On Notions of Causality and Distributed Knowledge*

Ron van der Meyden
University of New South Wales
meyden@cse.unsw.edu.au

Abstract

The notion of distributed knowledge is used to express what a group of agents would know if they were to combine their information. The paper considers the application of this notion to systems in which there are constraints on how an agent's actions may cause changes to another agent's observations. Intuitively, in such a setting, one would like that anything an agent knows about other agents must be distributed knowledge to the agents that can causally affect it. In prior work, we have argued that the definition of intransitive noninterference — a notion of causality used in the literature on computer security — is flawed because it fails to satisfy this property, and have proposed alternate definitions of causality that we have shown to be better behaved with respect to the theory of intransitive noninterference. In this paper we refine this understanding, and show that in order for the converse of the property to hold, one also needs a novel notion of distributed knowledge, as well as a new notion of what it means for a proposition to be "about" other agents.

Introduction

It is a commonly held intuition that information flows along causal lines: where there is no causal relationship, there will be no flow of information. In this paper, we attempt to give a precise characterization of this intuition using notions drawn from the literature on computer security (where lack of causality is called *noninterference*, and is considered in dealing with information flow security) and the literature on epistemic logic.

In particular, we start with the following idea. Suppose that a system is structured so that the only way that an agent u may be causally affected by the outside world is through the activity of a set of "interfering" agents I_u. Then any information that u has about the outside world, it must have obtained by somehow combining the information that it received from agents I_u. The notion of *distributed knowledge* is used in the epistemic logic literature to capture what could be deduced if we were to combine all the information available to a group of agents. Thus, we may express our intuition by saying that anything that u knows about the outside world must be distributed knowledge to I_u.

In previous work (van der Meyden 2007), we have used this formulation of the intuition to argue that *intransitive noninterference*, a notion of causality from the computer security literature, is flawed, and have proposed a number of other definitions of causality in response to this failure. We justifed these new definitions in that work primarily by showing that they lead to a more satisfactory account of the classical proof theory and applications for intransitive noninterference.

In this paper, we show that these new notions in fact satisfy the intuition much more generally than in the single example considered in (van der Meyden 2007), and give a refined statement of how this is the case. In doing so, we argue that distributed knowledge, as it has been defined in the epistemic logic literature, is not the only useful way to formalise the intuitive notion of what a group of agents would know if they were to combine their information. We develop several new distributed-knowledge like modalities for our application.

However, we also attempt to do more than show that the new notions of distributed knowledge satisfy the intuition concerning an agent's knowledge and the distributed knowledge of its interferers. We also consider the converse relationship between knowledge and causality. That is, we ask whether it is the case that when the expected relationship between the knowledge of an agent "about other agents" and the distributed knowledge of its interferers holds, we can conclude that the system has the expected causal structure.

We show that this converse relationship does in fact hold, provided one uses appropriate formulations of distributed knowledge and what it means for an agent to know something "about" other agents. Some of the details are subtle, and give fresh insight into how information flows in the class of systems we consider.

Intransitive Noninterference

We begin by recalling some notions of causality used in the computer security literature.

The definitions are cast in terms of a formal semantic model for multi-agent systems. Several different models have been used in the literature; we follow the state-observed machine formulation of (Rushby 1992). This

*Thanks to the Computer Science Department, Stanford University for hosting a sabbatical visit during which some of this research was conducted. Work of the author supported by an Australian Research Council Discovery grant.

model consists of deterministic machines of the form $\langle S, s_0, A, \mathtt{step}, \mathtt{obs}, \mathtt{dom}\rangle$, where S is a set of states, $s_0 \in S$ is the *initial state*, A is a set of actions, $\mathtt{dom} : A \to D$ associates each action to an element of the set D of agents, $\mathtt{step} : S \times A \to S$ is a deterministic transition function, and $\mathtt{obs} : S \times D \to O$ maps states to an observation in some set O, for each agent.

The nomenclature D and $\mathtt{dom}$ for agents arises from the fact that, in the security literature, D is though of as the set of *security domains*. For example, in a multi-level secure system, a security domain is a pair consisting of of a security level (e.g. Low, High, Secret) together with a set of classes (e.g. Nuclear, NATO, FBI) used to restrict information access on a need-to-know basis. Several agents might be assigned to such a security domain in this interpretation. For our purposes in this paper, we will think of each element of D as a single agent, since this better fits the perspective that we apply from the logic of knowledge.

We write $s \cdot \alpha$ for the state reached by performing the sequence of actions $\alpha \in A^*$ from state s, defined inductively by $s \cdot \epsilon = s$, and $s \cdot \alpha a = \mathtt{step}(s \cdot \alpha, a)$ for $\alpha \in A^*$ and $a \in A$. Here ϵ denotes the empty sequence.

Permitted causal relationships are expressed in the security literature using *noninterference policies*, wich are relations $\rightarrowtail \subseteq D \times D$, with $u \rightarrowtail v$ intuitively meaning that "actions of agent u are permitted to interfere with agent v", or "information is permitted to flow from agent u to agent v". Since, intuitively, an agent should be allowed to interfere with, or have information about, itself, this relation is assumed to be reflexive.[1]

Noninterference was given a formal semantics for transitive noninterference policies (which arise naturally from partially ordered security levels) by Goguen and Meseguer (Goguen and Meseguer 1982), using a definition based on a "purge" function. Given a set $E \subseteq D$ of agents and a sequence $\alpha \in A^*$, we write $\alpha \upharpoonright E$ for the subsequence of all actions a in α with $\mathtt{dom}(a) \in E$. Given a policy $\rightarrowtail$, we define the function $\mathtt{purge} : A^* \times D \to A^*$ by

$$\mathtt{purge}(\alpha, u) = \alpha \upharpoonright \{v \in D \mid v \rightarrowtail u\}.$$

(For clarity, we may use subscripting of agent arguments of functions, writing e.g., $\mathtt{purge}(\alpha, u)$ as $\mathtt{purge}_u(\alpha)$.)

Definition 1: A system M is *P-secure* with respect to a policy $\rightarrowtail$ if for all agents u and for all sequences $\alpha, \alpha' \in A^*$ such that $\mathtt{purge}_u(\alpha) = \mathtt{purge}_u(\alpha')$, we have $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. □

This can be understood as saying that agent u's observation depends only on the sequence of interfering actions that have been performed.

This definition is appropriate when the noninterference policy is transitive, but it has been considered to be inappropriate for the intransitive case. An example of this is systems consisting of a High security agent H, a low security agent L, and a downgrader agent D, whose role is to make declassification decisions that release High security information to L. Here the policy is $H \rightarrowtail D \rightarrowtail L$.[2] P-security says that L can learn about D actions, but will never know anything about H actions. Thus, in a P-secure system, L will not know about H activity even if D has chosen to downgrade it. For example, if h, d are actions of H, D, respectively, then $\mathtt{purge}_L(hdh) = d = \mathtt{purge}_L(d)$. Thus, according this definition, L cannot distinguish the sequences hdh and d, so D is unable to downgrade the fact that action h has occurred.

To avoid this problem, Haigh and Young (Haigh and Young 1987) generalised the definition of the purge function to intransitive policies as follows. Intuitively, the intransitive purge of a sequence of actions with respect to a domain u is the largest subsequence of actions that could form part of a causal chain of effects (permitted by the policy) ending with an effect on domain u. More formally, the definition makes use of a function $\mathtt{sources} : A^* \times D \Rightarrow \mathcal{P}(D)$ defined inductively by $\mathtt{sources}(\epsilon, u) = \{u\}$ and

$$\begin{aligned}&\mathtt{sources}(a\alpha, u)\\ &\quad = \mathtt{sources}(\alpha, u) \cup\\ &\qquad \{\mathtt{dom}(a) \mid \exists v \in \mathtt{sources}(\alpha, u)(\mathtt{dom}(a) \rightarrowtail v)\}\end{aligned}$$

for $a \in A$ and $\alpha \in A^*$. Intuitively, $\mathtt{sources}(\alpha, u)$ is the set of domains v such that there exists a sequence of permitted interferences from v to u within α. The *intransitive purge* function $\mathtt{ipurge} : A^* \times D \to A^*$ is then defined inductively by $\mathtt{ipurge}(\epsilon, u) = \epsilon$ and

$$\mathtt{ipurge}(a\alpha, u) = \begin{cases} a \cdot \mathtt{ipurge}(\alpha, u) & \text{if } \mathtt{dom}(a) \in \mathtt{sources}(a\alpha, u) \\ \mathtt{ipurge}(\alpha, u) & \text{otherwise} \end{cases}$$

for $a \in A$ and $\alpha \in A^*$. The intransitive purge function is then used in place of the purge function in Haigh and Young's definition:

Definition 2: M is IP-secure with respect to a policy $\rightarrowtail$ if for all $u \in D$ and all sequences $\alpha, \alpha' \in A^*$ with $\mathtt{ipurge}_u(\alpha) = \mathtt{ipurge}_u(\alpha')$, we have $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. □

It can be seen that $\mathtt{ipurge}_u(\alpha) = \mathtt{purge}_u(\alpha)$ when $\rightarrowtail$ is transitive, so IP-security is in fact a generalisation of the definition of security for transitive policies.

This definition solves the problem noted above, since now we have $\mathtt{ipurge}_L(hdh) = hd$. Here we see that L can now learn of the first occurrence of h. (The second h remains invisible to L. This is in accordance with the intent of the definition - L should only know of H events that have been explicitly downgraded.)

Knowledge and Distributed Knowledge

We have recently presented an argument against the definition of intransitive noninterfere (van der Meyden 2007), in

[1] We follow the terminology historically used in the area even though it is peculiar: note that the relation $\rightarrowtail$ specifies permitted interferences rather than required noninterferences, and when we speak of "intransitive noninterference", we mean that noninterference policies $\rightarrowtail$ are not assumed to be transitive (rather than assumed to be not transitive).

[2] When presenting policies we list only the nonreflexive relations; the policy should be taken to be the reflexive closure of the facts given.

which we exploit intuitions from the literature on the logic of knowledge (Fagin et al. 1995). In this section, we recall the relevant background from the latter area.

Let $Prop$ be a set of atomic propositions. We define a propositional modal logic based on a set Op of monadic modal operators (to be introduced below), with formulas defined as follows: if $p \in Prop$ then p is formula, and if ϕ and ψ are formulas and $X \in Op$ is a modal operator, then $\neg\phi$, $\phi \vee \psi$ and $X\phi$ are formulas. We use standard boolean abbreviations such as $\phi \Rightarrow \psi$ for $\neg\phi \vee \psi$. If $u \in D$ is an agent and $G \subseteq D$ is a nonempty set of agents, the set Op will contain the operators $\mathtt{K}_u$ and $\mathtt{D}_G$. Intuitively, the formula $\mathtt{K}_u\phi$ expresses that agent u knows ϕ, and $\mathtt{D}_G\phi$ expresses that ϕ is *distributed knowledge* to the group G, which means that the group G, collectively, knows ϕ. (We introduce additional operators below.)

We take the atomic propositions to be interpreted over sequences of actions of a system M with actions A, by means of an interpretation function $\pi : Prop \to \mathcal{P}(A^*)$. Formulas ϕ are then interpreted as being satisfied at a sequence of actions[3] $\alpha \in A^*$ by means of a relation $M, \pi, \alpha \models \phi$. For atomic propositions $p \in Prop$, this relation is defined by $M, \pi, \alpha \models p$ if $\alpha \in \pi(p)$.

The semantics of the operators for knowledge is defined using the following notion of *view*. The definition uses an absorbtive concatenation function $\circ$, defined over a set X by, for $s \in X^*$ and $x \in X$, by $s \circ x = s$ if x is equal to the final element of s (if any), and $s \circ x = s \cdot x$ (ordinary concatenation) otherwise. The view of agent u with respect to a sequence $\alpha \in A^*$ is captured using the function $\mathtt{view}_u : A^* \to (A \cup O)^*$ (where O is the set of observations in the system), defined by

$$\begin{array}{rcl} \mathtt{view}_u(\epsilon) & = & \mathtt{obs}_u(s_0), \text{ and} \\ \mathtt{view}_u(\alpha a) & = & (\mathtt{view}_u(\alpha) \cdot b) \circ \mathtt{obs}_u(s_0 \cdot \alpha), \end{array}$$

where $b = a$ if $\mathtt{dom}(a) = u$ and $b = \epsilon$ otherwise. That is, $\mathtt{view}_u(\alpha)$ is the sequence of all observations and actions of domain u in the run generated by α, compressed by the elimination of stuttering observations. Intuitively, $\mathtt{view}_u(\alpha)$ is the complete record of information available to agent u in the run generated by the sequence of actions α. The reason we apply the absorbtive concatenation is to capture that the system is asynchronous, with agents not having access to a global clock. Thus, two periods of different length during which a particular observation obtains are not distinguishable to the agent.

Using the notion of view, we may define for each agent u an equivalence relation $\sim_u$ on sequences of actions by $\alpha \sim_u \alpha'$ if $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\alpha')$. The semantics for the knowledge operators may then be given by

$M, \pi, \alpha \models \mathtt{K}_u\phi$ if $M, \pi, \alpha' \models \phi$ for all sequences α' such that $\alpha \sim_u \alpha'$.

[3]The reader may have expected to see propositions interpreted more generally at sequences alternating states and actions. We remark that there is in fact no loss of generality, because the assumption that actions are deterministic means that the sequence of states in a run is uniquely determined by the sequence of actions.

This is essentially the definition of knowledge used in the literature on reasoning about knowledge (Fagin et al. 1995) for an agent with *asynchronous perfect recall* .

The notion of distributed knowledge is defined in the literature as follows. First, define the relations $\sim_G$ on sequences of actions by $\alpha \sim_G \alpha'$ if $\alpha \sim_u \alpha'$ for all $u \in G$. The operators $\mathtt{D}_G$ are then given semantics by the clause

$M, \pi, \alpha \models \mathtt{D}_G\phi$ if $M, \pi, \alpha' \models \phi$ for all sequences α' such that $\alpha \sim_G \alpha'$.

Intuitively, this definition says that a fact is *distributed knowledge* to the set of agents G if it could be deduced after combining all the information that these agents have. Note that combination of the agents' information is here being modelled as the intersection of their "information sets" — we will argue below that there are alternatives to this definition that are both reasonable and useful.

The Key Intuition

Suppose that a system is secure with respect to a noninterference policy $\rightarrowtail$. Given an agent u, let $I_u = \{v \in D \mid v \rightarrowtail u, v \neq u\}$ be the set of all other agents that are permitted to interfere with u. Let p be a proposition that expresses a fact about some agent other than u, and suppose u knows p. Intuitively, if the system respects the noninterference policy, then since p is not "local information", the only way that u should be able to learn that p holds is by receiving information from the agents I_u. However, u may have deduced p by combining facts received from several sources. Since we have assumed agents have perfect recall, those sources should also know those facts. Hence, if we combine the information known to agents in I_u, then we should also be able to deduce p. Thus, we expect that if the system M satisfies the policy $\rightarrowtail$ and p is interpreted by π as being about agents other than u, then

$$M, \pi, \alpha \models K_u p \Rightarrow D_{I_u} p \tag{1}$$

for all $\alpha \in A^*$.

In (van der Meyden 2007), we presented an argument against IP-security on the grounds that this intuition can be false when we interpret "the system M satisfies the policy $\rightarrowtail$" as saying that M is IP-secure with respect to $\rightarrowtail$. The essential reason for this is that the intransitive purge $\mathtt{ipurge}_L(\alpha)$ preserves not just certain actions from the sequence α, but also their *order*. This allows L to "know" this order in situations where an intuitive reading of the policy would suggest that it ought not to know this order. The following reproduces the example that we used in (van der Meyden 2007) to show that IP-security does not satisfy the intuition concerning the relationship between causality and distributed knowledge.

Example 1: Consider the intransitive policy $\rightarrowtail$ given by $H_1 \rightarrowtail D_1$, $H_2 \rightarrowtail D_2$, $D_1 \rightarrowtail L$ and $D_2 \rightarrowtail L$. Intuitively, H_1, H_2 are two High security domains, D_1, D_2 are two downgraders, and L is an aggregator of downgraded information. For this policy, we have $I_L = \{D_1, D_2\}$ so (1) requires that

$$M, \pi, \alpha \models \mathtt{K}_L p \Rightarrow \mathtt{D}_{\{D_1, D_2\}} p \tag{2}$$

for all $\alpha \in A^*$. We show that if security is interpreted as IP-security, then this can be false.

Define the system M with actions $A = \{h_1, h_2, d_1, d_2, l\}$ with domains H_1, H_2, D_1, D_2, and L, respectively. The set of states of M is the set of all strings in A^*. The transition function is defined by concatenation, i.e. for a state $\alpha \in A^*$ and an action $a \in A$, $\mathtt{step}(\alpha, a) = \alpha a$. The observation functions are defined using the ipurge function associated to the above policy: $\mathtt{obs}_u(\alpha) = [\mathtt{ipurge}(\alpha, u)]$. (Here we put brackets around the sequence of actions when it is interpreted as an observation, to distinguish such occurrences from the actions themselves as they occur in a view.)

It is plain that M is IP-secure. For, if $\mathtt{ipurge}(\alpha, u) = \mathtt{ipurge}(\alpha', u)$ then $\mathtt{obs}_u(s_0 \cdot \alpha) = [\mathtt{ipurge}(\alpha, u)] = [\mathtt{ipurge}(\alpha', u)] = \mathtt{obs}_u(s_0 \cdot \alpha')$. We show that it does not satisfy condition (2).

Consider the sequences of actions $\alpha_1 = h_1 h_2 d_1 d_2$ and $\alpha_2 = h_2 h_1 d_1 d_2$. Note that these differ in the order of the events h_1, h_2. Let the atomic proposition p be interpreted by π as asserting that there is an occurrence of h_1 before an occurrence of h_2. That is, $\pi(p) = \{\alpha h_1 \beta h_2 \gamma \mid \alpha, \beta, \gamma \in A^*\}$.

Then we have $\mathtt{obs}_L(\alpha_1) = [\mathtt{ipurge}(\alpha_1, L)] = [h_1 h_2 d_1 d_2]$. Hence, for any sequence α' with $\alpha_1 \sim_L \alpha'$, we have $\mathtt{ipurge}_L(\alpha') = h_1 h_2 d_1 d_2$, so $\alpha' \in \pi(p)$. Thus $M, \pi, \alpha_1 \models \mathtt{K}_L p$, i.e., in α_1 agent L knows the ordering of the events h_1, h_2. We demonstrate that α_2 is a witness showing that it is not the case that $M, \pi, \alpha_1 \models \mathtt{D}_{\{D_1, D_2\}} p$, i.e., we have $\alpha_1 \sim_{\{D_1, D_2\}} \alpha_2$ and $\alpha_2 \notin \pi(p)$. The latter is trivial. For the former, note

$$\begin{aligned}
&\mathtt{view}_{D_1}(\alpha_1) \\
&\quad = [\epsilon] \circ [h_1] \circ [h_1] \circ d_1 \circ [h_1 d_1] \circ [h_1 d_1] \\
&\quad = [\epsilon] \circ [\epsilon] \circ [h_1] \circ d_1 \circ [h_1 d_1] \circ [h_1 d_1] \\
&\quad = \mathtt{view}_{D_1}(\alpha_2)
\end{aligned}$$

i.e., $\alpha_1 \sim_{D_1} \alpha_2$. By symmetry, we also have $\alpha_1 \sim_{D_2} \alpha_2$, hence $\alpha_1 \sim_{\{D_1, D_2\}} \alpha_2$. This means that D_1 and D_2 do not have distributed knowledge of the ordering of the events h_1, h_2, even with respect to the asynchronous perfect recall intepretation of knowledge, in which they reason based on everything that they learn in the run.

Thus, L has acquired information that cannot have come from the two sources D_1 and D_2 that are supposed to be, according to the policy, its only sources of information.□

Alternate Definitions of Causality

In response to the example of the previous section, we have defined in (van der Meyden 2007) several alternative notions of causality/security that are not only better behaved than IP-security with respect to the example, but also prove to be a much better fit to proof techniques and applications that had been developed for intransitive noninterference. We will not go into the latter here, but confine ourselves to stating the definitions of our alternative notions of causality.

All the alternatives are based on a concrete model of the maximal amount of information that an agent may have after some sequence of actions has been performed, and state that an agent's observation may not give it more than this maximal amount of information. The definitions differ in the modelling of the maximal information, but all take the view that an agent increases its information either by performing an action or by receiving information transmitted by another agent.

In the first model of the maximal information, what is transmitted when an agent performs an action is information about the actions performed by other agents. The following definition expresses this in a weaker way than the ipurge function.

Given sets X and A, let the set $\mathcal{T}(X, A)$ be the smallest set containing X and such that if $x, y \in \mathcal{T}$ and $z \in A$ then $(x, y, z) \in \mathcal{T}$. Intuitively, the elements of $\mathcal{T}(X, A)$ are are binary trees with leaves labelled from X and interior nodes labelled from A.

Given a policy $\rightarrowtail$, define, for each agent $u \in D$, the function $\mathtt{ta}_u : A^* \to \mathcal{T}(\{\epsilon\}, A)$ inductively by $\mathtt{ta}_u(\epsilon) = \epsilon$, and, for $\alpha \in A^*$ and $a \in A$,

1. if $\mathtt{dom}(a) \not\rightarrowtail u$, then $\mathtt{ta}_u(\alpha a) = \mathtt{ta}_u(\alpha)$,
2. if $\mathtt{dom}(a) \rightarrowtail u$, then $\mathtt{ta}_u(\alpha a) = (\mathtt{ta}_u(\alpha), \mathtt{ta}_{\mathtt{dom}(a)}(\alpha), a)$.

Intuitively, $\mathtt{ta}_u(\alpha)$ captures the maximal information that agent u may, consistently with the policy $\rightarrowtail$, have about the past actions of other agents. (The nomenclature is intended to be suggestive of *transmission* of information about *actions*.) Initially, an agent has no information about what actions have been performed. The recursive clause describes how the maximal information $\mathtt{ta}_u(\alpha)$ permitted to u after the performance of α changes when the next action a is performed. If a may not interfere with u, then there is no change, otherwise, u's maximal permitted information is increased by adding the maximal information permitted to $\mathtt{dom}(a)$ at the time a is performed (represented by $\mathtt{ta}_{\mathtt{dom}(a)}(\alpha)$), as well the fact that a has been performed. Thus, this definition captures the intuition that an agent may only transmit information that it is permitted to have, and then only to agents with which it is permitted to interfere.

Definition 3: A system M is TA-secure with respect to a policy $\rightarrowtail$ if for all agents u and all $\alpha, \alpha' \in A^*$ such that $\mathtt{ta}_u(\alpha) = \mathtt{ta}_u(\alpha')$, we have $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. □

Intuitively, this says that each agent's observations provide the agent with no more than the maximal amount of information that may have been transmitted to it, as expressed by the functions ta.

Like IP-security, the definition of TA-security views it as permissible for an agent to cause the transmission of information that it is permitted to have, but has not itself observed. For example, an agent may forward an email attachment without inspecting it. For at least some applications (e.g., the downgrading example mentioned above) this is undesirable. To prohibit such behaviour, a second alternative definition of causality from (van der Meyden 2007) restricts the transmitted information to that which has been observed. This alternative is defined as follows. Given a policy $\rightarrowtail$, for each domain $u \in D$, define the function

$\mathtt{to}_u : A^* \to \mathcal{T}(O(A \cup O)^*, A)$ by $\mathtt{to}_u(\epsilon) = \mathtt{obs}_u(s_0)$ and

$$\mathtt{to}_u(\alpha a) = \begin{cases} \mathtt{to}_u(\alpha) & \text{if } \mathtt{dom}(a) \not\rightarrowtail u, \\ (\mathtt{to}_u(\alpha), \mathtt{view}_{\mathtt{dom}(a)}(\alpha), a) & \text{otherwise.} \end{cases}$$

Intuitively, this definition takes the model of the maximal information that an action a may transmit after the sequence α to be the fact that a has occurred, together with the information that $\mathtt{dom}(a)$ *actually* has, as represented by its view $\mathtt{view}_{\mathtt{dom}(a)}(\alpha)$. By contrast, TA-security uses in place of this the maximal information that $\mathtt{dom}(a)$ *may* have. (The nomenclature '`to`' is intended to be suggestive of *transmission* of information about *observations*.)

We may now base the definition of security on either the function `to` or `ito` rather than `ta`.

Definition 4: The system M is TO-secure with respect to $\rightarrowtail$ if for all domains $u \in D$ and all $\alpha, \alpha' \in A^*$ with $\mathtt{to}_u(\alpha) = \mathtt{to}_u(\alpha')$, we have $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. □

The following result shows how these definitions are related:

Theorem 1 *(van der Meyden 2007) For state-observed systems, with respect to a given policy $\rightarrowtail$, P-security implies TO-security implies TA-security implies IP-security.*

Examples showing that all these notions are distinct are presented in (van der Meyden 2007).

This completes the background for the contributions of the present paper. In what follows, we show that these new definitions of security can in fact be shown to be closely related to our intuitions about distributed knowledge in such settings, provided we also develop some new notions of distributed knowledge. In order to capture common structure in our results, it is useful to take a more abstract perspective on the above definitions.

Define a *local-state assignment* in a system M to be a function $Y : A^* \times D \to L$, where L is some set. As usual, we write $Y_u(\alpha)$ for $Y(\alpha, u)$. The functions `view`, `ta` and `to` are all examples of local-state assignments.

Let X and Y be local-state assignments. We say that X is a *full-information local-state assignment based on Y and $\rightarrowtail$* if the following holds: $X_u(\epsilon) = \epsilon$ and for all agents $u \in D$, sequences $\alpha \in A^*$ and $a \in A$, we have

$$X_u(\alpha a) = \begin{cases} X_u(\alpha) & \text{if } \mathtt{dom}(a) \not\rightarrowtail u \\ (X_u(\alpha), Y_{\mathtt{dom}(a)}(\alpha), a) & \text{otherwise.} \end{cases}$$

That is, according to the local state assignment X, the effect of an action a after a sequence α is to transmit to all domains u with $\mathtt{dom}(a) \rightarrowtail u$ the information that the action a has been performed, as well as all the information in the local state $Y_{\mathtt{dom}(a)}(\alpha)$. Such domains u add this new information to the information $X_u(\alpha)$ already collected. The action has no effect on domains with which it is not permitted to interfere.

We can now identify some common structure in the above definitions by noting that `to` is a full-information local-state assignment with respect to `view` and $\rightarrowtail$, and `ta` is a full-information local-state assignment with respect to `ta` and $\rightarrowtail$.

The Example Revisited - A First Concern

We have presented Example 1 in terms of the notion of distributed knowledge as it is usually defined in the literature. We now note that it is reasonable to object to the use of this notion in the present context. Note that whereas L is, intuitively, permitted by the policy to observe the ordering of actions in the domains D_1, D_2 (and actually does observe this order in the example), the way that the definition of distributed knowledge combines the information available to D_1 and D_2 does not take into account the relative ordering of the actions in these domains. Indeed, the following example illustrates that to ask that information known to L be distributed knowledge to D_1 and D_2 may be too strong.

Example 2: Let the (transitive) policy $\rightarrowtail$ be defined by $L_1 \rightarrowtail H$ and $L_2 \rightarrowtail H$. Consider a system with actions $A = \{l_1, l_2, h\}$ of domains L_1, L_2, H, respectively, and let the states and transitions of M be given by $S = A^*$ and $\mathtt{next}(\alpha, a) = \alpha \cdot a$ as in Example 1, but define the observations by $\mathtt{obs}_{L_i}(\alpha) = \mathtt{purge}_{L_i}(\alpha)$ and $\mathtt{obs}_H(\alpha) = \alpha$. This system is intuitively secure, and is easily seen to be P-secure (hence secure for all the other definitions) but, taking p to mean "there is an occurrence of l_1 before l_2", we have $M, \pi, l_1 l_2 \models \mathtt{K}_H p$ but not $M, \pi, l_1 l_2 \models \mathtt{D}_{\{L_1, L_2\}} p$, since $l_1 l_2 \sim_{\{L_1, L_2\}} l_2 l_1$. □

The appropriate diagnosis here seems to be that distributed knowledge is too strong a notion for our present purposes. We are therefore motivated to define a variant, that takes into account H's observational powers in this example. Define the equivalence relation $\sim^p_G$ on A^* by $\alpha \sim^p_G \alpha'$ if $\alpha \sim_G \alpha'$ and $\alpha \restriction G = \alpha' \restriction G$. Intuitively, this relation combines the information in views of the agents G, not just by intersecting the information sets, but also taking into account the ordering of the actions in these views in the actual run. Extend the modal language by adding an operator $\mathtt{D}^p_G$, with semantics given by

$M, \pi, \alpha \models \mathtt{D}^p_G \phi$ if $M, \pi, \alpha' \models \phi$ for all $\alpha' \in A^*$ with $\alpha \sim^p_G \alpha'$.

Note that $\mathtt{D}_G \phi \Rightarrow \mathtt{D}^p_G \phi$, so this is a weaker notion of distributed knowledge.

Example 3: Reconsidering Example 2, we see that $l_1 l_2 \sim^p_{\{L_1, L_2\}} \alpha$ iff $\alpha \in \{h\}^* \cdot l_1 \cdot \{h\}^* \cdot l_2 \cdot \{h\}^*$, hence $M, \pi, l_1 l_2 \models \mathtt{D}^p_{\{L_1, L_2\}} p$. Thus, with this new interpretation of distributed knowledge, we recover the desired intuition in this example. □

Using this variant notion of distributed knowledge, the problem identified in the example from (van der Meyden 2007) can be shown to persist.

Example 4: Let the system M, the interpretation π of the proposition p and the sequences α_1 and α_2 be as in Example 1. Since $\alpha_1 \restriction \{D_1, D_2\} = d_1 d_2 = \alpha_2 \restriction \{D_1, D_2\}$, we have $\alpha_1 \sim^p_{\{D_1, D_2\}} \alpha_2$, so $M, \pi, \alpha_1 \models \mathtt{K}_L p$ but not $M, \pi, \alpha_1 \models \mathtt{D}^p_{\{D_1, D_2\}} p$. □

We therefore conclude that although the concern raised

above is valid, the problem we have identified with the definition of IP-security is real. Nevertheless, we would like to have a better understanding concerning the intuition than can be obtained from a single example. We pursue this in the following sections.

From Causality to Distributed Knowledge

Based on a problem identified in Example 1, we have constructed some alternative definitions for security with respect to intransitive policies, reflecting intuitions about the transmission of information in a concrete setting. These definitions were justified in (van der Meyden 2007) on the grounds that they lead to a better account of the classical proof theory and applications for intransitive noninterference. We now consider what these new definitions say about our motivating example. Indeed, we show that they support our intuition not just in this example, but more generally.

As discussed above, our key intuition is that if the system is secure, then any information that u has about other agents must be deducible from the information passed to it by the agents I_u, so should be distributed knowledge, in some sense, to I_u. We now check how each of our definitions fares with respect to this intuition. One qualification will be required: note that if u is able to have an effect on the agents I_u, then it may combine its knowledge of these effects with the distributed knowledge of I_u to deduce facts that are not distributed knowledge to I_u alone. We therefore require that u not be permitted to interfere, directly or indirectly, with I_u. This may be captured by the requirement that u not be part of any nontrivial cycle in $\rightarrowtail$.

To formalise the intuitive notion of "information about other agents", we use the following notion. For $F \subseteq D$ a set of agents, say that a property $\Pi \subseteq A^*$ *depends only on* F if for all $\alpha, \alpha' \in A^*$ with $\alpha \restriction F = \alpha' \restriction F$ we have $\alpha \in \Pi$ iff $\alpha' \in \Pi$. Similarly, say that an atomic proposition p is interpreted by π as *depending only on* F if $\pi(p)$ depends only on F. This expresses more precisely the fact that the proposition p is "about the agents F".

As already noted in the previous section, the appropriate notion of distributed knowledge to capture our intuition in the current semantic setting is one that takes into account the interleaving of the actions of the group. As different definitions of causality require different notions of distributed knowledge to capture the intuition, we generalize the operator $\mathtt{D}^p$ as follows. Suppose that Y is a local-state assignment in M and G is a set of agents. Define the relation $\sim^Y_G$ on sequences of actions in M by $\alpha \sim^Y_G \alpha'$ if $\alpha \restriction G = \alpha' \restriction G$ and $Y_v(\alpha) = Y_v(\alpha')$ for all $v \in G$, Then we define the new modal operator $\mathtt{D}^Y_G$ with semantics

$M, \pi, \alpha \models \mathtt{D}^Y_G\phi$ if for all $\alpha' \in A^*$ such that $\alpha \sim^Y_G \alpha'$ we have $M, \pi, \alpha' \models \phi$.

Intuitively, this is just the notion of distributed knowledge $\mathtt{D}^p_G$, except that instead of combining the information in the local states $\mathtt{view}_v(\alpha)$ for $v \in G$ relative to the interleaving of G actions in α, we combine the information in the local states $Y_v(\alpha)$. Indeed, it is easily seen that $\mathtt{D}^p_G\phi$ is equivalent to $\mathtt{D}^{\mathtt{view}}_G\phi$.

We may now express the key intuition abstractly using the following definition.

Definition 5: A system M is *confined with respect to a local-state assignment Y and noninterference policy $\rightarrowtail$* if for all sequences of actions $\alpha \in A^*$ and all agents u, if π interprets p as depending only on $D \setminus u$, then $M, \pi, \alpha \models \mathtt{K}_u p \Rightarrow \mathtt{D}^Y_{I_u} p$, where $I_u = \{v \in D \mid u \neq v,\ v \rightarrowtail u\}$. □

We would like to have that if a system is secure for a given notion of security, then it is confined with respect to an appropriate local-state assignment. In order to identify these local state assignments, note that, given the way that information is transmitted in the definitions of `ta` and `to`, an agent $v \in I_u$ may have acquired new information after it last transmitted information to u. Whereas u is not expected to have this information, the distributed knowledge of I_u takes this into account. Thus, in some sense, the relation $\sim_{I_u}$ takes *more* information from I_u into account than is required. We therefore develop a construct that helps to identify the latest point at which an agent may have transmitted information to other agents.

For each domain u define the mapping $m_u : A^* \rightarrow A^*$ so that $m_u(\alpha)$ is the prefix of α up to but excluding the rightmost action of agent u. More precisely, the definition is given inductively by $m_u(\epsilon) = \epsilon$ and $m_u(\alpha a) = m_u(\alpha)$ if $\mathtt{dom}(a) \neq u$ and $m_u(\alpha a) = \alpha$ otherwise. Intuitively, $m_u(\alpha)$ expresses the history of system at the point that the latest action of u occurs in α. If Y is a local-state assignment, we write $Y \circ m$ for the local-state assignment defined by $(Y \circ m)_u(\alpha) = Y_u(m_u(\alpha))$.

Consider now the relation $\sim^{TO}_G$ and the operator $\mathtt{D}^{TO}_G$, defined to be $\sim^Y_G$ and $\mathtt{D}^Y_G\phi$, respectively, where $Y = \mathtt{view} \circ m$. Intuitively, $\mathtt{D}^{TO}_G\phi$ says that ϕ is deducible by an agent that has all the information that was available to agents in G — at the time that they performed their latest action — together with the ordering of all the actions that have been performed by these agents. It is easily seen that this is a stronger notion of distributed knowledge than $\mathtt{D}^p_G$, in the sense that the formula $\mathtt{D}^{TO}_G\phi \Rightarrow \mathtt{D}^p_G\phi$ is valid. This notion supports our key intuition:

Theorem 2 *Suppose that $\rightarrowtail$ is acyclic. If M is TO-secure with respect to $\rightarrowtail$ then M is confined with respect to the local-state assignment* $\mathtt{view} \circ m$ *and $\rightarrowtail$.*

In order to obtain a similar result for TA-security, we need to take into account that TA-security is consistent with the transmission of information that is not contained in the sender's view. Thus, the best that we can expect in this case is that information known to u must be *permitted* to be distributed knowledge to the agents that may transmit information to u. As with the notion $\mathtt{D}^{TO}_G$, we also restrict the distributed knowledge to information that may have been transmitted, rather than that information currently held. Thus, we consider the notion of distributed knowledge $\mathtt{D}^{TA}_G$, defined as $\mathtt{D}^Y_G$ where $Y = \mathtt{ta} \circ m$. Using this, we can again support the key intuition:

Theorem 3 *Suppose that $\rightarrowtail$ is acyclic. If M is TA-secure with respect to $\rightarrowtail$ then M is confined with respect to* $\mathtt{ta} \circ m$ *and $\rightarrowtail$.*

Theorems 2 and 3 show that for both TO-security and TA-security, we are able to support the intuitive relationship between causal structure and distributed knowledge.

From Distributed Knowledge to Causality

In the preceeding, we have shown that the causal structure of a system implies a relationship between an agent's knowledge and the distributed knowledge of agents that may interfere with it. We now consider the converse direction. Suppose that security of a system implies that it is confined with respect to a local-state assignment Y and $\rightarrowtail$. Is it also the case that if a system is confined with respect to Y and $\rightarrowtail$ then it is secure? The following example shows that it does not.

Example 5: We show that it is possible for a system to be confined with respect to $\mathtt{view} \circ m$, yet still be TO-insecure. Thus, the converse to Theorem 2 does not hold. (A similar example may be constructed for $\mathtt{ta} \circ m$ and TA-security, we leave this as an exercise for the reader.)

Consider a system with agents H, D, L, each of which has a single action h, d, l, respectively. Thus $D = \{H, D, L\}$ and $A = \{h, d, l\}$. We take the set of states to be A^*, the initial state to be ϵ and state transitions to be defined by concatenation: $\alpha \cdot a = \alpha a$. Let the observations for agent H be given by $\mathtt{obs}_H(\alpha) = 0$, the observations for agent D be given by $\mathtt{obs}_D(\alpha) = \alpha \upharpoonright \{H, D\}$, and those for L be given by

1. $\mathtt{obs}_L(\alpha) = 0$ if α does not contain a d action,
2. $\mathtt{obs}_L(\alpha) = 1$ if α contains a d action and the first action of α is not l,
3. $\mathtt{obs}_L(\alpha) = 2$, otherwise.

Let the policy $\rightarrowtail$ be given by $H \rightarrowtail D \rightarrowtail L$. Then, if we take $\alpha_1 = hld$ and $\alpha_2 = lhd$, we have that

$$\begin{aligned} \mathtt{to}_L(\alpha_1) &= (\mathtt{to}_L(hl), \mathtt{view}_D(hl), d) \\ &= ((\mathtt{to}_L(h), \mathtt{view}_L(h), l), [\epsilon] \circ [h] \circ [h], d) \\ &= (\mathtt{to}_L(\epsilon), 0 \circ 0, l), [\epsilon][h], d) \\ &= ((\epsilon, 0, l), [\epsilon][h], d) \end{aligned}$$

and

$$\begin{aligned} \mathtt{to}_L(\alpha_2) &= (\mathtt{to}_L(lh), \mathtt{view}_D(lh), d) \\ &= (\mathtt{to}_L(l), [\epsilon] \circ [\epsilon] \circ [h], d) \\ &= ((\mathtt{to}_L(\epsilon), \mathtt{view}_L(\epsilon), l), [\epsilon] \circ [\epsilon] \circ [h], d) \\ &= ((\epsilon, 0, l), [\epsilon][h], d), \end{aligned}$$

where we include sequences in square braces for clarity. Thus, $\mathtt{to}_L(\alpha_1) = \mathtt{to}_L(\alpha_2)$. On the other hand, we have that $\mathtt{obs}_L(s_0 \cdot \alpha_1) = \mathtt{obs}_L(\alpha_1) = 1$ but $\mathtt{obs}_L(s_0 \cdot \alpha_2) = \mathtt{obs}_L(\alpha_2) = 2$. Thus, this system is not TO-secure.

We now show that, on the other hand, this system is confined with respect to the local state assignment $\mathtt{view} \circ m$. For this, we show that for each agent v, if $M, \pi, \alpha \models K_v p$ and π interprets p as depending only on $D \setminus \{v\}$ then $M, \pi, \alpha \models D^{TO}_{I_v} p$. For this, we show that if $\alpha \sim^{TO}_{I_v} \beta$ then there exists a sequence β' such that (1) $\mathtt{view}_v(\alpha) = \mathtt{view}_v(\beta')$ and (2) $\beta' \upharpoonright D \setminus \{v\} = \beta \upharpoonright D \setminus \{v\}$. It then follows using the fact that $M, \pi, \alpha \models K_v p$ and (1) that $M, \pi, \beta' \models p$, hence using (2) and the fact that π interprets p as depending only on $D \setminus \{v\}$ that $M, \pi, \beta \models p$. We consider several cases for v, with subcases for α:

`Case 1:` $v = L$. Here we have that π interprets p as depending only on $\{H, D\}$, $I_L = \{D\}$, and $\alpha \sim^{TO}_{I_L} \beta$ implies $\alpha \upharpoonright D = \beta \upharpoonright D$.

1. Suppose $v_L(\alpha) = 0(l0)^k$. Then there is no occurrence of d in α, and consequently, no such occurrence in β. Here we take $\beta' = (\beta \upharpoonright H)l^k$ which can easily be seen to satisfy (1) and (2).
2. Suppose $v_L(\alpha) = 0(l0)^k 1(l1)^j$. Then we can write $\alpha = \alpha_0 d \alpha_1$ where the occurrence of d is the first in α, $\alpha_0 \upharpoonright L = l^k$ and the first action in α_0 is not l, and $\alpha_1 \upharpoonright L = l^j$. Since $\alpha \upharpoonright D = \beta \upharpoonright D$, there is also an occurrence of d in β and we may write $\beta = \beta_0 d \beta_1$, where β_0 contains at least one h action. Define
$$\beta' = (\beta_0 \upharpoonright H) l^k d (\beta_1 \upharpoonright \{H, D\}) l^j.$$
This can be seen to satisfy (1) and (2).
3. Suppose $v_L(\alpha) = 0(l0)^k 2(l2)^j$. Then we can write $\alpha = \alpha_0 d \alpha_1$ where the occurrence of d is the first in α, $\alpha_0 \upharpoonright L = l^k$ and the first action in α_0 is l, and $\alpha_1 \upharpoonright L = l^j$. Since $\alpha \upharpoonright D = \beta \upharpoonright D$, there is also an occurrence of d in β and we may write $\beta = \beta_0 d \beta_1$, where β_0 does not contain d. Define $\beta' = l(\beta_0 \upharpoonright H) l^{k-1} d (\beta_1 \upharpoonright \{H, D\}) l^j$. This can be seen to satisfy (1) and (2).

`Case 2:` $v = D$. Here we have that π interprets p as depending only on $\{H, L\}$, $I_D = \{H\}$, and $\alpha \sim^{TO}_{I_D} \beta$ implies $\alpha \upharpoonright H = \beta \upharpoonright H$. Here $\mathtt{view}_H(\alpha)$ has the form $0(h0)^k$. Using the fact that $\alpha \upharpoonright H = \beta \upharpoonright H$, we can construct β' such that $\alpha \upharpoonright \{H, D\} = \beta' \upharpoonright \{H, D\}$ and $\alpha \upharpoonright \{H, L\} = \beta \upharpoonright \{H, L\}$. This yields (1) and (2).

`Case 3:` $v = H$. Here we have that π interprets p as depending only on $\{D, L\}$, $I_H = \emptyset$, and $\alpha \sim^{TO}_{I_H} \beta$ is trivially true for all β. We take $\beta' = (\beta \upharpoonright \{D, L\})(\alpha \upharpoonright H)$, which can be seen to satisfy (1) and (2). □

The appropriate diagnosis for this example appears to be that security of a system talks not just about what an agent knows about the actions of other agents, but also about how the agent's own actions are interleaved with those of other agents. Because we have formulated the notion of a proposition that "depends only on other agents" in a way that is not sensitive to such interleavings, we we are not able to express security of the system using our present formulation of the intuition. We now set about developing a variant formulation that does take such interleavings into account. We will show that this can be done in such as way as to completely characterize security in terms of the relationship between an agent's knowledge and the distributed knowledge of its interferers.

First, we relativize the notion of dependency to an agent. Let $F \subseteq D$ be a set of agents and let $u \in D$ be an agent. For a sequence $\alpha \in A^*$, define $\alpha \upharpoonright_u F$ to be the sequence in

$A^* \cup \{\bot_u\}$ otained from α by deleting occurences of action a with $\mathtt{dom}(a) \notin F$, and replacing each occurrence of an action a with $\mathtt{dom}(a) = u$ by $\bot_u$. We then say that a proposition $\Pi \subseteq A^*$ is *about F relative to u* if for all $\alpha, \alpha' \in A^*$, if $\alpha \restriction_u F = \alpha' \restriction_u F$ then $\alpha \in \Pi$ iff $\alpha' \in \Pi$.

We will similarly relativize the notions of distributed knowledge defined above. Suppose that Y is a local-state assignment in M and u is an agent. Then we define the new modal operator $\mathtt{D}^Y_{G,u}$ with semantics

$M, \pi, \alpha \models \mathtt{D}^Y_{G,u}\phi$ if for all $\alpha' \in A^*$ such that $\alpha \restriction_u G = \alpha' \restriction_u G$ and $Y_v(\alpha) = Y_v(\alpha')$ for all $v \in G$, we have $M, \pi, \alpha' \models \phi$.

Intuitively, this definition expresses that the group G is able to deduce ϕ from the information in the local states Y, given information about the actions of the group G and how they were interleaved, and the points in that interleaving at which agent u performed an action (but not the details of u's actions). In particular, we define the operators $\mathtt{D}^{TO}_{G,u}$ and $\mathtt{D}^{TA}_{G,u}$ by taking Y to be $\mathtt{view} \circ m$ and $\mathtt{ta} \circ m$ in this definition, respectively.

It is easily seen from the definitions and the fact that $\alpha \restriction_u G = \alpha' \restriction_u G$ implies $\alpha \restriction G = \alpha' \restriction G$ that $\mathtt{D}^Y_G\phi \Rightarrow \mathtt{D}^Y_{G,u}\phi$ is valid. Thus $\mathtt{D}^Y_{G,u}$ is a weaker notion of distributed knowledge than D^Y_G.

Using these new notions of dependence and distributed knowledge, we consider the following formulation of our intuition concerning the relationship between the knowledge of an agent u and the distributed knowledge of its interferers I_u:

Definition 6: A system M is *relatively confined with respect to a local-state assignment Y and a noninterference policy* $\rightarrowtail$ if for all sequences of actions $\alpha \in A^*$, for all agents u and all π that interpret p as depending only on $D \setminus u$ relative to u, we have $M, \pi, \alpha \models \mathtt{K}_u p \Rightarrow \mathtt{D}^Y_{I_u,u} p$, where $I_u = \{v \in D \mid u \neq v,\ v \rightarrowtail u\}$. □

In many cases of interest, this is a stronger statement than our previous formulation of the intuition using $\mathtt{D}^Y_G$. This is not obvious, since while it is plain that if Π depends only on $D \setminus u$ then Π depends only on $D \setminus u$ relative to u, in order to obtain our previous formulation we would also need to have $\mathtt{D}^Y_{G,u}\phi \Rightarrow \mathtt{D}^Y_G\phi$ which is the *converse* of the fact that $\mathtt{D}^Y_G\phi \Rightarrow \mathtt{D}^Y_{G,u}\phi$ noted above. However, the desired implication in fact often holds, because of the following.

Let $\rightarrowtail$ be a noninterference relation. If G is a set of agents, we write $G\downarrow$ for the set $\{v \mid v \rightarrowtail^* u \in G\}$, and write $u\downarrow$ for $\{u\}\downarrow$. Say that a local-state assignment Y *respects* $\rightarrowtail$ in M if for all agents u, if $\alpha \restriction (u\downarrow) = \beta \restriction (u\downarrow)$ then $Y_u(\alpha) = Y_u(\beta)$. Several of the local-state assignments we have introduced respect $\rightarrowtail$.

Lemma 1 *1. If $\alpha \restriction u\downarrow = \beta \restriction u\downarrow$ then $m_u(\alpha) \restriction u\downarrow = m_u(\beta) \restriction u\downarrow$.*

2. The local-state assignments ta *and* $\mathtt{ta} \circ m$ *respect* $\rightarrowtail$.

3. If M is TO-secure then the local-state assignments to, $\mathtt{to} \circ m$, view *and* $\mathtt{view} \circ m$ *respect* $\rightarrowtail$.

Subject to the condition that the local-state assignment Y respects $\rightarrowtail$, we obtain from the following that $\mathtt{D}^Y_{G,u}$ and $\mathtt{D}^Y_G$ are equivalent on the set of propositions of interest for confinement.

Lemma 2 *Suppose that Y respects $\rightarrowtail$ in M and let $u \notin G\downarrow$. Then if π interprets p as depending only on $D \setminus u$, we have $M, \pi \models \mathtt{D}^Y_{G,u}p \Rightarrow \mathtt{D}^Y_G p$.*

Proof: Assume that $M, \pi, \alpha \models \mathtt{D}^Y_{G,u}p$. We show that $M, \pi, \alpha \models \mathtt{D}^Y_G p$. Let $\alpha' \in A^*$ be a sequence such that $\alpha \restriction G = \alpha' \restriction G$ and $Y_v(\alpha) = Y_v(\alpha')$ for all $v \in G$. We need to show that $M, \pi, \alpha' \models p$. Define β to be a sequence obtained from α' by first deleting actions a with $\mathtt{dom}(a) = u$ and then inserting such actions in such a way that $\beta \restriction_u G = \alpha \restriction_u G$. (This is possible because $\alpha \restriction G = \alpha' \restriction G$.) Since Y respects $\rightarrowtail$ and $G\downarrow \subseteq D \setminus u$, it follows from this that also $Y_v(\alpha) = Y_v(\beta)$ for all $v \in G$. Thus, by the assumption that $M, \pi, \alpha \models \mathtt{D}^Y_{G,u}p$, we have $M, \pi, \beta \models p$. Since we also have that $\beta \restriction D\setminus u = \alpha' \restriction D\setminus u$, we obtain using that fact that π interprets p as depending only on $D \setminus u$ that $M, \pi, \alpha' \models p$, as required. □

From Lemma 2 we obtain the following.

Corollary 1 *If M is relatively confined with respect to Y and $\rightarrowtail$, and Y respects $\rightarrowtail$ in M, then M is confined with respect to Y and $\rightarrowtail$.*

Combining this result with with and Lemma 1, we obtain:

Corollary 2 *If M is relatively confined with respect to $\rightarrowtail$, then M is confined with respect to* $\mathtt{ta} \circ m$ *and $\rightarrowtail$. If M is TO-secure and relatively confined with respect to $\rightarrowtail$, then M is confined with respect to* $\mathtt{view} \circ m$ *and $\rightarrowtail$.*

Thus, the notion of relative confinement is stronger than the notion of confinement in the cases we considered above. We now establish that this stronger notion actually holds in these cases. We obtain this as a consequence of a more general result, for which we first need some technical definitions and results.

Say that a local-state assignment Y is *locally cumulative* if it satisfies the following conditions, for all agents $u \in D$, actions $a, b \in A$ and sequences of actions $\alpha, \beta \in A^*$:

[LC1.] If $Y_u(\alpha a) = Y_u(\epsilon)$ then $\mathtt{dom}(a) \neq u$ and $Y_u(\alpha) = Y_u(\epsilon)$.

[LC2.] If $\mathtt{dom}(a) \neq u$ and $\mathtt{dom}(b) \neq u$ and $Y_u(\alpha a) = Y_u(\beta b)$ then either $Y_u(\alpha) = Y_u(\beta b)$ or $Y_u(\alpha a) = Y_u(\beta)$ or $Y_u(\alpha) = Y_u(\beta)$.

[LC3.] If $\mathtt{dom}(a) \neq u$ and $\mathtt{dom}(b) = u$ and $Y_u(\alpha a) = Y_u(\beta b)$ then $Y_u(\alpha) = Y_u(\beta b)$.

[LC4.] If $\mathtt{dom}(a) = u$ and $\mathtt{dom}(b) = u$ and $Y_u(\alpha a) = Y_u(\beta b)$ then $Y_u(\alpha) = Y_u(\beta)$.

Some of the local-state assignments of interest to us have this property.

Lemma 3 *The local-state assignments* view *and* ta *are locally cumulative.*

The following results gives some technical properties that will be of use to us below.

Lemma 4 *If Y is locally cumulative, then for all sequences $\alpha, \beta \in A^*$ and agents u, $Y_u(\alpha) = Y_u(\beta)$ implies $Y_u(m_u(\alpha)) = Y_u(m_u(\beta))$.*

Lemma 5 *Let X be a full-information local-state assignment based on Y and $\rightarrowtail$. Then we have the following.*

1. *If M is X-secure with respect to $\rightarrowtail$ then for all $\alpha, \beta \in A^*$ and agents u, if $X_u(\alpha) = X_u(\beta)$ then $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\beta)$.*
2. *If $X_u(\alpha) = X_u(\beta)$ and $v \rightarrowtail u$, then $Y_v(m_v(\alpha)) = Y_v(m_v(\beta))$.*
3. *Suppose Y is locally cumulative. If $\alpha \restriction I_u \cup \{u\} = \beta \restriction I_u \cup \{u\}$ and $Y_v(m_v(\alpha)) = Y_v(m_v(\beta))$ for all $v \in I_u$, then $X_u(\alpha) = X_u(\beta)$.*

Lemma 6 *Suppose that Y respects $\rightarrowtail$ in M and $\rightarrowtail$ is acyclic. Then for all $\alpha, \beta \in A^*$, if $\alpha \restriction_u D \setminus u = \beta \restriction_u D \setminus u$ then for all $v \in I_u$, we have $Y_v(m_v(\alpha)) = Y_v(m_v(\beta))$.*

We can now state a result that generalizes Theorem 2 and Theorem 3.

Theorem 4 *Let $\rightarrowtail$ be acyclic, and let X be a full information local state assignment based on Y and $\rightarrowtail$. Suppose that Y respects $\rightarrowtail$ in M and is locally cumulative. Then if M is X-secure then M is relatively confined with respect to $Y \circ m$ and $\rightarrowtail$.*

Proof: Suppose π interprets p as depending only on $D \setminus u$ relative to u. We need to show that for all $\alpha \in A^*$ we have $M, \pi, \alpha \models \mathtt{K}_u p \Rightarrow \mathtt{D}^{Y \circ m}_{I_u,u} p$. We assume $M, \pi, \alpha \models \neg \mathtt{D}^{Y \circ m}_{I_u,u} p$, and show that $M, \pi, \alpha \models \neg \mathtt{K}_u p$.

Since $M, \pi, \alpha \models \neg \mathtt{D}^{Y \circ m}_{I_u,u} p$, there exists a sequence $\beta \in A^*$ such that $\alpha \restriction_u I_u = \beta \restriction_u I_u$ and $Y_v(m_v(\alpha)) = Y_v(m_v(\beta))$ for all $v \in I_u$ and $M, \pi, \beta \models \neg p$. Since $\alpha \restriction_u I_u = \beta \restriction_u I_u$, the sequences α and β have the same number k of occurrences of actions of agent u, and $\alpha \restriction I_u = \beta \restriction I_u$. Let γ be the sequence constructed from β by replacing the i-th occurrence of an action of agent u by the i-th action of agent u in α, for $i = 1 \ldots k$. Then we have that $\beta \restriction_u D \setminus u = \gamma \restriction_u D \setminus u$ and $\alpha \restriction I_u \cup \{u\} = \gamma \restriction I_u \cup \{u\}$.

By Lemma 6, since $\beta \restriction_u D \setminus u = \gamma \restriction_u D \setminus u$, we have that $Y_v(m_v(\beta)) = Y_v(m_v(\gamma))$ for all $v \in I_u$. It follows that $Y_v(m_v(\alpha)) = Y_v(m_v(\gamma))$ for all $v \in I_u$. Since also $\alpha \restriction I_u \cup \{u\} = \gamma \restriction I_u \cup \{u\}$ and Y is locally cumulative, we obtain using Lemma 5(3) that $X_u(\alpha) = X_u(\gamma)$. By X-security of M and Lemma 5(1), we conclude that $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\gamma)$. However, since $M, \pi, \beta \models \neg p$, $\beta \restriction_u D \setminus u = \gamma \restriction_u D \setminus u$ and π interprets p as depending only on $D \setminus u$ relative to u, we have that $M, \pi, \gamma \models \neg p$. Thus $M, \pi, \alpha \models \neg \mathtt{K}_u p$. □

Using Lemma 3 and Lemma 1 , we obtain the following result, which can be seen to strengthen Theorem 2 and Theorem 3 by Corollary 2.

Corollary 3 *If M is TA-secure with respect to $\rightarrowtail$ then M is relatively confined with respect to $\mathtt{ta} \circ m$ and $\rightarrowtail$. If M is TO-secure with respect to $\rightarrowtail$ then M is relatively confined with respect to $\mathtt{view} \circ m$ and $\rightarrowtail$.*

Admittedly, the added complexity of the operator $\mathtt{D}^Y_{G,u}$ used in the definition of local confinement makes this result somewhat less intuitive. However, the benefit that we obtain from this complexity is the ability to state the following converse to Theorem 4:

Theorem 5 *Let $\rightarrowtail$ be acyclic, and let X be a full-information local-state assignment based on Y and $\rightarrowtail$. If M is relatively confined with respect to $Y \circ m$ and $\rightarrowtail$ then M is X-secure.*

Proof: We need to show that for all sequences α, α' and agents u, if $X_u(\alpha) = X_u(\alpha')$ then $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. We proceed by induction on the combined length of α and α'. The base case of $\alpha = \alpha' = \epsilon$ is trivial, so we consider the case of sequences $\alpha a, \alpha'$, with $X_u(\alpha a) = X_u(\alpha')$, where $a \in A$.

Case 1: $\mathtt{dom}(a) \not\rightarrowtail u$. Then $X_u(\alpha) = X_u(\alpha a) = X_u(\alpha')$, so by the induction hypothesis, we have $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \alpha')$. We would like to show that $\mathtt{obs}_u(s_0 \cdot \alpha a) = \mathtt{obs}_u(s_0 \cdot \alpha' a)$. We suppose not, and obtain a contradiction. Note that it follows from the assumption and the conclusion above that $\mathtt{obs}_u(s_0 \cdot \alpha a) \neq \mathtt{obs}_u(s_0 \cdot \alpha)$. Define the interpretation π on p by $M, \pi, \gamma \models p$ iff $\gamma \restriction_u D \setminus u \neq \alpha a \restriction_u D \setminus u$. Then plainly π interprets p as depending only on $D \setminus u$ relative to u.

We show that $M, \pi, \alpha \models \mathtt{K}_u p$. Suppose that $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\gamma)$ and $M, \pi, \gamma \models \neg p$. From $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\gamma)$ it follows that $\alpha \restriction u = \gamma \restriction u$, and from $M, \pi, \gamma \models \neg p$ we have $\gamma \restriction_u D \setminus u = \alpha a \restriction_u D \setminus u$. It follows that $\gamma = \alpha a$. However, since $\mathtt{obs}_u(s_0 \cdot \alpha a) \neq \mathtt{obs}_u(s_0 \cdot \alpha)$, we have $\mathtt{view}_u(\gamma) = \mathtt{view}_u(\alpha a) \neq \mathtt{view}_u(\alpha)$, a contradiction. Thus, $\mathtt{view}_u(\alpha) = \mathtt{view}_u(\gamma)$ implies $M, \pi, \gamma \models p$. This shows that $M, \pi, \alpha \models \mathtt{K}_u p$.

Since π interprets p as depending only on $D \setminus u$ relative to u and M is relatively confined with respect to $Y \circ m$ and $\rightarrowtail$, it follows that $M, \pi \models \mathtt{D}^{Y \circ m}_{I_u,u} p$. However, since $\mathtt{dom}(a) \not\rightarrowtail u$, we have $\mathtt{dom}(a) \notin I_u \cup \{u\}$, so $\alpha a \restriction_u I_u = \alpha \restriction_u I_u$. Moreover, for $v \in I_u$, we have $m_v(\alpha a) = m_v(\alpha)$, so $Y_v(m_v(\alpha a)) = Y_v(m_v(\alpha))$. Plainly $M, \pi, \alpha a \models \neg p$. Thus, we have $M, \pi \models \neg \mathtt{D}^{Y \circ m}_{I_u,u} p$ This is a contradiction.

Case 2: $\mathtt{dom}(a) \rightarrowtail u$. Then $X_u(\alpha a) = (X_u(\alpha), Y_{\mathtt{dom}(a)}(\alpha), a)$. Since $X_u(\alpha a) = X_u(\alpha')$, we cannot have $\alpha' = \epsilon$, so write $\alpha' = \beta b$. If $\mathtt{dom}(b) \not\rightarrowtail u$, then we can switch the roles of αa and βb and apply the previous case. Thus, without loss of generality $\mathtt{dom}(b) \rightarrowtail u$, and we have $X_u(\beta b) = (X_u(\beta), Y_{\mathtt{dom}(a)}(\beta), b)$. It follows that $a = b$, $X_u(\alpha) = X_u(\beta)$ and $Y_{\mathtt{dom}(a)}(\alpha) = Y_{\mathtt{dom}(a)}(\beta)$.

Suppose that $\mathtt{obs}_u(s_0 \cdot \alpha a) \neq \mathtt{obs}_u(s_0 \cdot \beta a)$. By the induction hypothesis, we obtain from $X_u(\alpha) = X_u(\beta)$ that $\mathtt{obs}_u(s_0 \cdot \alpha) = \mathtt{obs}_u(s_0 \cdot \beta)$. Define π on p by $M, \pi, \gamma \models p$ iff $\gamma \restriction_u D \setminus u \neq \beta a \restriction_u D \setminus u$.

We show that $M, \pi, \alpha a \models \mathtt{K}_u p$. Suppose that $\mathtt{view}_u(\gamma) = \mathtt{view}_u(\alpha a)$ and $M, \pi, \gamma \models \neg p$. Then $\gamma \restriction u = \alpha a \restriction u$ and $\gamma \restriction_u D \setminus u = \beta a \restriction_u D \setminus u$. Now from $X_u(\alpha) = X_u(\beta)$ and the fact that X is a full-information local-state assignment, we obtain that $\alpha \restriction u = \beta \restriction u$. Hence $\gamma \restriction u = \alpha a \restriction u = \beta a \restriction u$. Combining this with $\gamma \restriction_u D \setminus u = \beta a \restriction_u D \setminus u$ we obtain $\gamma = \beta a$. This implies that $\mathtt{view}_u(\gamma) \neq \mathtt{view}_u(\alpha a)$, a contradiction,

since $\mathtt{obs}_u(s_0 \cdot \alpha a) \neq \mathtt{obs}_u(s_0 \cdot \beta a)$. This shows that $M, \pi, \alpha a \models \mathtt{K}_u p$. It is obvious that π interprets p as depending only on $D \setminus u$ relative to u, so by the assumption, we have $M, \pi, \alpha a \models \mathtt{D}^{Y \circ m}_{I_u, u} p$.

By an induction on the definition of X_u, we can see that $X_u(\alpha a) = X_u(\beta a)$ implies $\alpha a \restriction_u I_u \cup \{u\} = \beta a \restriction_u I_u \cup \{u\}$. By Lemma 5(2) we also have $Y_v(m_v(\alpha a)) = Y_v(m_v(\beta a))$ for all agents v with $v \in I_u$. Since $M, \pi, \beta a \models \neg p$, it follows that $M, \pi, \alpha a \models \neg \mathtt{D}^{Y \circ m}_{I_u, u} p$, a contradiction. □

In particular, combining this result with Corollary 3 we obtain the following.

Corollary 4 *M is relatively confined with respect to $\mathtt{ta} \circ m$ iff M is TA-secure. M is relatively confined with respect to $\mathtt{view} \circ m$ iff M is TO-secure.*

That is, we are able to give a complete characterization of the causal notions of TO-security and TA-security that is stated entirely in epistemic terms. The equivalence gives us a reduction of causal notions to epistemic notions.

Conclusion

While the notion of distributed knowledge has been studied in the literature on reasoning about knowledge since the 1980's, relatively little concrete use has been made of it. Notably, one of the few relates distributed knowledge to the notion of of Lamport causality, which is appropriate in asynchronous message passing systems (see (Fagin et al. 1995) Proposition 4.4.3). Our work provides a study of a slightly different nature, dealing with distributed knowledge and information flow in causally constrained asynchronous systems. Specifically, we have clarified the intuition that the causal structure of a system constrains agents to acquire only external knowledge that is distributed knowledge to the agents that may causally affect it.

In doing so, we have argued that the classical definition of distributed knowledge needs to be adapted in order to capture the intuition. We have shown that, besides intersection of information sets, there are other ways that a group of agents might combine their information. In particular, we have argued that such combination needs to be relativized to an agent to which the group is communicating its information in order to fully capture our intuitions concerning information flow and causal structure. We believe our new notions of distributed knowledge are intuitive and may find application in other contexts. We remark that the need for such variants has previously been argued by Moses and Bloom (Moses and Bloom 1994), who show that distributed knowledge it is too *weak* for certain applications, and propose a stronger notion I_G called *inherent knowledge*, such that $I_G\phi \Rightarrow D_G\phi$ is valid. This is a strengthening whereas we propose weakenings, and the context in this work is rather different from ours: systems in which communication is by asynchronous message passing, rather than the asynchronously operating systems with a synchronous communication mechanism, as in our definition of systems. Finally, we believe that the perspective from the logic of knowledge provides fresh insight into the definitions of causality being used in the literature on computer security. We have shown that it is possible to express these definitions precisely in terms of how information flows in the system.

There are several directions that could be explored in future research. We note that our definitions are suited to deterministic systems in which actions are interleaved and have immediate effects: different formulations may be required to capture similar intuitions in asynchronous message passing systems, or systems with simultaneous actions, such as the model used in (Fagin et al. 1995). It would also be of interest to consider nondeterministic systems, and systems in which actions have probabilistic effects. Our semantic model and notions of causality originate in the computer security literature and have a richer temporal structure than is generally considered in KR work in the area (Pearl 2000; Halpern and Pearl 2001), but it would be interesting to investigate the relationships. We remark also that we have confined our attention to acyclic noninterference policies. It remains to formulate generalizations of these results that encompass systems with causal cycles.

References

Fagin, R.; Halpern, J.; Moses, Y.; and Vardi, M. 1995. *Reasoning About Knowledge*. MIT-Press.

Goguen, J., and Meseguer, J. 1982. Security policies and security models. In *Proc. IEEE Symp. on Security and Privacy*, 11–20.

Haigh, J., and Young, W. 1987. Extending the noninterference version of MLS for SAT. *IEEE Trans. on Software Engineering* SE-13(2):141–150.

Halpern, J. Y., and Pearl, J. 2001. Causes and explanations: A structural-model approach - Part II: Explanations. In Nebel, B., ed., *IJCAI*, 27–34. Morgan Kaufmann.

Moses, Y., and Bloom, B. 1994. Knowledge, timed precedence and clocks (preliminary report). In *Proc. ACM Symp. on Principles of Distributed Computing*, 294–303.

Pearl, J. 2000. *Cauality: Models, reasoning and inference*. Cambridge University Press.

Rushby, J. 1992. Noninterference, transitivity, and channel-control security policies. Technical Report CSL-92-02, SRI International.

van der Meyden, R. 2007. What, indeed, is intransitive noninterference (extended abstract). In Biskup, J., and Lopez, J., eds., *Proc. European Symp. on Research in Computer Security (ESORICS)*, volume 4734 of *Springer LNCS*, 235–250. Full paper at `http://www.cse.unsw.edu.au/~meyden`.