A Knowledge-Based Interpretation of Possibilistic Information Flow

Riccardo Pucella

February 19, 2001

1 Introduction

In [12], Zakinthinos and Lee describe a generalization of the *secure interleaving functions* framework of McLean [7]. The generalization is couched in terms of Low Level Equivalence Sets (LLES) that, informally, represent the sequence of low level events appearing in a particular trace of the system. The low level events are the only events that a low level user can observe. The general approach to information flow security is to formally specify and restrict the information that a low level user can infer about high level events from the observation of low level events. We attempt here a knowledge-based formalization of [12]. For the time being, we assume a knowledge of possible-worlds semantics for modal logics, and a reading of [12] and related work on information flow.

Recall the definition of an event system [5] $S = (E, I, O, T)$ where E is a set of events, and $I \subseteq E$ the input events and $O \subseteq E$ the output events. (Note that $I \cap O = \emptyset$. We also assume that $I \cup O = E$, in other words that all the events are either input events or output events.[1]) The set $T \subseteq E^*$ is the set of traces of the system. Given a trace $\tau \in E^*$ and a set $E' \subseteq E$, let $\tau|E'$ be the restriction of the trace τ to only events in E'.

We assume a partition of the events into low level events L and high level events H. An event in I and L is said to be alow level input event, and similarly for low level output events, high level input events, and high level output events. By definition, given a trace τ, the trace $\tau|L$ of the low level events happening in τ is the only observation that L can make about τ.

Intuitively, information is said to flow from the high level user to the low level user if when the low level user observes a sequence of low level events corresponding to some unknown trace τ, it can infer information about high level activity in τ. There are two notions to be formalized in such a statement: what is meant by "inferring information", and what is meant by "high level activity".

In [12], the following definition of information flow is used. (In fact, this definition does not appear in the paper, but is given in [11, p. 30].) There is information flowing from high to low in an event sytem S if and only if "the low level user's observation of $\tau|L$ implies that at least one high level event sequence or interleaving is not possible." In other words, information flows from high level users to low level users if there exists a high level trace or interleaving such that if it had occurred, then $\tau|L$ could not have occurred. This definition is intuitive, and one is tempted to attempt to formalize it directly. However, there is a caveat. As Zakinthinos argues [11, p.30]:

[1] It is not clear whether allowing internal events is an issue. Presumably, we could assume that internal events are simply output events for a more general definition of output.

> “Care must be taken in interpreting this statement. If low level actions influence high level behaviour then it is possible for a particular sequence not to be possible because the low level influence precludes it. However, in this case no inference is possible.”

(We note that the low level user is assumed to have complete knowledge of all the traces of the system.)

Intuitively, the definition of information therefore should be read as: there is information flowing from high to low if and only if the low level user’s observation of $\tau|L$ implies that at least one high level event sequence or interleaving *that as far as the low level user can figure out is allowed* is not possible.

Recasting the work on possibilistic information that [12] attempt to capture in their framework in the light of this definition, the various definitions of security reviewed in [12] can be understood as giving different ways of defining what the low level user can figure out should be allowed.

2 Information flow and knowledge

Let us give a simple logical framework in which to interpret the above intuitions. We fix a set Φ_0 of primitive propositions. The logic is a modal propositional logic with two modal operators K and P. Intuitively, $K\varphi$ means that the low level user knows that φ is true (in the system under consideration), while $P\varphi$ means that the low level user considers φ possible or allowable. The set of formulas Φ is the closure of Φ_0 under negation, conjunction and application of the K and P operators. We will sometimes be interested in the set of *propositional formulas* Φ^P, that is the closure of Φ_0 under negation and conjunction.

We assign a semantic to the logic by viewing event systems as Kripke structures, as follows. Define a *ZL structure* M_S corresponding to $S = (E, I, O, T)$ as a tuple $(P, T, \mathcal{K}, \pi)$, where P is the set of allowed traces of the system, and $T \subseteq P$ is the set of actual traces of the system (taken from S). The relation $\mathcal{K}$ is an equivalence relation on traces defined by $\mathcal{K}(\tau, \tau')$ iff $\tau|L = \tau'|L$. The function π assigns a truth-value to the primitive propositions in Φ_0, for every trace in P. The satisfiability relation $\models$ is a variation on the standard Kripke semantics:

$$(M, \tau) \models p \text{ iff } \pi(\tau)(p) = \textbf{true}$$

$$(M, \tau) \models \neg\varphi \text{ iff } (M, \tau) \not\models \varphi$$

$$(M, \tau) \models \varphi_1 \wedge \varphi_2 \text{ iff } (M, \tau) \models \varphi_1 \text{ and } (M, \tau) \models \varphi_2$$

$$(M, \tau) \models P\varphi \text{ iff } \forall \tau' \in \mathcal{K}(\tau), (M, \tau') \models \varphi$$

$$(M, \tau) \models K\varphi \text{ iff } \forall \tau' \in \mathcal{K}(\tau) \cap T, (M, \tau') \models \varphi$$

A knowledge-based interpretation of a security property with respect to information flow is to ensure that for a set of formulas $\Phi' \subseteq \Phi$ (intuitively, the set of formulas representing the information that should not be inferred by the low level user), a ZL structure $M = (P, T, \mathcal{K}, \pi)$ should be such that for all $\varphi \in \Phi'$, and for all $\tau \in T$ (that is, for all the traces that *actually exist* in the system), $(M, \tau) \models K\varphi \Rightarrow P\varphi$. (Intuitively, if the low level user knows φ, then it could have inferred φ by simply considering the traces that he considers allowed, as opposed to those that he knows could happen.)

A reasonable choice of Φ' to model the current work on information flow is to take Φ' to be Φ^P, the propositional formulas over Φ_0.

There are two advantages with this view of security. First, we separate the issues of the sequence of events happening in the system (the trace) from what the low level user can express (the language derived from Φ_0). This allows us to be much more precise in our statements: a system is secure with respect to the knowledge-based notion of information flow if there is no formula of interest (in the set Φ') that can distinguish between the real system and the allowed system. As we will see in the next section, this allows a system to be qualified of secure even if it does not satisfy a security property in its full generality. The second advantage is that we can readily extract the information leaked by a system deemed insecure at any given trace: a trace τ leaks information φ if $(M, \tau) \not\models K\varphi \Rightarrow P\varphi$, that is $(M, \tau) \models K\varphi \wedge \neg P\varphi$.

3 An interpretation of security properties

Many different notions of security have been proposed in the literature, each trying to capture the notion of information flow. In [12] and subsequent work (for example, [4]), such security properties are interpreted as closure properties over the set of traces in the system. In this section, we reinterpret those security properties in the logical framework of Section 2. As we will see, a security property will correspond to different assumptions about the *a priori* knowledge that the low level user has of the system under consideration.

For an event system S, we will view a security property as a mean of deriving the set P of allowed traces from the set T of actual traces in the system. Intuitively, since P represents the set of traces that the low level user considers allowed, this captures the knowledge of the low level user about the system. This *a priori* knowledge is extracted from the actual traces of the system, and not from some analysis of the system being modeled. As we shall see, this makes it difficult to describe precisely what some of the security properties actually mean.

We now review some "classical" security properties, which describe the minimal structure of P given T. We say structure $M = (P, T, \mathcal{K}, \pi)$ satisfies property X if the set P satisfies property X.

> **Noninference** [8]: For all $\tau \in P$, there is a $\tau' \in P$ such that $\tau|L = \tau'|L$ and $\tau'|H = \langle\rangle$. The intuition here is that the low level user does not know anything about the occurence of high level events; for any trace τ, he allows in the system a trace with the same low level sequence, but no high-level events.
>
> **Generalized noninference** [7]: For all $\tau \in P$, there is a $\tau' \in P$ such thath $\tau|L = \tau'|L$ and $\tau'|(H \cap I)$. This is similar to noninference, but here the low level user is only ignorant of the high level input events, while presumably being able to predict the output events.
>
> **Generalized noninterference** [5, 6]: For all $\tau \in P$ and for all τ' in $\mathit{interleave}((H \cap I)^*, \tau|L)$, there is a $\tau'' \in P$ such that $\tau''|L = \tau|L$ and $\tau''|(L \cup (H \cap I)) = \tau'$. This is a nondeterministic version of Goguen and Meseguer original noninterference property [1, 2]. It captures the intuition that the high level events are prevented to affect the low level events. It is not clear how to interpret this knowledge-wise. (In fact, the statement of generalized noninterference given here should probably be called strong generalized noninterference, since it is slightly more general than the definition of generalized noninterference given by McCullough.)

Non-deducible output [3]: ...

Separability [7]: ...

PSP [12]: ...

Given an event system with traces T, the security properties above attempt to infer a plausible notion of "what the low level user knows *a priori*", from which to derive the concept of information flow: information flows from high to low if, roughly speaking, the low level user knows something that he did not know *a priori*. For example, noninference captures the fact that the low level user knows nothing about high level events, while separability captures the fact that the low level user knows nothing about high level events *except* for the allowable sequences of high level events.

With this perspective, the various security properties appear rather arbitrary. (Mantel [4] dissects various security properties into basic security blocks, shedding some light on the various choices made.) In all fairness, those security properties also attempt to address the issue of *composition*, which we have not touched in this note.

To summarize, a system is deemed secure under possibilistic information flow (that is, a system does not exhibit information flow from high to low) if the low level user does not learn anything from witnessing the sequence of low level events of a given trace. To make this precise, we need to formalize what the low level user knows about the system. This formalization can be seen as the construction of P for a ZL structure.

4 Comments

1. Can we somehow formalize the intuition underlying all the security properties of Section 3?
2. Can we figure out ways of coming with P by "external" means?
3. Figure out the notions of unwinding and composability. How do they fit in the picture we're trying to paint?
4. Zakinthinos seems to have missed the boat at some point. In [11], he refers to a paper by Wittbold and Johnson [10] as containing a good discussion of why a previous theory of Sutherland [9] was inadequate for capturing information flow security. Zakinthinos seem-s to dismiss Wittbold and Johnson's conclusion that Sutherland's work was flawed because he did not consider strategies (in the game-theoretic sense, possibly). At the very least, that thread should be followed up on.

References

[1] J. A. Goguen and J. Meseguer. Security policies and security models. In *Proceedings of the 1982 IEEE Symposium on Research in Security and privacy*, pages 11–20. IEEE Computer Society Press, 1982.

[2] J. A. Goguen and J. Meseguer. Unwinding and inference control. In *Proceedings of the 1984 IEEE Symposium on Research in Security and Privacy*, pages 75–86. IEEE Computer Society Press, 1984.

[3] J. D. Guttman and M. E. Nadal. What needs securing? In *Proceedings of the Computer Security Foundations Workshop*, pages 34–57. IEEE Computer Society Press, 1988.

[4] H. Mantel. Possibilistic Definitions of Security—An Assembly Kit. In *Proceedings of the IEEE Computer Security Foundations Workshop*, pages 185–199, 2000.

[5] D. McCullough. Specifications for multi-level security and a hook-up property. In *IEEE 1987 Symposium on Security and Privacy*, pages 161–166. IEEE Computer Society Press, 1987.

[6] D. McCullough. Noninterference and the composability of security properties. In *Proceedings of the 1988 IEEE Symposium on Research in Security and Privacy*, pages 177–186. IEEE Computer Society Press, 1988.

[7] J. McLean. A general theory of composition for trace sets closed under selective interleaving functions. In *Proc. IEEE Symposium on Security and Privacy*, pages 79–93. 1994.

[8] C. O'Halloran. A calculus of information flow. In *Proceedings of the European Symposium on Research in Computer Security*, 1990.

[9] D. Sutherland. A model of information. In *Proceedings of the 9th National Computer Security Conference*, pages 175–183, 1986.

[10] J. T. Wittbold and D. Johnson. Information flow in nondeterministic systems. In *Proceedings of the 1990 IEEE Symposium on Research in Security and Privacy*. IEEE Computer Society Press, 1990.

[11] A. Zakinthinos. *On the Composition of Security Properties.* PhD thesis, University of Toronto, 1996.

[12] A. Zakinthinos and E. S. Lee. A general theory of security properties. In *Proc. IEEE Symposium on Security and Privacy*, pages 94–102. 1997.