F}_{\mathcal{X}}(Z)(\mathit{\Pi}) \quad = \quad J$$

Proof. One way to decompose $J{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$ is through the marginal-diagonal-conditional decomposition [10]. *Let $\mathit{\Pi}{:}\,\mathbb{D}\mathcal{X}^2$ be the diagonal matrix such that*

$$\mathit{\Pi}_{xx'} = \begin{cases} \sum_z J_{xz} & \textit{if } x = x' \\ 0 & \textit{otherwise} \end{cases} ,$$

and $Z{:}\,\mathcal{X}{\rightarrow\!\!\triangleright}\mathcal{Z}$ such that

$$Z_{xz} = \begin{cases} J_{xz}/\mathit{\Pi}_{xx} & \textit{if } \mathit{\Pi}_{xx} > 0 \\ 1/|\mathcal{Z}| & \textit{otherwise} \end{cases} .$$

Since $\mathit{\Pi}$ is a diagonal matrix, we have

$$\mathit{\Pi}\rangle Z = \mathit{\Pi}_{xx} Z_{xz} = \sum_{x'} \mathit{\Pi}_{xx'} Z_{x'z} = \mathbb{F}_{\mathcal{X}}(Z)(\mathit{\Pi}) \ .$$

Moreover, for $\mathit{\Pi}_{xx} > 0$, we have $\mathit{\Pi}_{xx} Z_{xz} = J_{xz}$. If $\mathit{\Pi}_{xx} = 0$ for some x, then $J_{xz} = 0$ for every z. In either cases, $J = \mathbb{F}_{\mathcal{X}}(Z)(\mathit{\Pi})$.

Note that the decomposition in Lem. 3 is not necessarily unique. For instance, the convention $Z_{xz} = 1/|\mathcal{Z}|$ is arbitrary and can be replaced with any collection of numbers such that $\sum_z Z_{xz} = 1$.

The following corollary of Lem. 3 shows that a natural transformation $h{:}\,\mathbb{F}_{\mathcal{X}}{\Rightarrow}\mathbb{D}\mathbb{F}_{\mathcal{X}}$ is fully characterised by its realisation $h^{\mathcal{X}}{:}\,\mathbb{F}_{\mathcal{X}}(\mathcal{X}){\rightarrow}\mathbb{D}\mathbb{F}_{\mathcal{X}}(\mathcal{X})$.

Corollary 1. *$h^{\mathcal{Z}}$ is determined by $h^{\mathcal{X}}$ i.e.*

$$h^{\mathcal{Z}}.J = \mathbb{D}\mathbb{F}_{\mathcal{X}}(Z) \circ h^{\mathcal{X}}(\mathit{\Pi})$$

for every $J{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$ where $Z, \mathit{\Pi}$ is a decomposition of J.

Proof. Lem. 3 and Defn. 3.

This corollary has practical importance as it allows us to focus any formal analysis of a program H on the arrow $[\![H]\!]^{\mathcal{X}}$ without any thought about the possibly unknown collateral type $\mathcal{Z}$. Thus, we do not need to exhaust all possible collateral types to know the collateral effect of executing H. We only need to compute the leaked information regarding the initial and final values of the mutable secret with type $\mathcal{X}$ [10].

In particular, if two programs H, K leak the exact same amount of information about the initial and final values of the secret with type $\mathcal{X}$, then $[\![H]\!]^{\mathcal{Z}}$ and $[\![K]\!]^{\mathcal{Z}}$ leak the exact same amount of information about the collateral type $\mathcal{Z}$. For instance, once A understands that she needs to account for *both* the initial and final value of `pw`, she realises that the collateral secret `npw` is not safe anymore well before she creates the new password.

Let us now explore a few special cases which reinforce the fact that our natural transformation semantics fully captures the notion of collateral information leakage.

3.1 Z is a null channel: $\mathcal{Z}_2 = \{*\}$

This first case explains that the closed semantics of a *HMM* in [9] is a particular instance of our natural transformation semantics. We obtain this by replacing $\mathcal{Z}_2$ with the singleton set $\{*\}$.

An arrow $Z{:}\,\mathcal{Z}_1 {\rightarrowtriangle} \{*\}$ is equivalent to the **null channel** on $\mathcal{Z}_1$ (i.e. single column filled with ones). Defn. 3 reduces to

$$\begin{tikzcd}
\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_1) \arrow[r, "h^{\mathcal{Z}}"] \arrow[d, "\mathbb{F}_{\mathcal{X}}(Z)"'] & \mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z}_1) \arrow[d, "\mathbb{D}\mathbb{F}_{\mathcal{X}}(Z)"] \\
\mathbb{D}\mathcal{X} \arrow[r, "h^*"'] & \mathbb{D}^2\mathcal{X}
\end{tikzcd}$$

where $\mathbb{F}_{\mathcal{X}}(Z){:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_1){\to}\mathbb{D}\mathcal{X}$ such that, for $J{:}\,\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_1)$ we have

$$\mathbb{F}_{\mathcal{X}}(Z)(J)_x = \sum_z J_{xz} \quad ,$$

which is the **$\mathcal{X}$-marginal** of J and becomes our prior, which is then fed into the closed semantics h^*. In other words, h^* *is obtained by forgetting any correlated type i.e. working with the $\mathcal{X}$-marginals of priors and posteriors.*

3.2 Z selects a collateral prior: $\mathcal{Z}_1 = \{*\}$

Our second case shows that the closed semantics can only express independent correlations. Two independent secrets are expressed as a joint distribution in $\mathbb{D}(\mathcal{X}{\times}\mathcal{Z})$ that is obtained as the product of marginals on $\mathcal{X}$ and $\mathcal{Z}$. We obtain this by setting $\mathcal{Z}_1$ to be the singleton set $\{*\}$.

An arrow $Z{:}\,\{*\}{\rightarrowtriangle}\mathcal{Z}_2$ is identified with a distribution in $\mathbb{D}\mathcal{Z}_2$. In this case, Defn. 3 reduces to

$$\begin{tikzcd}
\mathbb{D}\mathcal{X} \arrow[r, "h^*"] \arrow[d, "\mathbb{F}_{\mathcal{X}}(Z)"'] & \mathbb{D}^2\mathcal{X} \arrow[d, "\mathbb{D}\mathbb{F}_{\mathcal{X}}(Z)"] \\
\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_2) \arrow[r, "h^{\mathcal{Z}}"'] & \mathbb{D}^2(\mathcal{X}{\times}\mathcal{Z}_2)
\end{tikzcd}$$

where $\mathbb{F}_{\mathcal{X}}(Z){:}\,\mathbb{D}\mathcal{X}{\to}\mathbb{D}(\mathcal{X}{\times}\mathcal{Z}_2)$ such that, for $\pi{:}\,\mathbb{D}\mathcal{X}$ we have

$$\mathbb{F}_{\mathcal{X}}(Z)(\pi)_{xz} = \pi_x Z_z \quad . \tag{9}$$

So what does this mean? Well, we know that the denotation h^* cannot express the information flow of variables correlated to the secret with type $\mathcal{X}$ [10]. However, Eqn. (9) says that $\mathbb{F}_{\mathcal{X}}(Z)$ simply constructs the independent product of Z and a prior π, and similarly for $\mathbb{D}\mathbb{F}_{\mathcal{X}}(Z)$. Thus the above commutative diagram says that, when $\mathcal{Z}_2$ is independent of $\mathcal{X}$, we can construct the joint distributions and encapsulating hypers by simply multiplying the marginals pointwise. This is the standard definition of a joint distribution expressing independence of the column and row random variables. This is an expected sanity check.

In other words, the non-compositional denotation h^* is also a special case of the compositional semantics where every other variable not specified in the underlying program is independent of the secret of type $\mathcal{X}$. In this case, and this case only, $h^{\mathcal{Z}}$ can be expressed via h^*.

3.3 Collateral only: $\mathcal{X} = \{*\}$

In this case, we will always know the value of the mutable secret variable (which is $*$, of course) and the uncertainty about any correlated secret should remain the same pre- and post execution of a program h.

Let us make that intuition precise. Let $Z\colon \mathcal{Z}_1 \rightarrowtriangle \mathcal{Z}_2$. We have $\mathbb{F}_*(Z)\colon \mathbb{D}\mathcal{Z}_1 \to \mathbb{D}\mathcal{Z}_2$ (up to the isomorphism $\{*\}\times\mathcal{Z} \simeq \mathcal{Z}$) where, for every $\pi\colon \mathbb{D}\mathcal{Z}_1$

$$\mathbb{F}_*(Z)(\pi) = \sum_z \pi_z Z_{z,z'} \ .$$

That h is natural gives us

$$\begin{tikzcd}
\mathbb{D}\mathcal{Z}_1 \arrow[r, "h^{\mathcal{Z}_1}"] \arrow[d, "\mathbb{F}_*(Z)"'] & \mathbb{D}^2\mathcal{Z}_1 \arrow[d, "\mathbb{D}\mathbb{F}_*(Z)"] \\
\mathbb{D}\mathcal{Z}_2 \arrow[r, "h^{\mathcal{Z}_2}"'] & \mathbb{D}^2\mathcal{Z}_2
\end{tikzcd}$$

But what is $h^{\mathcal{Z}_1}$? Well, Corollary 1 tells us that $h^{\mathcal{Z}}$ is determined by $h^{\mathcal{X}}$ for every $\mathcal{Z}$. But since $\mathcal{X} = \{*\}$, we deduce that $h^*\colon \mathbb{D}\{*\} \to \mathbb{D}^2\{*\}$. The only arrow of this type is $\mathbb{D}\eta$ where η maps the point $*$ to the point distribution η_*. That is, $h^* = \mathbb{D}\eta$ which means that it is the **Skip** program.

Finally, for every $\pi\colon \mathbb{D}\mathcal{Z}_1$ (actually $\mathbb{D}(\{*\}\times\mathcal{Z}_1)$), the only decomposition given by Lem 3 is given by the joint distribution matrix $(1) \in \mathbb{D}\{*\}^2$ and channel $Z\colon \{*\} \rightarrowtriangle \mathcal{Z}_1$ with $Z_* = \pi$. So

$$h^{\mathcal{Z}}.\pi = \mathbb{D}\mathbb{F}_*(Z) \circ h^*((1)) = \mathbb{D}\mathbb{F}_*(Z)(\eta_{(1)}) = \eta_{\mathbb{F}_*(Z)((1))} = \eta_\pi \ .$$

Thus $h^{\mathcal{Z}}$ is also **Skip** as expected.

4 Conclusion

We have shown that programs computing with secrets can be interpreted as natural transformations acting on a category Corr of correlations. This semantic space is compositional. For each collateral type $\mathcal{Z}$, the natural transformation $[\![P]\!]$, for an arbitrary program P, can be instantiated on $\mathcal{Z}$ by working with the arrow $[\![P]\!]^{\mathcal{Z}}$. Moreover, the natural transformation $[\![P]\!]$ is completely characterised by the component $[\![P]\!]^{\mathcal{X}}$. This implies that, instead of trying to capture all possible correlations (known and unknown), one can simply look at programs as constructing correlations between initial and final secret values. Thus a fresh correlation can be attached to the output using the preserved initial secret values. This is captured abstractly by the notion of polymorphism.

Our main goal was to provide a categorical compositional semantics that can be used as a foundation for further frameworks suitable for program with secrets. In particular, we did not elaborate on the explicit construction of a refinement ordering on the space of natural transformation nor did we study any notion of convergence of a sequence of natural transformations. Though both of these items are important to obtain a sound logical framework for programs, we leave them for future work.

References

1. Jiří Adámek, Horst Herrlich, and George Strecker. *Abstract and Concrete Categories.* Wiley-Interscience, New York, NY, USA, 1990.
2. M. Alvim, M. Andrés, and C. Palamidessi. Probabilistic information flow. In *Proc. of the 25th IEEE Symposium on Logic in Computer Science*, pages 314–321, 2010.
3. M. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith. Additive and multiplicative notions of leakage, and their capacities. In *Proc. of the IEEE 27th Computer Security Foundations Symposium*, pages 308–322, 2014.
4. M. Alvim, K. Chatzikokolakis, C. Palamidessi, and G. Smith. Measuring information leakage using generalized gain functions. In *Proc. of the 25th IEEE Computer Security Foundations Symposium*, pages 265–279, June 2012.
5. C. Braun, K. Chatzikokolakis, and C. Palamidessi. Quantitative notions of leakage for one-try attacks. In *Proc. of the 25th International Conference on Mathematical Foundations of Programming Semantics*, pages 75–91, 2009.
6. M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer, 1981.
7. Chris Karlof and David Wagner. Hidden Markov Model cryptanalysis. In *Cryptographic Hardware and Embedded Systems - CHES 2003, 5th International Workshop, Cologne, Germany, September 8-10, 2003, Proceedings*, pages 17–34, 2003.
8. A. McIver, L. Meinicke, and C. Morgan. Compositional closure for Bayes Risk in probabilistic noninterference. In *Proc. of the 37th international colloquium conference on Automata, languages and programming: Part II*, pages 223–235, 2010.
9. A. McIver, C. Morgan, and T. Rabehaja. Abstract Hidden Markov Models: a monadic account of quantitative information flow. In *Proc. of the 30th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 597–608, 2015.
10. A. McIver, C. Morgan, and T. Rabehaja. Algebra for quantitative information flow. In *Proc. of the 16th International Conference on Relational and Algebraic Methods in Computer Science*, pages 3–23, 2017.
11. A. McIver, C. Morgan, T. Rabehaja, and N. Bordenabe. Reasoning about distributed secrets. In *Proc. of the 37th IFIP WG 6.1 International Conference on Formal Techniques for Distributed Objects, Components, and Systems*, pages 156–170, 2017.

12. A. McIver, C. Morgan, G. Smith, B. Espinoza, and L. Meinicke. Abstract channels and their robust information-leakage ordering. In *Proc. of the Third International Conference on Principles of Security and Trust*, Lecture Notes in Computer Science, pages 83–102, 2014.
13. K. R. Parthasarathy. *Probability Measures on Metric Spaces.* Academic Press, 1967.
14. G. Smith. On the foundations of quantitative information flow. In *Proc. of the 12th International Conference on Foundations of Software Science and Computational Structures*, Lecture Notes in Computer Science, pages 288–302, 2009.
15. G. Smith. Quantifying information flow using min-entropy. In *Proc. of the 8th International Conference on Quantitative Evaluation of SysTems*, pages 159–167, 2011.
16. F. van Breugel. The metric monad for probabilistic nondeterminism. Draft available at *http://www.cse.yorku.ca/∼franck/research/drafts/monad.pdf*, 2005.

A Compositional sequential composition

The sequential composition (8) assumes that the mutable secrets of H and K are represented by the exact same variables ranging in $\mathcal{X}$. A slightly more general version of that result is to assume that H and K may share a hidden variable ranging in $\mathcal{X}$ and that they also have other hidden variables specific to each *HMM*. In this case, if H contains another hidden secret with type $\mathcal{X}_1$ then this secret must be considered as a collateral secret in K — thus K leaks collateral information about it, but does not update that secret. More precisely, we have the following lemma.

Lemma 4. *Let* $H\colon(\mathcal{X}\times\mathcal{X}_1)\mathrel{-\!\!\triangleright}\mathcal{Y}_1\times(\mathcal{X}\times\mathcal{X}_1)$ *with* $K\colon(\mathcal{X}\times\mathcal{X}_2)\mathrel{-\!\!\triangleright}\mathcal{Y}_2\times(\mathcal{X}\times\mathcal{X}_2)$ *be two HMM matrices. We have*

$$[\![H;K]\!]^{\mathcal{Z}} = [\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}};[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}\ . \tag{10}$$

where the composition $H;K$ *is given by Eqn.* (2) *and the composition operator* (;) *on the right takes the isomorphism between* $\mathcal{X}_1\times\mathcal{X}_2$ *and* $\mathcal{X}_2\times\mathcal{X}_1$ *into account.*

Proof. Firstly, we have the following types

- $[\![H;K]\!]^{\mathcal{Z}}\colon\mathbb{D}(\mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Z})\rightarrow\mathbb{D}^2(\mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Z})$
- $[\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}}\colon\mathbb{D}(\mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Z})\rightarrow\mathbb{D}^2(\mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Z})$
- $[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}\colon\mathbb{D}(\mathcal{X}\times\mathcal{X}_2\times\mathcal{X}_1\times\mathcal{Z})\rightarrow\mathbb{D}^2(\mathcal{X}\times\mathcal{X}_2\times\mathcal{X}_1\times\mathcal{Z})$

Let us assume that we have applied the isomorphism that permutes the positions of the $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *in the domain of* $[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}$ *so that the right hand side composition is well defined and is given by the sequential composition in the Giry monad [6]. This type matching function is implicitly embedded into the right hand side composition in* (10).

Let $\mathit{\Pi}\colon\mathbb{D}(\mathcal{X}\times\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{Z})$, *Eqns.* (1) *and* (2) *give*

$$(\mathit{\Pi}\rangle(H;K))_{y_1y_2,x''x_1'x_2'z} = \sum_{xx_1x_2}\mathit{\Pi}_{xx_1x_2z}\sum_{x'}H_{xx_1y_1x'x_1'}K_{x'x_2y_2x''x_2'} \tag{11}$$

On the other hand, $[\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}};[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}} = \mathsf{avg}\circ\mathbb{D}[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}\circ[\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}}$ *uses the Kleisli lifting. Let us consider an inner* δ *of* $[\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}};[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}.\mathit{\Pi}$*. There exists an inner* α *of* $[\![H]\!]^{\mathcal{X}_2\times\mathcal{Z}}.\mathit{\Pi}$ *such that* δ *is an inner of* $[\![K]\!]^{\mathcal{X}_1\times\mathcal{Z}}.\alpha$*. That is, there exist two observations* y *and* y' *such that*

$$\alpha^{y_1}_{x'x'_1x_2z} = \frac{\sum_{xx_1} \mathit{\Pi}_{xx_1x_2z} H_{xx_1y_1x'x'_1}}{\overline{\alpha}^{y_1}}$$

where $\overline{\alpha}^{y_1} = \sum_{x'x'_1x_2z}\left(\sum_{xx_1}\mathit{\Pi}_{xx_1x_2z}H_{xx_1y_1x'x'_1}\right)$*, and*

$$\delta^{y_1y_2}_{x''x'_1x'_2z} = \frac{\sum_{x'x_2}\alpha^{y_1}_{x'x'_1x_2z}K_{x'x_2y_2x''x'_2}}{\overline{\delta}^{y_1y_2}}$$

where

$$\overline{\delta}^{y_1y_2} = \sum_{x''x'_1x'_2z}\left(\sum_{x'x_2}\alpha^{y_1}_{x'x'_1x_2z}K_{x'x_2y_2x''x'_2}\right)$$

By substituting α *into the expression of* δ *and simplifying* $\overline{\alpha}^{y_1}$*, we have*

$$\begin{aligned}\delta^{y_1y_2}_{x''x'_1x'_2z} &= \frac{\sum_{x'x_2}\left(\sum_{xx_1}\mathit{\Pi}_{xx_1x_2z}H_{xx_1y_1x'x'_1}\right)K_{x'x_2y_2x''x'_2}}{\sum_{x''x'_1x'_2z}\left(\sum_{x'x_2}\left(\sum_{xx_1}\mathit{\Pi}_{xx_1x_2z}H_{xx_1y_1x'x'_1}\right)K_{x'x_2y_2x''x'_2}\right)}\\ &= \frac{\sum_{xx_1x_2}\mathit{\Pi}_{xx_1x_2z}\sum_{x'}H_{xx_1y_1x'x'_1}K_{x'x_2y_2x''x'_2}}{\sum_{x''x'_1x'_2z}\left(\sum_{xx_1x_2}\mathit{\Pi}_{xx_1x_2z}\sum_{x'}H_{xx_1y_1x'x'_1}K_{x'x_2y_2x''x'_2}\right)}\end{aligned}$$

The inner $\delta^{y_1y_2}$ *corresponds exactly to a normalized* (y_1,y_2)*-column of* $\mathit{\Pi}\rangle(H;K)$ *as per Eqn.* (11).

Lem. 4 is a straightforward generalisation our sequential composition for systems independent of any other collateral type [8,9] as well as our compositional, but single secret, semantics [10,11]. That is, the closed sequential composition is obtained by setting $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Z} = \{*\}$ while the context aware version is obtained by setting $\mathcal{X}_1 = \mathcal{X}_2 = \{*\}$. Lem. 4 is slightly more general than the composition we define in [10] because it allows the declaration of new variables "on the go". For instance, if we set $\mathcal{X}_1 = \{*\}$, then a new secret variable of type $\mathcal{X}_2$ that is declared in K does not change in H. The lifted map $[\![H]\!]^{\mathcal{X}_2}$ is aware of this upcoming new secret and accounts for the correlation between the new variable and current secret of type $\mathcal{X}$.