Toward a Mathematical Foundation for Information Flow Security

James W. Gray, III
Center for Secure Information Technology, Code 5543
Naval Research Lab, Washington, DC 20375

Abstract

We describe a general purpose, probabilistic state machine model which can be used to model a large class of nondeterministic (as well as deterministic) computer systems. We develop the necessary probability theory to rigorously state and prove probabilistic properties of modeled systems. Then we give a definition of information flow security making use of this formalism. Intuitively, *information flow security* is the aspect of computer security concerned with how information is permitted to flow through a computer system. We prove that our definition of information flow security implies an information theoretic definition. Finally, we give a verification condition for information flow security and prove that it implies our definition of information flow security.

1 Introduction

Intuitively, *information flow security* is the aspect of computer security concerned with how information is permitted to flow through a computer system. Recently, there has been a general belief that information flow security models should be founded on—or at least justified by—the field of information theory. This is primarily due to the fact that nondeterministic systems may exhibit probabilistic covert channels which are not ruled out by standard computer security models (see [15] for a discussion of this). To date, there have been several efforts relating information theory to computer security [6, 7, 9, 15].

McLean [6] gives a very general treatment of information flow security models. However, since his work is intended to provide a means for evaluating security models (rather than a means for evaluating real systems), it is difficult to see how to apply it to real systems. For example, his definition of security (which he calls FM) is given in the form of an equation involving conditional probabilities, where his probability space includes user inputs. Since the probability of a given user input occurring is typically unknown, it is unclear (and McLean makes no claims either way) whether his definition is generally verifiable.

On the other hand, the efforts of Millen [7], Moskowitz [9], and Wittbold and Johnson [15] have applied standard information theoretic concepts to concrete examples. However, their examples tend to be so simple that it is unclear whether their work can be applied to more complex systems. In particular, none of their example systems has a large, general purpose memory (such as a database system or a file system has) . The example that comes closest is from [7] where a system with a one bit memory is analyzed.

In this paper, we attempt to bridge this gap between the general, but abstract (i.e., McLean) and the concrete, but limited (i.e., Millen, Moskowitz, and Wittbold and Johnson). Although the present paper is mainly concerned with the special case where no leakage of classified information is tolerated (i.e., the capacity of all covert channels must be zero), the framework is intended to be general enough to deal with the more general case where the tolerance for information leakage is given by a rate such as n bits per second. We hope to pursue this more general problem in future work.

We begin by describing a general purpose, probabilistic state machine model. Our model has a finite—but potentially very large—set of states (as do real computer systems) and a set of communication channels which provide the only interface to the external environment. Our system transition function T, is general enough to describe internal (i.e., internal to the system being considered) probabilistic choices. This model can be used to model a large class of nondeterministic (as well as deterministic) computer systems.

Then, in Section 3, we develop the probability theory needed to state and prove probabilistic properties (such as information flow security) of computer systems. We separate internal probabilistic behavior (i.e., internal probabilistic choices) from external probabilistic behavior and then make the relationship between the two explicit. In this way, we can model the system's probabilistic behavior and the environment's probabilistic behavior separately and then state and prove properties involving joint probabilities of the two. In order to avoid ambiguity and careless errors, we define our probability space (actually an infinite set of probability spaces) and probability measure (again an infinite set of them) rigorously and then prove (in Appendix A) that what we call our "probability measure" is in fact a probability measure.

In Section 4 we present our definition of information flow security and give a theorem that lends support to its strength. In Section 5 we state and prove a relationship between our definition of information flow security and an information theoretic definition of channel capacity. And finally, in Section 6 we give a verification condition for information flow security. A proof that

CH2986-8/91/0000/0021$01.00 © 1991 IEEE

our verification condition implies our definition of information flow security is given in Appendix B.

Throughout this paper, we partition the set of system communication channels into two sets of channels H and L representing the channels connected to high processes (or users) and the channels connected to low processes (or users), respectively. The motivation for this work (as in [2, 4, 7, 9, 14, 15]) is the need to prevent information flow from high inputs to low outputs.

2 System Model

Our system model is a nonstandard finite state machine model. There is a finite set of inputs, a finite set of outputs, a finite set of states, and an initial state. The nonstandard features are a finite set of communication channels and a *probabilistic* transition function. We also have a rather unusual interpretation of the inputs and outputs. We discuss these ideas in more detail below.

Our model is similar to Millen's *synchronous state machine* model [8] in style and motivation. His treatment of time and his use of channels, inputs, and outputs are in essence identical to ours. The two models differ in some minor details but can be used to model the same class of computer systems. The only difference between the two models that is essential to our work is that our model has a probabilistic transition function rather than merely a nondeterministic transition function. This additional detail about the system transition function will allow us to reason about the probabilistic behavior of the system.

Regarding our interpretation of outputs, we think of a finite set of signals that the system can produce and that are discernable by the environment. We call these signals "outputs" but the set of outputs should not be interpreted as containing only outputs in the conventional sense. For example, we will need a distinguished output **null**, which can be thought of as the system performing no action. (Note that "no action" can be used as a signal to the environment.) Also, if for example the external environment can discern when the system accepts an input from a bounded buffer (as in an example from [5]), then the action of accepting an input would be considered a system output. On each transition of the system and on each communication channel, the system produces one output. Note that it is possible (and maybe even likely) that most outputs on most channels will be **null** (i.e., most systems do not produce interesting outputs on each communication channel on every transition).

We interpret the set of "inputs" in the same way as for outputs. In particular, any signal (including **null**—the lack of an action) that the environment can produce and that is discernable by the system, is considered an "input". Again, on each transition of the system and on each channel of the system, the environment produces one input.

We interpret the system transitions as occurring once for each tick of the system's internal clock. We assume that the system's internal clock runs independently of any other processing by the system. In this way, the transitions of the system mark the passing of time and we will use this fact to deal with covert timing channels.

Systems that are implemented by probabilistic algorithms (an increasingly popular practice; see for example [10, 11]) may make internal, probabilistic choices between various events. We assume (as in [3]) that these internal probabilistic choices are made using a pseudo-random number generator and that all choices are made independently of one another. To model internal probabilistic choices, the system transition function gives for any time t, the probability of the system producing an output vector b_t (Note: an output vector contains one output for each communication channel) and moving to state s_t, given that the system was in state s_{t-1} and the environment produced the input vector a_{t-1} at time $t-1$.

Now we introduce a little notation to set out our system model precisely.

Notation: Let X and Y be finite sets. Then $X[Y]$ denotes the set of all one-dimensional arrays of X indexed by Y (i.e., each $a \in X[Y]$ is an array with length equal to the cardinality of Y where each item in the array is an element of X). We will sometimes refer to arrays as "vectors" or "sequences" depending on the context. For example, suppose O is the set of outputs and C is the set of communication channels. Then $O[C]$ is the set of all output vectors. As another example, suppose S is the set of states. Then $S[1..t]$ is the set of all state sequences from time 1 through time t.

We will also need to make use of two-dimensional arrays (i.e., matrices). We will use the straightforward generalization as follows; suppose X, Y, and Z are finite sets. Then $X[Y, Z]$ is the set of all two-dimensional arrays of X indexed by $Y \times Z$.

We will pick out a particular element of an array using elements from the index set. For example, if $b \in O[C]$ is an output vector and $c_1 \in C$ is a particular channel, then $b[c_1]$ is the single output for c_1. Similarly, if $\gamma \in S[1..t]$ is a state sequence, then $\gamma[1]$ will mean the first state in the sequence, $\gamma[2]$ will mean the second state in the sequence, etc..

For any set X, we denote the power set of X (i.e., the set of all subsets) as $\mathcal{P}(X)$.

We denote the set of all positive integers by $\mathbf{N}^+$.

Given two real numbers x and y, we denote the closed interval from x to y (i.e., the set of all real numbers between (and including) x and y) by $[x, y]$. In particular, we will be making frequent use of the closed interval $[0, 1]$.

For any set of channels $\lambda \subseteq C$, we will use $\mathbf{null}_\lambda$ to denote the constant vector consisting of **null** inputs (or outputs depending on context) on all channels in λ.

Now we can define our system model.

Definition 2.1 *A system Σ is given by a six tuple $(C, I, O, S, s0, T)$ where*

C is a finite set of communication channels;
I is a finite set of input signals;
O is a finite set of output signals;
S is a finite set of internal system states;
$s0$ is the initial state; and

$T : S \times I[C] \times S \times O[C] \rightarrow [0,1]$ is the function that describes how the system moves from state to state while engaging in input and output events. T should be interpreted as follows: given that the system is in state $s \in S$ at some point in time, and that the input at that time, on all communication channels is the vector $a \in I[C]$, then for any state $s' \in S$ and for any output vector $b \in O[C]$, the probability of the system producing b and transitioning to state s' is $T(s,a,s',b)$. The input vector for the system's first transition (i.e., the transition that takes place at time 0), is assumed to be **null***C (i.e., the environment cannot produce any input before time 1).*

To support the above interpretation of T, we require of all systems $\Sigma = (C, I, O, S, T)$ that for all $s \in S$, and for all $a \in I[C]$, the function $P_{s,a} : \mathcal{P}(S \times O[C]) \rightarrow [0,1]$ defined by

$$P_{s,a}(x) \equiv \sum_{(s',b) \in x} T(s,a,s',b)$$

to be a probability measure.

This system model is general enough to model a very large class of computer systems. For example, depending on how we interpret the communication channels, we can model a single process executing (among other processes) on a multitasking operating system (where the communication channels would be provided by the operating system in the form of an interprocess communication mechanism) or an actual piece of hardware (where the communication channels would be physical wires).

For our purposes (i.e., information flow security), we will assume that the communication channels provide the only interface to the external environment (i.e., a system cannot communicate with the environment in any other way). In the case of a process executing on a multitasking operating system, this assumption amounts to the assumption that the operating system acts as a separation kernel [12].

3 Probability Theory

In this section we develop an infinite set of probability measures to be used in stating and proving probabilistic properties (e.g., in our case, information flow security) of computer systems. We define a probability measure P_t for each time t, where the sample space for P_t is the set of all possible "executions" of the system up to time t. P_t is determined by the particular system under consideration and the probabilistic behavior of the environment. This will all be set out precisely below. We prove in Appendix A that for any time t, P_t is a probability measure.

Definition 3.1 *Given a finite set of output signals O, containing at least the distinguished element* **null***, for any set of channels λ, and any positive integer t, let $O_{\lambda,t} = O[\lambda, 1..t]$ be the set of all possible system output histories on channels in λ during times 1 through t. Note: some of these output histories may occur with probability 0 (i.e., be essentially impossible) for a given system.*

For convenience, we adopt the convention that for any set of channels λ, $O_{\lambda,0}$ is the empty history (i.e., the history of length zero) of the channels in λ. The empty history represents the history of the system before it does anything.

Definition 3.2 *Given a finite set of input signals I containing at least the distinguished element* **null** *(here meaning that the environment performed no action), a set of channels λ, and a positive integer t, let $I_{\lambda,t} = I[\lambda, 1..t]$ be the set of all possible environment input histories on channels in λ during times 1 through t.*

As for output histories, we will write $I_{\lambda,0}$ for the empty input history on the channels in λ. The empty input history represents the history of the environment before the environment has a chance to do anything.

Definition 3.3 *Given a finite set of internal system states S and a positive integer t, let $S_t = S[1..t]$ be the set of all possible histories of internal system states during times 1 through t.*

Definition 3.4 *For every positive integer t, let $\Omega_t = I_{C,t} \times O_{C,t} \times S_t$ be a sample space for outcomes of the system (i.e., an outcome is a triple (α, β, γ) where α is an input history, β is an output history, and γ is an internal state history) up to time t.*

Note that since O, I, and S are finite, for any t, we have that $O_{C,t}$, $I_{C,t}$, and S_t are finite; and therefore, Ω_t is finite. Therefore, $\mathcal{P}(\Omega_t)$ possesses the necessary properties for it to be a suitable event space for our probability measure (i.e., $\langle \mathcal{P}(\Omega_t), \cup, \cap \rangle$ forms a σ-field).

We make the following assumption, which essentially constitutes the assumption that the only information about the system and its environment that is directly accessible to the high (low, resp.) environment is the inputs and outputs that have previously occurred on the high (low, resp.) channels. I.e., if the high environment obtains information about the low environment, it must obtain it indirectly through its interaction with the system; similarly if the low environment obtains information about the high environment, it must obtain it indirectly through its interaction with the system.

Assumption 3.1 *Let H and L be disjoint sets of channels representing the channels connected to high processes and the channels connected to low processes, respectively. For any time $t \in \mathbf{N}^+$, any input history $\alpha \in I_{H,t-1}$, and any output history $\beta \in O_{H,t-1}$,*

$H_{t,\alpha,\beta} : \mathcal{P}(I[H]) \rightarrow [0,1]$ is a probability measure that completely describes the probabilistic behavior of the environment external to H at time t, given the input history α and the output history β, prior to time t. We call the set of $H_{t,\alpha,\beta}$'s (i.e., the complete description of the behavior of the environment external to H) **H**.

For any time $t \in \mathbf{N}^+$, any input history $\alpha \in I_{L,t-1}$, and any output history $\beta \in O_{L,t-1}$, we make the analogous assumption for $L_{t,\beta,\sigma} : \mathcal{P}(I[L]) \rightarrow [0,1]$ and the behavior external to L. Also, we call the set of $L_{t,\alpha,\beta}$'s **L**.

Since the only element of $I_{H,0}$ is $\mathbf{null}_H$ and the only element of $O_{H,0}$ is $\mathbf{null}_H$ we will use $H_1(s)$ as a shorthand for $H_{1,\mathbf{null}_H,\mathbf{null}_H}$. Similarly, we will use $L_1(s)$ as a shorthand for $L_{1,\mathbf{null}_L,\mathbf{null}_L}$.

Note that the above assumption allows for the possibility that the environment external to H (L, resp.) has memory of what it has already done, as well as memory of what the system has already done on channels in H (L, resp.). Therefore, this assumption is general enough to allow for the possibility that the environment is acting according to some strategy that involves feedback from the system (as in an example from [15]). This assumption also allows for multiple processes external to H (and likewise for L) to be working cooperatively with shared memory.

Now we introduce a little more notation.

Notation: Given an input vector $a \in I[C]$ (or a sequence of input vectors $\alpha \in I_{C,t}$, resp.), and a set of channels $\lambda \subseteq C$, let $\pi_\lambda(a)$ (or $\pi_\lambda(\alpha)$, resp.) be the projection of a (or α, resp.) onto the channels in λ.

Given a sequence α of length t, for any $i \leq t$, let $\alpha_{\rightarrow i}$ be the subsequence of α from the first element up through (and including) the i^{th} element. Further, we will use $\zeta(\alpha)$ (read "front of α") as shorthand for $\alpha_{\rightarrow t-1}$.

We can now define a probability measure that allows us to reason about the probabilistic behavior of both the system and its environment. The following definition is consistent with the above assumptions that (1) the probabilistic behavior of the high environment is described completely by **H**; (2) the probabilistic behavior of the low environment is described completely by **L**; and (3) the probabilistic behavior of the system is described completely by T.

We define $P_{\Sigma,\mathbf{H},\mathbf{L},t} : \mathcal{P}(\Omega_t) \rightarrow [0,1]$ inductively, as follows.

Definition 3.5 *Given a system Σ, and high and low environment behaviors* **H** *and* **L**, *resp., for all $\omega \in \mathcal{P}(\Omega_1)$,*

$$P_{\Sigma,\mathbf{H},\mathbf{L},1}(\omega) \equiv \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \qquad \textit{if } (\alpha,\beta,\gamma) \in \omega; \\ 0, \quad \textit{otherwise.} \end{cases}$$

For all $t \geq 2$ and for all $\omega \in \mathcal{P}(\Omega_t)$,

$$P_{\Sigma,\mathbf{H},\mathbf{L},t}(\omega) \equiv \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \begin{cases} P_{\Sigma,\mathbf{H},\mathbf{L},t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t]), \\ \qquad \textit{if } (\alpha,\beta,\gamma) \in \omega; \\ 0, \quad \textit{otherwise.} \end{cases}$$

Thus, $P_{\Sigma,\mathbf{H},\mathbf{L},t}(\omega)$ is the sum over all system executions $(\alpha,\beta,\gamma) \in \omega$ of the probability of (α,β,γ) occurring, where the probability of (α,β,γ) occurring is obtained by multiplying the following together: the probability of the environment producing $\alpha[1]$ (as given by **H** and **L**), the probability of the system producing $\beta[1]$ and $\gamma[1]$ (as given by T), the probability of the environment producing $\alpha[2]$ (again given by **H** and **L**), the probability of the system producing $\beta[2]$ and $\gamma[2]$ (again given by T), ... the probability of the environment producing $\alpha[t]$ (again given by **H** and **L**), and the probability of the system producing $\beta[t]$ and $\gamma[t]$ (again given by T).

To make our expressions more compact, we will use P_t or $P_{\mathbf{H},t}$ as a shorthand for $P_{\Sigma,\mathbf{H},\mathbf{L},t}$ when no ambiguity results.

We can now prove that P_t is a probability measure.

Theorem 3.1 *For any system Σ, high and low environment behaviors,* **H**, **L**, *and time $t \in \mathbf{N}^+$, $P_{\Sigma,\mathbf{H},\mathbf{L},t}$ is a probability measure.*

Proof: see appendix A. □

Although proving this result does not prove that $P_{\Sigma,\mathbf{H},\mathbf{L},t}$ is the probability measure corresponding to our intuition, it may give us a some confidence in the definition. We can gain additional confidence in the definition by proving such facts as $P_t(\gamma[t] \cap \beta[t] \mid \gamma[t-1] \cap \alpha[t-1]) = T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t])$.

4 Definition of Information Flow Security

In this section, we give a definition of "information flow security" in terms of the probability theory developed above.

Recall from probability theory that an *event* is a set of outcomes (for our purposes, an outcome is a system history triple $(\alpha,\beta,\gamma) \in \Omega_t$). Suppose we have an event $x \in \mathcal{P}(\Omega_t)$. We say that x *occurs* on a particular trial of the system if the outcome of the trial (up to time t) is in x. Also, $P_t(x)$ is the probability that on any given trial of the system, the outcome of the trial up to time t will be in x.

There are three particular types of events that we will be making frequent use of. For convenience, we define these three types of events here.

Definition 4.1 *For every positive integer t, and for any set of channels, $\lambda = \{c_1, c_2, \ldots c_n\}$, define In-Seq-Event$_{\lambda,t}$ as follows.*

$$
\begin{array}{l}
\textit{In-Seq-Event}_{\lambda,t} \equiv \\
\quad \{\, s \in \mathcal{P}(\Omega_t) \mid \\
\qquad (\exists\, i_{c_1,1}, i_{c_1,2}, \ldots i_{c_1,t-1}, \\
\qquad\quad i_{c_2,1}, i_{c_2,2}, \ldots i_{c_2,t-1}, \\
\qquad\quad \ldots \\
\qquad\quad i_{c_n,1}, i_{c_n,2}, \ldots i_{c_n,t-1} \in I) \\
\qquad (\forall (\alpha, \beta, \gamma) \in \Omega_t) \\
\qquad\quad [(\alpha, \beta, \gamma) \in s \Longleftrightarrow \\
\qquad\qquad (\forall c \in \lambda)(\forall j \in 1..t-1) \\
\qquad\qquad\quad [\alpha[c, j] = i_{c,j}]]\,\}
\end{array}
$$

Each event in *In-Seq-Event*$_{\lambda,t}$ represents the occurrence of a particular input history (i.e., the $i_{c,j}$'s) on the channels in λ up to and including time $t-1$ (Note: *not* up to time t—this is due to the intended application of this definition).

Note that for any set of channels λ, *In-Seq-Event*$_{\lambda,1}$ is the singleton set $\{\Omega_1\}$. In other words, an event in *In-Seq-Event*$_{\lambda,1}$ represents the occurence of a particular input history on λ up to time 0; the only such event is that "nothing happened before time 1" since nothing *can* happen before time 1 and all outcomes in Ω_1 are members of the event specifying that "nothing happened before time 1".

Note: we call this set "*In-Seq-Event*" because each of its elements is characterized by a unique predicate on the *sequence of inputs* on the channels in λ.

Definition 4.2 *For every positive integer t, and for any set of channels, $\lambda = \{c_1, c_2, \ldots c_n\}$, define Out-Seq-Event$_{\lambda,t}$ as follows.*

$$
\begin{array}{l}
\textit{Out-Seq-Event}_{\lambda,t} \equiv \\
\quad \{\, s \in \mathcal{P}(\Omega_t) \mid \\
\qquad (\exists\, o_{c_1,1}, o_{c_1,2}, \ldots o_{c_1,t-1}, \\
\qquad\quad o_{c_2,1}, o_{c_2,2}, \ldots o_{c_2,t-1}, \\
\qquad\quad \ldots \\
\qquad\quad o_{c_n,1}, o_{c_n,2}, \ldots o_{c_n,t-1} \in O) \\
\qquad (\forall (\alpha, \beta, \gamma) \in \Omega_t) \\
\qquad\quad [(\alpha, \beta, \gamma) \in s \Longleftrightarrow \\
\qquad\qquad (\forall c \in \lambda)(\forall j \in 1..t-1) \\
\qquad\qquad\quad [\beta[c, j] = o_{c,j}]]\,\}
\end{array}
$$

Definition 4.3 *For every positive integer t, and for any set of channels, $\lambda = \{c_1, c_2, \ldots c_n\}$, define Final-Out-Event$_{\lambda,t}$ as follows.*

$$
\begin{array}{l}
\textit{Final-Out-Event}_{\lambda,t} \equiv \\
\quad \{\, s \in \mathcal{P}(\Omega_t) \mid \\
\qquad (\exists o_{c_1}, o_{c_2}, \ldots o_{c_n} \in O) \\
\qquad (\forall (\alpha, \beta, \gamma) \in \Omega_t) \\
\qquad\quad [(\alpha, \beta, \gamma) \in s \Longleftrightarrow \\
\qquad\qquad (\forall c \in \lambda) \\
\qquad\qquad\quad [\beta[c, t] = o_c]]\,\}
\end{array}
$$

Now we can give our definition of information flow security.

Definition 4.4 *Given a system $\Sigma = (C, I, O, S, s0, T)$, a set of high channels H and a set of low channels L such that H and L are disjoint and $H \cup L = C$, we say $H \not\rightarrow L$ (read H does not flow to L) if and only if for any high and low behaviors $\mathbf{H}$ and $\mathbf{L}$, any $t \in \mathbf{N}^+$, any $\alpha_H \in$ In-Seq-Event$_{H,t}$, any $\beta_H \in$ Out-Seq-Event$_{H,t}$, any $\alpha_L \in$ In-Seq-Event$_{L,t}$, any $\beta_L \in$ Out-Seq-Event$_{L,t}$, and any $l_t \in$ Final-Out-Event$_{L,t}$,*

$$
\begin{array}{l}
P_t(\alpha_L \cap \beta_L \cap \alpha_H \cap \beta_H) > 0 \Rightarrow \\
\quad P_t(l_t \mid \alpha_L \cap \beta_L \cap \alpha_H \cap \beta_H) = P_t(l_t \mid \alpha_L \cap \beta_L)
\end{array}
$$

The basic intuition of this definition is that the probability of a low output may depend on previous low events, but not on previous high events. In effect, this prevents low users from gaining any information about the high environment's behavior by observing low outputs.[1] Below, we give a theorem that supports this intuition. Further, in the next section we will show that this intuition is also supported by an information theoretic formulation of information flow.

Wittbold and Johnson define the notion of *nondeducibility on strategies* [15]. We can define nondeducibility on strategies in the terminology of this paper as follows.

Definition 4.5 *Given a system $\Sigma = (C, I, O, S, s0, T)$, a set of high channels H and a set of low channels L such that H and L are disjoint and $H \cup L = C$, Σ is* nondeducible on strategies *if for any time $t \in \mathbf{N}^+$, any $e \in \mathcal{P}(\Omega_t)$ such that e is characterized by a predicate on the inputs and outputs occuring on low channels up to time t (i.e., e is an event that is observable to the low*

[1] Given that (as stated in section 1) our goal is to prevent information flow from high inputs to low outputs, one might think (and in fact for some time the author thought) that the weaker condition $P_t(\alpha_L \cap \beta_L \cap \alpha_H) > 0 \Rightarrow P_t(l_t \mid \alpha_L \cap \beta_L \cap \alpha_H) = P_t(l_t \mid \alpha_L \cap \beta_L)$ would be sufficient. However this is not the case. For example, [15, example 2.3] satisfies this weaker condition but contains a covert channel. Intuitively, the problem with the weaker condition is as follows. The nondeterministic transition function T may be legitimately used to obscure the high input from the low output (e.g., as in an encryption device). However, if the high environment gains knowledge (via its output from the system) of the (nondeterministically generated) value to be used in obscuring its input ahead of time, then it can modify its input value accordingly and successfully transmit information to the low environment. Including the previous history of high outputs in the lefthand conditional probability ensures that internally generated probabilistic behavior that is crucial to the security of the system cannot be observed by the high environment and compensated for in its input.

On the other hand, the definition as stated is too strong. For example, a system that probabilistically generates a random number n (via the transition function T), outputs n to a high channel, and then outputs n to a low channel is intuitively secure (with respect to our requirements) but does not satisfy our definition of security. We hope to address this concern in future work.

environment), any low environment behavior **L**, *and any two high environment behaviors* **H** *and* **H'**,

$$P_{\mathbf{H},t}(e) > 0 \Rightarrow P_{\mathbf{H}',t}(e) > 0$$

Intuitively, nondeducibility on strategies ensures that the high environment's behavior cannot influence the *possibility* of a particular low event occurring. With the mathematical machinery developed in this paper thus far, we can state a stronger, probabilistic version of nondeducibility on strategies that ensures that the high environment's behavior cannot influence the *probability* of a particular low event occurring. We believe that this probabilistic version rules out all probabilistic covert channels in addition to the covert channels ruled out by standard nondeducibility on strategies.

Definition 4.6 *Given a system* $\Sigma = (C, I, O, S, s0, T)$, *a set of high channels* H *and a set of low channels* L *such that* H *and* L *are disjoint and* $H \cup L = C$, Σ *is* probabilistically nondeducible on strategies *if for any time* $t \in \mathbf{N}^+$, *any* $e \in \mathcal{P}(\Omega_t)$ *such that* e *is characterized by a predicate on the inputs and outputs occuring on low channels up to time* t, *any low environment behavior* **L**, *and any two high environment behaviors* **H** *and* **H'**,

$$P_{\mathbf{H},t}(e) = P_{\mathbf{H}',t}(e)$$

The following theorem lends support to the claim that our definition of security prevents low users from gaining any information about the high environment's behavior.

Theorem 4.1 *Given a system* $\Sigma = (C, I, O, S, s0, T)$, *a set of high channels* H *and a set of low channels* L *such that* H *and* L *are disjoint and* $H \cup L = C$, *if* $H \not\rightarrow L$ *then* Σ *is probabilistically nondeducible on strategies.*

Proof: omitted due to page limitation. □

Corollary 4.1 *Given a system* $\Sigma = (C, I, O, S, s0, T)$, *a set of high channels* H *and a set of low channels* L *such that* H *and* L *are disjoint and* $H \cup L = C$, *if* $H \not\rightarrow L$ *then* Σ *is nondeducible on strategies.*

Proof: follows trivially from theorem 4.1 and definitions 4.5 and 4.6. □

5 Relationship to Information Theory

The goal of this section is to show that $H \not\rightarrow L$ implies that the capacity (in information theoretic terms) of the channel from H to L is zero. Unfortunately, due to the generality of our system model (e.g., it has memory (i.e., an internal state), inputs from the receiving end of the channel (i.e., L) and feedback to the sending end of the channel (i.e., H)), there is no existing information theoretic formulation of its channel capacity. For example, Shannon's original formulation of channel capacity is for discrete memoryless channels (with no inputs from the receiver and no feedback to the sender) [13]. So, in this section, we motivate and formulate a definition of the channel capacity for our system model. In future work, we plan to justify our definition of channel capacity by a coding theorem analogous to Shannon's coding theorem for discrete memoryless channels.

We begin with a standard definition from information theory—that of the *mutual information* between two systems of events (see for example [1]). Intuitively, the mutual information between two systems of events is the average amount of information that is gained about one system of events by observing the other system of events.

Definition 5.1 *Let* $S_1 = \{E_1, E_2, \ldots, E_m\}$ *and* $S_2 = \{F_1, F_2, \ldots, F_n\}$ *be sets of mutually disjoint events. The mutual information between* S_1 *and* S_2 *is defined as:*

$$I(S_1; S_2) \equiv \sum_{i=1}^{m} \sum_{j=1}^{n} P(E_i \cap F_j) \log \left(\frac{P(E_i \cap F_j)}{P(E_i)P(F_j)} \right)$$

Mutual information is the basic information theoretic definition that is used to determine a communication channel's capacity. In the typical application of this definition, the E_i's model inputs to the communication channel being analyzed, and the F_j's model outputs from the channel. This interpretation of the E_i's and the F_j's follows from the common assumption that each output (or output block) is associated with (i.e., derived from) a single input (or input block). This assumption is true when the communication channel being analyzed is *memoryless* (viz., memoryless in the sense that after a given input has been transmitted, the channel does not make any further use of that input), which is commonly the case for communication channels. In fact, this is the assumption made by Shannon in his original definition of the capacity of discrete channels with noise [13]. However, the covert channels that we are concerned with may or may not be memoryless. In general, a high input may affect a low output at any later time (i.e., a general purpose computer system may have memory of previous inputs for an indefinite period of time).

To account for the fact that we need to analyze systems with memory and feedback, we interpret the E_i's as histories of high inputs and outputs from time 0 through time $t-1$ (i.e., $S_1 = \textit{In-Seq-Event}_{H,t} \times \textit{Out-Seq-Event}_{H,t}$) and the F_j's as low outputs at (only) time t (i.e., $S_2 = \textit{Final-Out-Event}_{L,t}$). In this way, we can calculate the mutual information between the low output at time t and the entire history of high inputs and outputs from time 0 through time $t-1$.

Even with this interpretation, we cannot apply the above definition of mutual information directly. The reason is that the recipient of the low outputs has some additional information—namely the previous history (i.e., the history from time 0 through time $t-1$) of inputs and outputs on the low channels. This additional information may increase or decrease the quantity of information that is gained about the previous history of high inputs by observing the low output at

time t. Fortunately, there is another concept from information theory that captures this situation perfectly—*conditional mutual information*. Intuitively, the mutual information between two systems of events, conditioned on a third system of events, is the average amount of information that is gained about the first system of events by observing the second system of events, given that the third system of events is already known.

Definition 5.2 *Let* $S_1 = \{E_1, E_2, \ldots, E_l\}$, $S_2 = \{F_1, F_2, \ldots, F_m\}$, *and* $S_3 = \{G_1, G_2, \ldots, G_n\}$ *be sets of mutually disjoint events. The mutual information between* S_1 *and* S_2, *conditioned on* S_3 *is defined as:*

$$I(S_1; S_2 \mid S_3) \equiv \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} P(E_i \cap F_j \cap G_k) \log\left(\frac{P(E_i \cap F_j \mid G_k)}{P(E_i \mid G_k)P(F_j \mid G_k)}\right)$$

Interpreting the E_i's and the F_j's as above, and the G_k's as histories of low inputs and low outputs from time 0 through time $t-1$ (i.e., $S_3 = \textit{In-Seq-Event}_{L,t} \times \textit{Out-Seq-Event}_{L,t}$) we can make the following definition of the "capacity" of the communication channel from high to low. Intuitively, the capacity of a channel is the least upper bound on the rate at which information can be reliably transmitted over the channel.

Definition 5.3 *The* channel capacity from H to L *is defined as:*

$$C \equiv \lim_{n\to\infty} C_n$$

where C_n *is given by:*

$$C_n \equiv \max_{\mathbf{H},\mathbf{L}} \left(\frac{1}{n} \sum_{i=1}^{n} I\left(\begin{array}{l} \textit{In-Seq-Event}_{H,i}, \\ \textit{Out-Seq-Event}_{H,i}; \\ \textit{Final-Out-Event}_{L,i} \mid \\ \textit{In-Seq-Event}_{L,i}, \\ \textit{Out-Seq-Event}_{L,i} \end{array} \right) \right)$$

where $\max_{\mathbf{H},\mathbf{L}}(X)$ *is the maximum value of the expression* X *over all possible values of* **H** *and* **L**.

Note that the conditional mutual information I appearing in the definition of C_n is defined in terms of P_n, which is defined in terms of Σ, **H**, and **L**. Since C_n is maximized over all **H** and **L** (i.e., **H** and **L** are maximized out of the expression), C_n, and hence C, is a function of only Σ.

Gallager [1, pages 97-111] gives a probabilistic model of a "discrete finite state channel" and a definition of channel capacity for his model. His model differs from ours in that there is no input from the receiver (i.e., L) and he considers state machines with a set of possible initial states (rather than a single initial state). Gallager also proves a coding theorem for his model stating that information can be reliably transmitted at rates arbitrarily close to a channel's capacity but not above the capacity. This coding theorem provides justification for the "correctness" of his definition of channel capacity.

Our definition of channel capacity for our model is closely related to Gallager's and we believe that an analogous coding theorem can be stated and proved to provide justification for our definition. As mentioned above, we plan to prove such a coding theorem in future work. For the time being, the skeptical reader may convince himself of the plausibility of this definition by comparing it with [1, equations 4.6.6 and 4.6.7].

Given our definition of channel capacity, we can now state and prove a theorem relating our definition of information flow security to our definition of channel capacity.

Theorem 5.1 *Given a system* $\Sigma = (C, I, O, S, s0, T)$, *a set of high channels* H *and a set of low channels* L *such that* H *and* L *are disjoint and* $H \cup L = C$, *If* $H \not\rightsquigarrow L$ *then the channel capacity from* H *to* L *is* 0.

Proof: Suppose that $H \not\rightsquigarrow L$ for the given Σ, H and, L. Let **H** and **L** be arbitrary high and low environment behaviors, respectively, and let $t \in \mathbf{N}^+$ be an arbitrary time. Then, by definition,

$$\begin{aligned} &I(\textit{In-Seq-Event}_{H,t}, \textit{Out-Seq-Event}_{H,t}; \\ &\quad \textit{Final-Out-Event}_{L,t} \mid \\ &\quad \textit{In-Seq-Event}_{L,t}, \textit{Out-Seq-Event}_{L,t}) \\ &= \sum_{\alpha_H,\beta_H,b_L,\alpha_L,\beta_L} \left(\begin{array}{l} P_t(\alpha_H \cap \beta_H \cap b_L \cap \alpha_L \cap \beta_L) \cdot \\ \log\left(\frac{P_t(\alpha_H \cap \beta_H \cap b_L \mid \alpha_L \cap \beta_L)}{P_t(\alpha_H \cap \beta_H \mid \alpha_L \cap \beta_L) P_t(b_L \mid \alpha_L \cap \beta_L)}\right) \end{array} \right) \\ &= \sum_{\alpha_H,\beta_H,b_L,\alpha_L,\beta_L} \left(\begin{array}{l} P_t(\alpha_H \cap \beta_H \cap b_L \cap \alpha_L \cap \beta_L) \cdot \\ \log\left(\frac{P_t(\alpha_H \cap \beta_H \cap b_L \cap \alpha_L \cap \beta_L) P_t(\alpha_L \cap \beta_L)}{P_t(\alpha_H \cap \beta_H \cap \alpha_L \cap \beta_L) P_t(b_L \cap \alpha_L \cap \beta_L)}\right) \end{array} \right) \\ &= \sum_{\alpha_H,\beta_H,b_L,\alpha_L,\beta_L} \left(\begin{array}{l} P_t(\alpha_H \cap \beta_H \cap b_L \cap \alpha_L \cap \beta_L) \cdot \\ \log\left(\frac{P_t(b_L \mid \alpha_H \cap \beta_H \cap \alpha_L \cap \beta_L)}{P_t(b_L \mid \alpha_L \cap \beta_L)}\right) \end{array} \right) \end{aligned}$$

Now, since $H \not\rightsquigarrow L$, we have:

$$\begin{aligned} &= \sum_{\alpha_H,\beta_H,b_L,\alpha_L,\beta_L} P_t(\alpha_H \cap \beta_H \cap b_L \cap \alpha_L \cap \beta_L) \log(1) \\ &= 0 \end{aligned}$$

Since this holds for any **H**, **L**, and t, we know that for any n,

$$\begin{aligned} C_n &= \max_{\mathbf{H},\mathbf{L}} \left(\frac{1}{n} \sum_{i=1}^{n} I\left(\begin{array}{l} \textit{In-Seq-Event}_{H,i}, \\ \textit{Out-Seq-Event}_{H,i}; \\ \textit{Final-Out-Event}_{L,i} \mid \\ \textit{In-Seq-Event}_{L,i}, \\ \textit{Out-Seq-Event}_{L,i} \end{array} \right) \right) \\ &= 0 \end{aligned}$$

Therefore, $C = 0$. □

6 Verification

It is not immediately obvious that $H \not\rightsquigarrow L$ can in principle (let alone in practice) be verified for nontrivial sys-

tems. In particular showing that $H \not\rightarrow L$ entails verifying an expression involving P_t (which is defined in terms of **H** and **L**) *for any high and low behaviors* **H** *and* **L**. Since **H** and **L** are sets of probability measures, and a probability measure has $[0,1]$ as its image, **H** and **L** range over an *uncountable* set of possibilities. Hence, we cannot enumerate all possible behaviors **H** and **L** and verify that the expression holds for each.

Rather than a condition involving P_t, what we need is some condition(s) involving T (i.e., the transition function that defines the system's internal probabilistic choices) that implies the condition involving P_t. Since T is not defined in terms of **H** and **L**, we should then, in principle, be able to verify that $H \not\rightarrow L$. Such conditions are commonly called *verification conditions*. In this section, we give such a verification condition and in appendix B, we prove that it implies $H \not\rightarrow L$.

First, we define a convenient shorthand for the product of the probabilities associated with a sequence of internal system choices (as given by the system transition function T).

Definition 6.1 *For any* $t \in \mathbf{N}^+$, *define* $T^t : I_{t-1} \times O_t \times S_t \rightarrow [0,1]$ *by:*

for any $\alpha_{t-1} \in I_{t-1}$, *any* $\beta_t \in O_t$, *and any* $\gamma_t \in S_t$,

$$T^t(\alpha_{t-1}, \beta_t, \gamma_t) = T(s0, \mathbf{null}_C, \gamma_t[1], \beta_t[1]) \cdot \prod_{i=2}^{t} T(\gamma_t[i-1], \alpha_{t-1}[i-1], \gamma_t[i], \beta_t[i])$$

For $t = 0$, *define* $T^t : I_0 \times O_0 \times S_0 \rightarrow [0,1]$ *as the constant function* $T(\alpha, \beta, \gamma) = 1$, *for all* α, β, *and* γ.

Now we need some notation for "composing" sequences.

Notation: For any time t, given two input sequences $\alpha_{H,t} \in I_{H,t}$ and $\alpha_{L,t} \in I_{L,t}$, let $\alpha_{H,t} \circ \alpha_{L,t} \in I_{C,t}$ be the element-wise composition of the two sequences yielding the input sequence that is the same as $\alpha_{H,t}$ on channels in H and the same as $\alpha_{L,t}$ on channels in L. We will also use the element-wise composition operator $\circ$ on output sequences and state sequences.

For any time t, any set of channels λ, any input sequence $\alpha_{\lambda,t} \in I_{\lambda,t}$, and any input vector $a \in I[\lambda]$, let $\alpha_{\lambda,t}@a \in I_{\lambda,t+1}$ be the concatenation of a onto the end of $\alpha_{\lambda,t}$. We will also use the concatenation operation @ on output sequences and state sequences.

Theorem 6.1 *Given a system* $\Sigma = (C, I, O, S, s0, T)$, *a set of high channels* H *and a set of low channels* L *such that* H *and* L *are disjoint and* $H \cup L = C$, $H \not\rightarrow L$ *if the following condition holds:*

For any time $t \in \mathbf{N}^+$, *for any low output vector* $b_{L,t} \in O[L]$ *(i.e., the low output vector at time* t*), for any low input and output histories (up to time* $t-1$*)* $\alpha_{L,t-1} \in I_{L,t-1}$ *and* $\beta_{L,t-1} \in O_{L,t-1}$, *for any high input and output histories* $\alpha_{H,t-1} \in I_{H,t-1}$ *and* $\beta_{H,t-1} \in O_{H,t-1}$, *and for any "alternate" high input and output histories* $\alpha'_{H,t-1} \in I_{H,t-1}$ *and* $\beta'_{H,t-1} \in O_{H,t-1}$, *if*

$$\sum_{\gamma \in S_{t-1}} \left(T^{t-1}\left(\begin{array}{l} (\alpha_{L,t-1})_{\rightarrow t-2} \circ (\alpha_{H,t-1})_{\rightarrow t-2}, \\ \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma \end{array} \right) \right) > 0$$

and

$$\sum_{\gamma \in S_{t-1}} \left(T^{t-1}\left(\begin{array}{l} (\alpha_{L,t-1})_{\rightarrow t-2} \circ (\alpha'_{H,t-1})_{\rightarrow t-2}, \\ \beta_{L,t-1} \circ \beta'_{H,t-1}, \gamma \end{array} \right) \right) > 0$$

then

$$\frac{\displaystyle\sum_{b_{H,t}} \sum_{\gamma \in S_t} \left(T^t\left(\begin{array}{l} \alpha_{L,t-1} \circ \alpha_{H,t-1}, \\ (\beta_{L,t-1}@b_{L,t}) \circ (\beta_{H,t-1}@b_{H,t}), \gamma \end{array} \right) \right)}{\displaystyle\sum_{\gamma \in S_{t-1}} \left(T^{t-1}\left(\begin{array}{l} (\alpha_{L,t-1})_{\rightarrow t-2} \circ (\alpha_{H,t-1})_{\rightarrow t-2}, \\ \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma \end{array} \right) \right)} =$$

$$\frac{\displaystyle\sum_{b_{H,t}} \sum_{\gamma \in S_t} \left(T^t\left(\begin{array}{l} \alpha_{L,t-1} \circ \alpha'_{H,t-1}, \\ (\beta_{L,t-1}@b_{L,t}) \circ (\beta'_{H,t-1}@b_{H,t}), \gamma \end{array} \right) \right)}{\displaystyle\sum_{\gamma \in S_{t-1}} \left(T^{t-1}\left(\begin{array}{l} (\alpha_{L,t-1})_{\rightarrow t-2} \circ (\alpha'_{H,t-1})_{\rightarrow t-2}, \\ \beta_{L,t-1} \circ \beta'_{H,t-1}, \gamma \end{array} \right) \right)}$$

We will refer to the above condition as the "verification condition".

Proof: see Appendix B. □

To convey some of the intuition behind this theorem, we give a rough sketch of the proof here. The interested reader can of course find all of the details in Appendix B.

In Lemma B.2 (in Appendix B) we show that

$$\frac{\displaystyle\sum_{b_{H,t}} \sum_{\gamma \in S_t} \left(T^t\left(\begin{array}{l} \alpha_{L,t-1} \circ \alpha_{H,t-1}, \\ (\beta_{L,t-1}@b_{L,t}) \circ (\beta_{H,t-1}@b_{H,t}), \gamma \end{array} \right) \right)}{\displaystyle\sum_{\gamma \in S_{t-1}} \left(T^{t-1}\left(\begin{array}{l} (\alpha_{L,t-1})_{\rightarrow t-2} \circ (\alpha_{H,t-1})_{\rightarrow t-2}, \\ \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma \end{array} \right) \right)}$$

is (under certain conditions) equal to

$$P_t(b_{L,t} \mid \alpha_{L,t-1} \cap \beta_{L,t-1} \cap \alpha_{H,t-1} \cap \beta_{H,t-1})$$

Therefore, the verification condition says (roughly) that for any low output vector $b_{L,t}$ (at time t), for any low input and output histories (up to time $t-1$) $\alpha_{L,t-1}$ and $\beta_{L,t-1}$, for any high input and output histories $\alpha_{H,t-1}$ and $\beta_{H,t-1}$, and for any "alternate" high input and output histories $\alpha'_{H,t-1}$ and $\beta'_{H,t-1}$,

$$P_t(b_{L,t} \mid \alpha_{L,t-1} \cap \beta_{L,t-1} \cap \alpha_{H,t-1} \cap \beta_{H,t-1}) = P_t(b_{L,t} \mid \alpha_{L,t-1} \cap \beta_{L,t-1} \cap \alpha'_{H,t-1} \cap \beta'_{H,t-1})$$

This can then be used to show (still speaking somewhat roughly) that for any low output vector $b_{L,t}$ (at time t), for any low input and output histories (up to time $t-1$) $\alpha_{L,t-1}$ and $\beta_{L,t-1}$, for any high input and output histories $\alpha_{H,t-1}$ and $\beta_{H,t-1}$,

$$P_t(b_{L,t} \mid \alpha_{L,t-1} \cap \beta_{L,t-1} \cap \alpha_{H,t-1} \cap \beta_{H,t-1}) = P_t(b_{L,t} \mid \alpha_{L,t-1} \cap \beta_{L,t-1})$$

which says that $H \not\rightarrow L$, thus completing the proof.

Note that as desired, this theorem provides us with a verification condition that is stated solely in terms of the system's internal transition function T. Therefore, it seems that (at least in principle) it is possible to verify that $H \not\rightarrow L$. Furthermore, the verification condition looks like it is conducive to a proof by induction, in which case we can further reduce the burden on the system verifier by providing a general purpose "unwinding" of this verification condition into a set of verification conditions that are even easier to verify.

7 Conclusions and Future Work

Since theorem 6.1 gives a verification condition that is stated solely in terms of the system's transition function (rather than in terms of probabilities that depend on the environment's behavior), it and theorem 5.1 together provide a condition that can in principle be verified (and we believe that with automated support it can be verified in practice) and that implies that the capacity of all covert channels put together (including storage channels, timing channels, noiseless channels, and noisy channels) is zero.

However, there are (at least) two remaining links that need to be taken care of before this work can be called complete. First of all, we need to prove that our definition of capacity is actually an upper bound on the rate at which information can be transmitted from high to low. Second, we need to provide a connection between source code and our system model. We plan to provide this connection in the form of an operational or denotational semantics for a high level progamming language. After these two remaining links are complete, it will be possible to establish that a given piece of source code introduces no covert channels of any kind into a computer system.

Acknowledgements

Through both published papers and private discussions, John McLean has had a big influence on this work. In particular, the definition of security in section 4 is essentially McLean's FM applied to the present problem and with some additional details fleshed out. Also, I owe much to Ira Moskowitz for help with probability theory, information theory, and for simplifying the proof of Theorem 6.1. Thanks to Todd Wittbold, and Robert Morris for asking questions at the Franconia workshop that led to this work. Also, thanks to Todd Wittbold for suggesting Theorem 4.1, to Paul Syverson for several helpful discussions, and to Jeremy Jacob, Cathy Meadows, and the anonymous reviewers for useful reviews of a previous draft.

A The Proof of Theorem 3.1

Proof: We must show that for all $t \in \mathbf{N}^+$ the following three conditions hold: (1) for all $A \in \mathcal{P}(\Omega_t)$, $P_t(A)$ is nonnegative, (2) $P_t(\Omega_t) = 1$, and (3) for any countable set of mutually disjoint events $\{A_i\}$, $P_t(\bigcup_i A_i) = \sum_i P_t(A_i)$. The proof is by induction.

Base case: To show (1), let $A \in \mathcal{P}(\Omega_1)$ be an arbitrary event. By definition,

$$P_1(A) = \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \qquad \text{if } (\alpha,\beta,\gamma) \in A; \\ 0, \qquad \text{otherwise.} \end{cases}$$

Since H_1 and L_1 are probability measures, each of the $H_1(\{\pi_H(\alpha[1])\})$ and $L_1(\{\pi_L(\alpha[1])\})$ terms are nonnegative. Also, we know from definition 2.1 (i.e., the definition of a system), that $T(s0, \mathbf{null}_C, \gamma[1], \beta[1]) = P_{s0,\mathbf{null}_C}(\gamma[1], \beta[1])$ and that $P_{s0,\mathbf{null}_C}$ is a probability measure. Therefore, $T(s0, \mathbf{null}_C, \gamma[1], \beta[1])$ is nonnegative. Since all terms in all summands are nonnegative, $P_1(A)$ is nonnegative.

Now we show (2). By definition,

$$P_1(\Omega_1)$$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \qquad \text{if } (\alpha,\beta,\gamma) \in \Omega_1; \\ 0, \qquad \text{otherwise.} \end{cases}$$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{array}{l} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]) \end{array}$$

$$= \sum_{\substack{\alpha_H\in I[H] \\ \alpha_L\in I[L] \\ \beta\in O_{C,1} \\ \gamma\in S_1}} \begin{array}{l} H_1(\{\alpha_H\}) \cdot \\ L_1(\{\alpha_L\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]) \end{array}$$

$$= \begin{array}{l} \displaystyle\sum_{\alpha_H\in I[H]} H_1(\{\alpha_H\}) \cdot \\ \displaystyle\sum_{\alpha_L\in I[L]} L_1(\{\alpha_L\}) \cdot \\ \displaystyle\sum_{\substack{\beta\in O_{C,1} \\ \gamma\in S_1}} T(s0, \mathbf{null}_C, \gamma[1], \beta[1]) \end{array}$$

$$= \sum_{\alpha_H \in I[H]} H_1(\{\alpha_H\}) \cdot \sum_{\alpha_L \in I[L]} L_1(\{\alpha_L\}) \cdot \sum_{\substack{\beta \in O_{C,1} \\ \gamma \in S_1}} P_{s0,\mathbf{null}_C}(\gamma[1], \beta[1])$$

Since H_1, L_1 and $P_{s0,\mathbf{null}_C}$ are all probability measures, and each is being summed over its entire sample space, each summation sums to 1. Therefore, $P_1(\Omega_1) = 1$.

Now to show (3), let $\{A_i\}$ be a countable set of mutually disjoint events in $\mathcal{P}(\Omega_1)$. By definition,

$P_1(\bigcup_i A_i)$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in \bigcup_i A_i; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= \sum_i \sum_{(\alpha,\beta,\gamma)\in\Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot \\ L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in A_i; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= \sum_i P_1(A_i)$$

Therefore, P_1 is a probability measure.

Induction case: Suppose $t \geq 2$. The induction hypothesis is that P_{t-1} is a probability measure.

To show (1), let $A \in \Omega_t^1$ be an arbitrary event. By definition,

$$P_t(A) = \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \begin{cases} P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ \quad \text{if } (\alpha,\beta,\gamma) \in A; \\ 0, \quad \text{otherwise.} \end{cases}$$

Since P_{t-1} is a probability measure (by the induction hypothesis), $P_{t-1}(\{(\zeta(\alpha))\})$ is nonnegative. Since all of the other terms are nonnegative (by arguments analogous to the base case), $P_t(A)$ is nonnegative.

Now we show (2). By definition,

$P_t(\Omega_t)$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \begin{cases} P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in \Omega_t; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \left(\begin{array}{l} P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t]) \end{array} \right)$$

$$= \sum_{\substack{(\alpha,\beta,\gamma)\in\Omega_{t-1} \\ \beta_t \in O[C] \\ \gamma_t \in S \\ \alpha_{H,t} \in I[H] \\ \alpha_{L,t} \in I[L]}} \left(\begin{array}{l} P_{t-1}(\{(\alpha,\beta,\gamma)\}) \cdot \\ H_{t,\pi_H(\alpha),\pi_H(\beta)}(\{\alpha_{H,t}\}) \cdot \\ L_{t,\pi_L(\alpha),\pi_L(\beta)}(\{\alpha_{L,t}\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma_t, \beta_t) \end{array} \right)$$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_{t-1}} P_{t-1}(\{(\alpha,\beta,\gamma)\}) \cdot \sum_{\substack{\beta_t \in O[C] \\ \gamma_t \in S}} T(\gamma[t-1], \alpha[t-1], \gamma_t, \beta_t) \cdot \sum_{\alpha_{H,t} \in I[H]} H_{t,\pi_H(\alpha),\pi_H(\beta)}(\{\alpha_{H,t}\}) \cdot \sum_{\alpha_{L,t} \in I[L]} L_{t,\pi_L(\alpha),\pi_L(\beta)}(\{\alpha_{L,t}\})$$

Since by the induction hypothesis P_{t-1} is a probability measure, and $H_{t,\pi_H(\alpha),\pi_H(\beta)}$, $L_{t,\pi_L(\alpha),\pi_L(\beta)}$, and $P_{s0,\mathbf{null}_C}$ are all probability measures, and each is being summed over its entire sample space, each summation sums to 1. Therefore, $P_t(\Omega_t) = 1$.

Now to show (3), let $\{A_i\}$ be a countable set of mutually disjoint events in $\mathcal{P}(\Omega_t)$. By definition,

$P_t(\bigcup_i A_i)$

$$= \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \begin{cases} P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in \bigcup_i A_i; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= \sum_i \sum_{(\alpha,\beta,\gamma)\in\Omega_t} \begin{cases} P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot \\ T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in A_i; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= \sum_i P_t(A_i)$$

Thus, P_t is a probability measure; and by induction, for all $t \in \mathbf{N}^+$, P_t is a probability measure. □

B The Proof of Theorem 6.1

We begin by stating and proving two lemmas relating T^t and P_t.

Lemma B.1 *Let $t \in \mathbf{N}^+$ be any arbitrary time. Let $(\alpha, \beta, \gamma) \in \Omega_t$ be an arbitrary history of the system up to time t. Then,*

$$P_t(\{(\alpha,\beta,\gamma)\})$$

$$= T^t(\zeta(\alpha),\beta,\gamma) \cdot H_1(\{\pi_H(\alpha[1])\}) \cdot L_1(\{\pi_L(\alpha[1])\}) \cdot \prod_{i=2}^{t} H_{i,\pi_H(\alpha_{\to i-1}),\pi_H(\beta_{\to i-1})}(\pi_H(\alpha[i])) \cdot \prod_{i=2}^{t} L_{i,\pi_L(\alpha_{\to i-1}),\pi_L(\beta_{\to i-1})}(\pi_L(\alpha[i]))$$

Proof: The proof is by induction.

Base case: $t = 1$. Let $(\alpha, \beta, \gamma) \in \Omega_1$ be an arbitrary history of the system up to time 1. By definition,

$$P_1(\{(\alpha,\beta,\gamma)\})$$

$$= \sum_{(\alpha',\beta',\gamma') \in \Omega_1} \begin{cases} H_1(\{\pi_H(\alpha'[1])\}) \cdot \\ L_1(\{\pi_L(\alpha'[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma'[1], \beta'[1]), \\ \qquad \text{if } (\alpha',\beta',\gamma') \in \{(\alpha,\beta,\gamma)\}; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= H_1(\{\pi_H(\alpha[1])\}) \cdot L_1(\{\pi_L(\alpha[1])\}) \cdot T(s0, \mathbf{null}_C, \gamma[1], \beta[1])$$

$$= T^1(\alpha_{\to 0},\beta,\gamma) \cdot H_1(\{\pi_H(\alpha[1])\}) \cdot L_1(\{\pi_L(\alpha[1])\}) \cdot \prod_{i=2}^{1} H_{i,\pi_H(\alpha_{\to i-1}),\pi_H(\beta_{\to i-1})}(\pi_H(\alpha[i])) \cdot \prod_{i=2}^{1} L_{i,\pi_L(\alpha_{\to i-1}),\pi_L(\beta_{\to i-1})}(\pi_L(\alpha[i]))$$

Induction case: $t \geq 2$. The induction hypothesis is for any $(\alpha, \beta, \gamma) \in \Omega_{t-1}$,

$$P_{t-1}(\{(\alpha,\beta,\gamma)\})$$

$$= T^{t-1}(\alpha_{\to t-2},\beta,\gamma) \cdot H_1(\{(\pi_H(\alpha[1])\}) \cdot L_1(\{(\pi_L(\alpha[1])\}) \cdot \prod_{i=2}^{t-1} H_{i,\pi_H(\alpha_{\to i-1}),\pi_H(\beta_{\to i-1})}(\pi_H(\alpha[i])) \cdot \prod_{i=2}^{t-1} L_{i,\pi_L(\alpha_{\to i-1}),\pi_L(\beta_{\to i-1})}(\pi_L(\alpha[i]))$$

Let $(\alpha, \beta, \gamma) \in \Omega_t$ be an arbitrary history of the system up to time t. By definition,

$$P_t(\{(\alpha,\beta,\gamma)\})$$

$$= \sum_{(\alpha',\beta',\gamma') \in \Omega_t} \begin{cases} P_{t-1}(\{(\zeta(\alpha'),\zeta(\beta'),\zeta(\gamma'))\}) \cdot \\ H_{t,\pi_H(\zeta(\alpha')),\pi_H(\zeta(\beta'))}(\{\pi_H(\alpha'[t])\}) \cdot \\ L_{t,\pi_L(\zeta(\alpha')),\pi_L(\zeta(\beta'))}(\{\pi_L(\alpha'[t])\}) \cdot \\ T(\gamma'[t-1], \alpha'[t-1], \gamma'[t], \beta'[t]), \\ \qquad \text{if } (\alpha',\beta',\gamma') \in \{(\alpha,\beta,\gamma)\}; \\ 0, \quad \text{otherwise.} \end{cases}$$

$$= P_{t-1}(\{(\zeta(\alpha),\zeta(\beta),\zeta(\gamma))\}) \cdot H_{t,\pi_H(\zeta(\alpha)),\pi_H(\zeta(\beta))}(\{\pi_H(\alpha[t])\}) \cdot L_{t,\pi_L(\zeta(\alpha)),\pi_L(\zeta(\beta))}(\{\pi_L(\alpha[t])\}) \cdot T(\gamma[t-1], \alpha[t-1], \gamma[t], \beta[t])$$

Applying the induction hypothesis, we have:

$$= T^{t-1}(\alpha_{\to t-2},\zeta(\beta),\zeta(\gamma)) \cdot H_1(\{(\pi_H(\alpha[1])\}) \cdot L_1(\{(\pi_L(\alpha[1])\}) \cdot \prod_{i=2}^{t-1} H_{i,\pi_H(\alpha_{\to i-1}),\pi_H(\beta_{\to i-1})}(\pi_H(\alpha[i])) \cdot \prod_{i=2}^{t-1} L_{i,\pi_L(\alpha_{\to i-1}),\pi_L(\beta_{\to i-1})}(\pi_L(\alpha[i])) \cdot H_{t,\pi_H(\zeta(\alpha')),\pi_H(\zeta(\beta'))}(\{\pi_H(\alpha'[t])\}) \cdot L_{t,\pi_L(\zeta(\alpha')),\pi_L(\zeta(\beta'))}(\{\pi_L(\alpha'[t])\}) \cdot T(\gamma'[t-1], \alpha'[t-1], \gamma'[t], \beta'[t])$$

$$= T^t(\zeta(\alpha),\beta,\gamma) \cdot H_1(\{\pi_H(\alpha[1])\}) \cdot L_1(\{\pi_L(\alpha[1])\}) \cdot \prod_{i=2}^{t} H_{i,\pi_H(\alpha_{\to i-1}),\pi_H(\beta_{\to i-1})}(\pi_H(\alpha[i])) \cdot \prod_{i=2}^{t} L_{i,\pi_L(\alpha_{\to i-1}),\pi_L(\beta_{\to i-1})}(\pi_L(\alpha[i]))$$

Therefore, by induction the lemma holds for all $t \in \mathbf{N}^+$. □

Lemma B.2 *Let $t \in \mathbf{N}^+$ be any arbitrary time. Let $b_{L,t} \in O[L]$ be an arbitrary low output vector ($b_{L,t}$ should be thought of as the low output vector at time t). Let $\alpha_{L,t-1} \in I_{L,t-1}$, $\beta_{L,t-1} \in O_{L,t-1}$, $\alpha_{H,t-1} \in I_{H,t-1}$, and $\beta_{H,t-1} \in O_{H,t-1}$ be arbitrary input and output histories of the low and high channels, respectively. Further, let $E(b_{L,t}) \in$ Final-Out-Event$_{L,t}$ be the event representing the occurrence of the output vector $b_{L,t}$ on the low channels at time t. Similarly, let $E(\alpha_{L,t-1}) \in$ In-Seq-Event$_{L,t-1}$, $E(\beta_{L,t-1}) \in$ Out-Seq-Event$_{L,t-1}$, $E(\alpha_{H,t-1}) \in$ InputSequence-*

$Event_{H,t-1}$, and $E(\beta_{H,t-1}) \in$ Out-Seq-$Event_{H,t-1}$ be the events representing the occurence of $\alpha_{L,t-1}$, $\beta_{L,t-1}$, $\alpha_{H,t-1}$, and $\beta_{H,t-1}$, respectively. Then,

$$P_t(\quad E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1})) > 0 \Rightarrow$$

$$P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1}))$$

$$= \frac{\sum_{b_{H,t} \in O[H]} \sum_{\gamma \in S_t} T^t(\alpha_{L,t-1} \circ \alpha_{H,t-1}, (\beta_{L,t-1} @ b_{L,t}) \circ (\beta_{H,t-1} @ b_{H,t}), \gamma)}{\sum_{\gamma \in S_{t-1}} T^{t-1}((\alpha_{L,t-1})_{\to t-2} \circ (\alpha_{H,t-1})_{\to t-2}, \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma)}$$

Proof: The proof is in two cases.

Case 1: $t = 1$.

$$P_1(E(b_{L,1}) \mid E(\alpha_{L,0}) \cap E(\beta_{L,0}) \cap E(\alpha_{H,0}) \cap E(\beta_{H,0}))$$

$$= \frac{P_1(E(b_{L,1}) \cap E(\alpha_{L,0}) \cap E(\beta_{L,0}) \cap E(\alpha_{H,0}) \cap E(\beta_{H,0}))}{P_1(E(\alpha_{L,0}) \cap E(\beta_{L,0}) \cap E(\alpha_{H,0}) \cap E(\beta_{H,0}))}$$

$$= \frac{P_1(E(b_{L,1}))}{P_1(\Omega_1))}$$

$$= \frac{\sum_{(\alpha,\beta,\gamma) \in \Omega_1} \begin{cases} H_1(\{\pi_H(\alpha[1])\}) \cdot L_1(\{\pi_L(\alpha[1])\}) \cdot \\ T(s0, \mathbf{null}_C, \gamma[1], \beta[1]), \\ \quad \text{if } (\alpha,\beta,\gamma) \in E(b_{L,1}); \\ 0, \quad \text{otherwise.} \end{cases}}{1}$$

$$= \frac{\sum_{b_{H,1} \in O[H]} \sum_{\gamma \in S_1} T(s0, \mathbf{null}_C, \gamma[1], b_{L,1} \circ b_{H,1})}{1}$$

$$= \frac{\sum_{b_{H,1} \in O[H]} \sum_{\gamma \in S_1} T^1(\alpha_{L,0} \circ \alpha_{H,0}, b_{L,1} \circ b_{H,1}, \gamma)}{\sum_{\gamma \in S_0} T^0(\alpha_{L,0} \circ \alpha_{H,0}, \beta_{L,0} \circ \beta_{H,0}, \gamma)}$$

Therefore, the lemma holds for $t = 1$.

Case 2: $t \geq 2$.

$$P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1}))$$

$$= \frac{P_t(E(b_{L,t}) \cap E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1}))}{P_t(E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1}))}$$

$$= \frac{\sum_{\substack{a_{L,t} \in I[L] \\ a_{H,t} \in I[H] \\ b_{H,t} \in O[H]}} \sum_{\gamma \in S_t} \left(\begin{array}{l} P_{t-1}(\{(\alpha_{L,t-1} \circ \alpha_{H,t-1}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \zeta(\gamma))\}) \cdot \\ H_{t,\alpha_{H,t-1}),\beta_{H,t-1})}(\{a_{H,t}\}) \cdot \\ L_{t,\alpha_{L,t-1}),\beta_{L,t-1})}(\{a_{L,t}\}) \cdot \\ T(\gamma[t-1], (\alpha_{L,t-1} \circ \alpha_{H,t-1})[t-1], \\ \quad \gamma[t], b_{L,t} \circ b_{H,t}) \end{array} \right)}{\sum_{\substack{a_{L,t} \in I[L] \\ a_{H,t} \in I[H] \\ b_{H,t} \in O[H]}} \sum_{\gamma \in S_t} \left(\begin{array}{l} P_{t-1}(\{(\alpha_{L,t-1} \circ \alpha_{H,t-1}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \zeta(\gamma))\}) \cdot \\ H_{t,\alpha_{H,t-1}),\beta_{H,t-1})}(\{a_{H,t}\}) \cdot \\ L_{t,\alpha_{L,t-1}),\beta_{L,t-1})}(\{a_{L,t}\}) \cdot \\ T(\gamma[t-1], (\alpha_{L,t-1} \circ \alpha_{H,t-1})[t-1], \\ \quad \gamma[t], b_t) \end{array} \right)}$$

$$= \frac{\sum_{\substack{b_{H,t} \in O[H] \\ \gamma \in S_t}} \left(\begin{array}{l} P_{t-1}(\{(\alpha_{L,t-1} \circ \alpha_{H,t-1}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \zeta(\gamma))\}) \cdot \\ T(\gamma[t-1], (\alpha_{L,t-1} \circ \alpha_{H,t-1})[t-1], \\ \quad \gamma[t], b_{L,t} \circ b_{H,t}) \end{array} \right)}{\sum_{\gamma \in S_{t-1}} P_{t-1}(\{(\alpha_{L,t-1} \circ \alpha_{H,t-1}, \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma)\})}$$

Applying Lemma 6.1, we have:

$$= \frac{\sum_{\substack{b_{H,t} \in O[H] \\ \gamma \in S_t}} \left(\begin{array}{l} T^{t-1}((\alpha_{L,t-1} \circ \alpha_{H,t-1})_{\to t-2}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \zeta(\gamma)) \cdot \\ H_1(\{\alpha_{H,t-1}[1]\}) \cdot L_1(\{\alpha_{L,t-1}[1]\}) \cdot \\ \prod_{i=2}^{t-1} H_{i,\alpha_{H,t-1\to i-1},\beta_{H,t-1\to i-1}}(\alpha_{H,t-1}[i]) \cdot \\ \prod_{i=2}^{t-1} L_{i,\alpha_{L,t-1\to i-1},\beta_{L,t-1\to i-1}}(\alpha_{L,t-1}[i]) \cdot \\ T(\gamma[t-1], (\alpha_{L,t-1} \circ \alpha_{H,t-1})[t-1], \\ \quad \gamma[t], b_{L,t} \circ b_{H,t}) \end{array} \right)}{\sum_{\gamma \in S_{t-1}} \left(\begin{array}{l} T^{t-1}((\alpha_{L,t-1} \circ \alpha_{H,t-1})_{\to t-2}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma) \cdot \\ H_1(\{\alpha_{H,t-1}[1]\})\}) \cdot L_1(\{\alpha_{L,t-1}[1]\})\}) \cdot \\ \prod_{i=2}^{t-1} H_{i,\alpha_{H,t-1\to i-1},\beta_{H,t-1\to i-1}}(\alpha_{H,t-1}[i]) \cdot \\ \prod_{i=2}^{t-1} L_{i,\alpha_{L,t-1\to i-1},\beta_{L,t-1\to i-1}}(\alpha_{L,t-1}[i]) \end{array} \right)}$$

$$= \frac{\sum_{\substack{b_{H,t} \in O[H] \\ \gamma \in S_t}} \left(\begin{array}{l} T^{t-1}((\alpha_{L,t-1} \circ \alpha_{H,t-1})_{\to t-2}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \zeta(\gamma)) \cdot \\ T(\gamma[t-1], (\alpha_{L,t-1} \circ \alpha_{H,t-1})[t-1], \\ \quad \gamma[t], b_{L,t} \circ b_{H,t}) \end{array} \right)}{\sum_{\gamma \in S_{t-1}} \left(\begin{array}{l} T^{t-1}((\alpha_{L,t-1} \circ \alpha_{H,t-1})_{\to t-2}, \\ \quad \beta_{L,t-1} \circ \beta_{H,t-1}, \gamma) \cdot \end{array} \right)}$$

$$= \frac{\sum_{\substack{b_{H,t}\in O[H]\\ \gamma\in S_t}} \left(\begin{array}{c} T^t(\zeta(\alpha_{L,t-1}\circ\alpha_{H,t-1}),\\ \beta_{L,t-1}@b_{L,t}\circ\beta_{H,t-1}@b_{H,t},\gamma) \end{array} \right)}{\sum_{\gamma\in S_{t-1}} \left(\begin{array}{c} T^{t-1}((\alpha_{L,t-1}\circ\alpha_{H,t-1})_{\rightarrow t-2},\\ \beta_{L,t-1}\circ\beta_{H,t-1},\gamma) \end{array} \right)}$$

Therefore, the lemma holds for all $t \in \mathbf{N}^+$. □

We can now prove theorem 6.1.

Proof (of Theorem 6.1): Assume that the verification condition is true. Let $t \in \mathbf{N}^+$ be an arbitrary time and let $b_{L,t} \in O[L]$ be an arbitrary low output vector at time t. Let $E(b_{L,t}) \in$ *Final-Out-Event*$_{L,t}$ be the event representing the occurrence of the ouput vector $b_{L,t}$ on the low channels at time t.

Note: in the following, we will continue to use the notational convention that when $E(x)$ is an event representing the occurrence of a particular part of the system's history (e.g., the history of low inputs up to time t), then x will represent that particular part of the system's history. For example, if $E(\alpha_{L,t})$ is the *event* representing that the history of high inputs up to time t was some particular sequence α, then we will freely use $\alpha_{L,t}$ to represent the sequence α.

Now let $E(\alpha_{L,t-1}) \in$ *In-Seq-Event*$_{L,t}$, $E(\beta_{L,t-1}) \in$ *Out-Seq-Event*$_{L,t}$, and $E(\alpha_{H,t-1}) \in$ *In-Seq-Event*$_{H,t}$ be arbitrary histories of the low input, low output, and high input, respectively, such that $P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})) > 0$. Consider the set Φ of all events corresponding to possible high output histories up to time $t-1$ given $E(\alpha_{L,t-1})$, $E(\beta_{L,t-1})$, and $E(\alpha_{H,t-1})$:

$$\Phi \equiv \{\ E(\beta_{H,t-1}) \mid E(\beta_{H,t-1}) \in \textit{Out-Seq-Event}_{H,t} \text{ and } P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) > 0\ \}$$

First, note that the elements of Φ are mutually disjoint and that

$$\sum_{E(\beta_{H,t-1})\in\Phi} \left(\begin{array}{c} P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{L,t-1})) \end{array} \right)$$

$$= P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1}))$$

Next, note that

$$P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) > 0$$

implies that

$$\sum_{\gamma\in S_{t-1}} \left(\begin{array}{c} T^{t-1}((\alpha_{L,t-1})_{\rightarrow t-2}\circ(\alpha_{H,t-1})_{\rightarrow t-2},\\ \beta_{L,t-1}\circ\beta_{H,t-1},\gamma) \end{array} \right) > 0$$

Now, consider any two elements of Φ: $E(\beta_{H,t-1})$ and $E(\beta'_{H,t-1})$. By Lemma 6.2,

$$P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta_{H,t-1}))$$

$$= \frac{\sum_{\substack{b_{H,t}\in O[H]\\ \gamma\in S_t}} \begin{array}{c} T^t(\alpha_{L,t-1}\circ\alpha_{H,t-1},\\ (\beta_{L,t-1}@b_{L,t})\circ(\beta_{H,t-1}@b_{H,t}),\gamma) \end{array}}{\sum_{\gamma\in S_{t-1}} \begin{array}{c} T^{t-1}((\alpha_{L,t-1})_{\rightarrow t-2}\circ(\alpha_{H,t-1})_{\rightarrow t-2},\\ \beta_{L,t-1}\circ\beta_{H,t-1},\gamma) \end{array}}$$

and

$$P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta'_{H,t-1}))$$

$$= \frac{\sum_{\substack{b_{H,t}\in O[H]\\ \gamma\in S_t}} \begin{array}{c} T^t(\alpha_{L,t-1}\circ\alpha_{H,t-1},\\ (\beta_{L,t-1}@b_{L,t})\circ(\beta'_{H,t-1}@b_{H,t}),\gamma) \end{array}}{\sum_{\gamma\in S_{t-1}} \begin{array}{c} T^{t-1}((\alpha_{L,t-1})_{\rightarrow t-2}\circ(\alpha_{H,t-1})_{\rightarrow t-2},\\ \beta_{L,t-1}\circ\beta'_{H,t-1},\gamma) \end{array}}$$

And so, by the verification condition (which holds by assumption),

$$P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta_{H,t-1}))$$

$$= P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap E(\alpha_{H,t-1})\cap E(\beta'_{H,t-1}))$$

Since this holds for *any* two elements of Φ, there must exist some constant $C \in [0,1]$ such that for any $x \in \Phi$, $P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap x) = C$.

Now,

$$\left(\begin{array}{c} P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) \end{array} \right) = C,$$

for any $E(\beta_{H,t-1}) \in \Phi$

$$\Rightarrow \left(\frac{\begin{array}{c} P_t(E(b_{L,t})\cap E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) \end{array}}{\begin{array}{c} P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) \end{array}} \right) = C,$$

for any $E(\beta_{H,t-1}) \in \Phi$

$$\Rightarrow \left(\begin{array}{c} P_t(E(b_{L,t})\cap E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) \end{array} \right)$$

$$= C \cdot \left(\begin{array}{c} P_t(E(\alpha_{L,t-1})\cap E(\beta_{L,t-1})\cap\\ E(\alpha_{H,t-1})\cap E(\beta_{H,t-1})) \end{array} \right),$$

for any $E(\beta_{H,t-1}) \in \Phi$

$$\Rightarrow \sum_{E(\beta_{H,t-1})\in\Phi} \left(\begin{array}{c} P_t(E(b_{L,t}) \cap E(\alpha_{L,t-1}) \cap \\ E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1})) \end{array} \right)$$

$$= \sum_{E(\beta_{H,t-1})\in\Phi} C \cdot \left(\begin{array}{c} P_t(E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap \\ E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1})) \end{array} \right)$$

$$\Rightarrow P_t(E(b_{L,t}) \cap E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}))$$
$$= C \cdot P_t(E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}))$$

$$\Rightarrow P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}))$$
$$= C$$

Therefore, we know that for any $E(\alpha_{L,t-1})$, $E(\beta_{L,t-1})$, $E(\alpha_{H,t-1})$, and $E(\beta_{H,t-1})$, $P_t(E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1})) > 0 \Rightarrow P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}) \cap E(\beta_{H,t-1})) = P_t(E(b_{L,t}) \mid E(\alpha_{L,t-1}) \cap E(\beta_{L,t-1}) \cap E(\alpha_{H,t-1}))$, and hence, $H \not\rightarrow L$. □

References

[1] Robert G. Gallager. *Information Theory and Reliable Communication.* John Wiley and Sons, Inc., New York, 1968.

[2] J. A. Goguen and J. Meseguer. Security policies and security models. In *Proceedings of the 1982 IEEE Computer Society Symposium on Computer Security and Privacy*, Oakland, CA, 1982.

[3] James W. Gray, III. Information sharing in secure systems. In *Proc. Computer Security Foundations Workshop III*, Franconia, NH, June 1990.

[4] Daryl McCullough. Specifications for multilevel security and a hook-up property. In *Proceedings of the 1987 IEEE Computer Society Symposium on Computer Security and Privacy*, Oakland, CA, 1987.

[5] Daryl McCullough. Noninterference and the composability of security properties. In *Proceedings of the 1988 IEEE Computer Society Symposium on Computer Security and Privacy*, Oakland, CA, 1988.

[6] John McLean. Security models and information flow. In *Proc. 1990 IEEE Symposium on Security and Privacy*, Oakland, CA, May 1990.

[7] Jonathan K. Millen. Covert channel capacity. In *Proceedings of the 1987 IEEE Computer Society Symposium on Computer Security and Privacy*, Oakland, CA, 1987.

[8] Jonathan K. Millen. Hookup security for synchronous machines. In *Proceedings of the Computer Security Foundations Workshop III*, Franconia, NH, June 1990.

[9] Ira S. Moskowitz. Noise effects upon a simple timing channel. *NRL Memorandum Report 6740*, Washington, DC. Submitted 9 July 1990 (To appear).

[10] Amir Pnueli and Lenore Zuck. Verification of Multiprocess Probabilistic Protocols. *Distributed Computing*, 1:53–72, 1986.

[11] William Pugh. Skip Lists: A Probabilistic Alternative to Balanced Trees. *CACM*, 33(6):668–676, June 1990.

[12] John Rushby. A trusted computing base for embedded systems. In *Proceedings 7th DoD/NBS Computer Security Initiative Conference*, pages 120–136, Gaithersburg, MD, September 1984.

[13] C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423, July 1948. Republished in: C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, IL 1949.

[14] David Sutherland. A model of information. In *Proceeding of the 9th National Computer Security Conference*, Baltimore, MD, September 1986.

[15] J. Todd Wittbold and Dale M. Johnson. Information flow in nondeterministic systems. In *Proceedings of the 1990 IEEE Computer Society Symposium on Computer Security and Privacy*, Oakland, CA, 1990.