The Logic of Secrets

Extended Abstract

Zuojun Xiong
Southwest University, China

Thomas Ågotnes
Southwest University, China
University of Bergen, Norway

Yuzhi Zhang
Southwest University, China

In this paper we formalise the notion of a *secret*, in epistemic logic – a standard framework for formal reasoning about information in multi-agent systems. The notion of a secret is fundamental in areas such as safety and security, in particular in cryptography, authentication and access control. What is a secret? While dictionary definitions of the noun varies somewhat – "a piece of knowledge that is hidden and intended to be kept hidden" (Wiktionary); "a piece of information that is only known by one person or a few people and should not be told to others" (Cambridge Dictionary); "something that is kept or meant to be kept unknown or unseen by others" (Oxford English Dictionary); "something kept from the knowledge of others" (Merriam-Webster) – it is clear that secrets are fundamentally about *knowledge* and *ignorance* (the lack of knowledge). For example, when we say that "Ann keeps her pin code secret" or "Bill has a secret girlfriend" we mean that there is something (Ann's pin code or the identity of Bill's girlfriend) that is (1) known by someone (Ann or Bill) and (2) not known by others.

In this paper we formalise secrets, or more precisely the notion of "secretly knowing", in the standard framework for formalising knowledge, namely epistemic logic [7]. We introduce a modality S_a, such that $S_a\varphi$ is intended to mean that agent a secretly knows φ. We study the properties of secretly knowing that follow from the basic definition based on knowledge and ignorance.

We focus here on the *epistemic* properties of secrets. As discussed above these are quite fundamental, but it should be mentioned that there are other aspects of secrets, such as *intentionality* ("*. . . intended to be kept hidden*") that we abstract away from here. Furthermore, in this paper we focus on formalising a basic notion of secretly knowing: we assume that the secret is exclusively known by a single person. As seen above definitions of secrets also allow for the secret to be known by a (small) number of people. We focus here on the simplest case in order to clarify the basic principles of secretly knowing as much as possible.

For characterising secrets in terms of knowledge, neccessary conditions for "agent a secretly knows φ" includes (1) that a knows φ and (2) that any other agent b does not know φ. We argue, however, that a stronger condition is needed: (2') not only should any other agent b not know whether φ, but *a should know that b doesn't know*. This property of secrets is not explicitly mentioned in the definitions cited above, but it is usually implicitly assumed. Indeed, if Bill believes that other people know who his girlfriend is, or if he merely doesn't know that they don't know who his girlfriend is, the identify of his girlfriend wouldn't be called a secret. We can now use the language of epistemic logic [7], where $K_a\varphi$ intuitively means that agent a knows φ, to express the fact that "agent a secretly knows φ":

$$K_a\varphi \wedge K_a\Big(\bigwedge_{b\neq a} \neg K_b\varphi\Big). \tag{SKs}$$

We note that in standard epistemic logic this is equivalent to

$$K_a\varphi \wedge K_a\Big(\bigwedge_{b\neq a} (\neg K_b\varphi \wedge \neg K_b\neg\varphi)\Big) \tag{SK}$$

where "b not knowing *that* φ" has been replaced with "b not knowing *whether* φ" [8, 11].

In order to formally study the logical properties of secret knowledge, we introduce new modalities S_a, such that $S_a\varphi$ means that agent a secretly knows φ in the precise sense defined above. $S_a\varphi$ is, of course, definable in terms of K_a and K_b, but we introduce it as a primary operator because we are interested not only in the interaction properties of secrets and knowledge, but also in the main principles of secrets in a language without the knowledge operators. Thus we define two formal languages $\mathcal{L}_{SK}$ and $\mathcal{L}_S$, both parameterised by a non-empty set $\mathtt{Prop}$ of propositional letters and a finite non-empty set $\mathtt{Agt}$ of agents and defined for $\varphi \in \mathcal{L}_{SK}$ and $\psi \in \mathcal{L}_S$ as follows:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid S_a\varphi \qquad \psi ::= p \mid \neg\psi \mid (\psi \wedge \psi) \mid S_a\psi$$

where $p \in \mathtt{Prop}, a \in \mathtt{Agt}$.

An *epistemic (S5) model* $M = (W, \sim, V)$ consists of a set of states W, an equivalence relation $\sim_a$ on W for each $a \in \mathtt{Agt}$ and a valuation function $V : W \to 2^{\mathtt{Prop}}$. Satisfaction of a formula $\varphi \in \mathcal{L}_{SK}$ in a state w of a model $M = (W, \sim, V)$ is defined as follows:

$M, w \models p$ iff $w \in V(p)$.
$M, w \models \neg\varphi$ iff $M, w \not\models \varphi$.
$M, w \models \varphi \wedge \psi$ iff $M, w \models \varphi$ and $M, w \models \psi$.
$M, w \models K_a\varphi$ iff $\forall w' \in W$, if $w \sim_a w'$, then $M, w' \models \varphi$.
$M, w \models S_a\varphi$ iff $\forall w' \sim_a w \in W$ such that $M, w' \models \varphi$ and $\forall b \neq a$, $\exists u \in W$ such that $w' \sim_b u$ and $M, u \models \neg\varphi$.

It is easy to see that $S_a\varphi$ holds iff $K_a\varphi \wedge K_a\big(\bigwedge_{b\in\mathtt{Agt}\setminus\{a\}} \neg K_b\varphi\big)$ holds.

Tables 1 and 2 shows logical principles of secretly knowing in the form of valid axioms and validity preserving rules. In addition there are the usual instances of propositional tautologies/rules and S5 axioms and necessitation rule for the K modalities (not shown). Of course, in the full language $\mathcal{L}_{SK}$ the S_a modalities are derivable from K_a, but Table 1 also illustrates interesting derived properties such as introspection about secrets. Perhaps more interesting are the validities and rules in the language without K_a, shown in Table 2, the core principles of secrets without refering explicitly to knowledge.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 2020, Auckland, New Zealand

https://doi.org/doi

Table 3 show some non-valid axiom schemas and rules that are not validity preserving that we for various reasons find interesting.

The most obvious thing to observe is that the secretly knowing modality S_a is not *normal* [6]: while it does distribute over implication, it does not satisfy the necessitation rule (from φ infer $S_a\varphi$). Neither is the dual or the negation of S_a (while both satisfy necessitation neither distribute over implication).

In the vocabulary of non-normal modal logics (see, e.g., [12]), the S_a operator is an ECKT4 modality: it satisfies the axioms and rules of the basic ECK system, i.e., the axioms K, C and the rule RE, in addition to T and 4. ECK, in turn, is the extension of the weakest modal logic with *neighbourhood semantics* E with the axioms C and K. In addition to the ECKT4 properties, our modality also satisfies the $\top$ axiom.

One consequence of "negative necessitation" (derivable) is that there are no tautological secrets – there are no formulas φ such that $S_a\varphi$ is valid. The reader can also observe that formulas with nested secrets often can be reduced to formulas with more shallow nesting. For example, both $S_aS_b\varphi$ and $S_a\neg S_b\varphi$ for $a \neq b$ are equivalent to $\bot$, while $S_aS_a\varphi$ is equivalent to $S_a\varphi$. This illustrate that we get certain *normal forms* with limited nestings; we leave out details due to the limited space.

Going one step further in simplifying the language, one could have just a single S modality, interpreted in the same structures but where S refers implicitly to a fixed designated agent a. This logic perhaps gives us the most distilled principles of secretly knowing, found in the two first parts of Table 2: ECKT4 plus the $\top$ axiom. Unfortunately we don't know whether these axioms and rules (plus instances of propositional tautologies and modus ponens) are *complete*, or whether the full table is complete for the language $\mathcal{L}_S$. As mentioned these logics are extensions of the weakest non-normal logic with neighbourhood semantics E, which indicates that standard techniques for proving completeness can be used. However, at time of writing proving completeness for even the standard systems EK and ECK, not to mention of ECKT4, are still open problems[1]. Of course, for the language $\mathcal{L}_{SK}$ we get a trivial completeness result by simply extending multi-agent S5 with the S axiom.

Of related work, secrets play a key role in work on *gossip protocols* [3–5] which use logic to formalise reasoning about information flow. However, secrets are taken as a primary notion rather than derived from the underlying notion of knowledge, and the focus is not on the properties of secretly knowing. Also related are modal logics of access control [1, 2, 9, 10]. Some works in this area are concerned with properties of secrets of the type we consider in this paper, but they are (again) mostly taken as primary rather than derived from an underlying abstract epistemic framework. A more detailed discussion of the relationship appears in the full version of this paper. As mentioned in the introduction, sometimes secrets are known by a small number of people rather than a single person. In this case *common knowledge* seems to play an important role. We leave further discussion for the full paper.

[1] A manuscript with a completeness proof by Frederik van der Putte and Paul McNamara for both EK and ECK is currently under review for a journal, and has been shown to us. It acknowledges the extension with the 4 axiom as non-trivial.

Table 1: Secret principles in the language $\mathcal{L}_{SK}$. The following axiom and rule schemas ($a \neq b$) are valid/validity preserving.

Interaction axioms for S_a and K_a		
(S)	$S_a\varphi \leftrightarrow K_a\varphi \wedge K_a(\bigwedge_{b\neq a} \neg K_b\varphi)$	Def. of S_a
(4SK)	$S_a\varphi \rightarrow K_aS_a\varphi$	Positive secret knowledge introspection
(5SK)	$\neg S_a\varphi \rightarrow K_a\neg S_a\varphi$	Negative secret knowledge introspection
(P)	$S_a\varphi \rightarrow (K_a\varphi \wedge \neg K_b\varphi)$	Secret privacy
(NKS)	$\neg K_bS_a\varphi$	Secret unknowability
(NSK1)	$\neg S_aK_b\varphi$	Knowledge is no secret
(NSK2)	$\neg S_a\neg K_b\varphi$	Ignorance is no secret
(NC)	$K_aS_a\varphi \vee K_a\neg S_a\varphi$	Secret neg. completeness

Table 2: Secret principles in the language $\mathcal{L}_S$ (and $\mathcal{L}_{SK}$). The following axiom and rule schemas ($a \neq b$) are valid/validity preserving. Core axioms/rules have names written in bold. Names not written in bold are derivable from the core axioms/rules.

Axioms for S_a		
(K)	$S_a(\varphi \rightarrow \psi) \rightarrow (S_a\varphi \rightarrow S_a\psi)$	Secret distribution
(T)	$S_a\varphi \rightarrow \varphi$	Secret veridicality
(4)	$S_a\varphi \rightarrow S_aS_a\varphi$	Secret introspection
(C)	$(S_a\varphi \wedge S_a\psi) \rightarrow S_a(\varphi \wedge \psi)$	Secret combination
(D)	$S_a\varphi \rightarrow \neg S_a\neg\varphi$	Secrets partiallity
($\top$)	$\neg S_a\top$	No tautological secrets
($\bot$)	$\neg S_a\bot$	No contradictory secrets
Rules for S_a		
(RE)	From $\varphi \leftrightarrow \psi$ infer $S_a\varphi \leftrightarrow S_a\psi$	Replacement of equivalents
(Nnec)	From φ infer $\neg S_a\varphi$	Negative necessitation
(Dnec)	From φ infer $\neg S_a\neg\varphi$	Diamond necessitation
Interaction axioms for S_a and S_b		
(Ex1)	$S_a\varphi \rightarrow \neg S_b\varphi$	Secret exclusivity
(Ex2)	$S_a\neg S_a\varphi \rightarrow \neg S_b\neg S_b\varphi$	Higher-order secret exclusivity
(N1)	$\neg S_aS_b\varphi$	No secret secrets
(N2)	$\neg S_a\neg S_b\varphi$	No secret non-secrets

Table 3: Non-validities and unsound rules, when $|\mathtt{Agt}| \geq 2$. For each of the following, there are formulas φ and ψ that demonstrates the non-validity/that the rule is not validity preserving.

$\not\models \neg S_a(\varphi \rightarrow \psi) \rightarrow (\neg S_a\varphi \rightarrow \neg S_a\psi)$
$\not\models S_a(\varphi \wedge \psi) \rightarrow S_a\varphi$
$\not\models \neg S_a\neg(\varphi \rightarrow \psi) \rightarrow (\neg S_a\neg\varphi \rightarrow \neg S_a\neg\psi)$
$\not\models \neg S_a\neg\varphi \rightarrow S_a\varphi$
$\not\models \neg S_a\neg S_a\varphi$
$\models \varphi \quad \not\Rightarrow \quad \models S_a\varphi$
$\models \varphi \rightarrow \psi \quad \not\Rightarrow \quad \models S_a\varphi \rightarrow S_a\psi$
$\models \varphi \rightarrow \psi \quad \not\Rightarrow \quad \models \neg S_a\varphi \rightarrow \neg S_a\psi$
$\models \varphi \rightarrow \psi \quad \not\Rightarrow \quad \models S_a\neg\varphi \rightarrow S_a\neg\psi$
$\models \varphi \rightarrow \psi \quad \not\Rightarrow \quad \models \neg S_a\neg\varphi \rightarrow \neg S_a\neg\psi$

REFERENCES

[1] Martín Abadi. Logic in access control. In *18th Annual IEEE Symposium of Logic in Computer Science, 2003. Proceedings.*, pages 228–233. IEEE, 2003.

[2] Martin Abadi, Michael Burrows, Butler Lampson, and Gordon Plotkin. A calculus for access control in distributed systems. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 15(4):706–734, 1993.

[3] Krzysztof R Apt, Davide Grossi, and Wiebe van der Hoek. Epistemic protocols for distributed gossiping. *arXiv preprint arXiv:1606.07516*, 2016.

[4] Krzysztof R Apt and Dominik Wojtczak. Verification of distributed epistemic gossip protocols. *Journal of Artificial Intelligence Research*, 62:101–132, 2018.

[5] Maduka Attamah, Hans Van Ditmarsch, Davide Grossi, and Wiebe van der Hoek. Knowledge and gossip. In *ECAI*, pages 21–26, 2014.

[6] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge University Press, 2001.

[7] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. Cambridge, MA: The MIT Press, 1995.

[8] Jie Fan, Yanjing Wang, and Hans Van Ditmarsch. Contingency and knowing whether. *The Review of Symbolic Logic*, 8(1):75–107, 2015.

[9] Philip WL Fong. Relationship-based access control: protection model and policy language. In *Proceedings of the first ACM conference on Data and application security and privacy*, pages 191–202, 2011.

[10] Deepak Garg and Martín Abadi. A modal deconstruction of access control logics. In *International Conference on Foundations of Software Science and Computational Structures*, pages 216–230. Springer, 2008.

[11] Sergiu Hart, Aviad Heifetz, and Dov Samet. "knowing whether","knowing that," and the cardinality of state spaces. *journal of economic theory*, 70(1):249–256, 1996.

[12] Eric Pacuit. *Neighborhood semantics for modal logic*. Springer, 2017.