# Bounded Forgetting

**Yi Zhou**
Intelligent Systems Laboratory
School of Computing and Mathematics
University of Western Sydney, NSW, Australia

**Yan Zhang**
Intelligent Systems Laboratory
School of Computing and Mathematics
University of Western Sydney, NSW, Australia

**Abstract**

The result of forgetting some predicates in a first-order sentence may not exist in the sense that it might not be captured by any first-order sentences. This, indeed, severely restricts the usage of forgetting in applications. To address this issue, we propose a notion called $k$-forgetting, also called bounded forgetting in general, for any fixed number $k$. We present several equivalent characterizations of bounded forgetting and show that the result of bounded forgetting, on one hand, can always be captured by a single first-order sentence, and on the other hand, preserves the information that we are concerned with.

## Introduction

The notion of forgetting has attracted many researchers from different disciplines under different names, e.g., forgetting in Artificial Intelligence, uniform interpolation in Mathematical Logic, etc. In Artificial Intelligence, the notion of forgetting has been extensively studied in many different knowledge representation paradigms such as first-order logic (Lin and Reiter 1994; Zhang and Zhou 2010), propositional logic (Lang, Liberatore, and Marquis 2003), modal logic (van Ditmarsch et al. 2008; Herzig and Mengin 2008; Zhang and Zhou 2009), description logic (Wang et al. 2009; Konev, Walther, and Wolter 2009), logic programming (Zhang and Foo 2006; Eiter and Wang 2008). The reason why forgetting is so remarkable is not only because of its own theoretical importance but also its practical value for a wide range of applications, for instance, conflict solving (Lang and Marquis 2002; Eiter and Wang 2008), belief change (Zhang and Foo 2006), conservative extension (Ghilardi et al. 2006; Lutz, Walther, and Wolter 2007), progression in situation calculus (Lin and Reiter 1997; Liu and Lakemeyer 2009; Vassos and Levesque 2008), partition-based reasoning (Amir and McIlraith 2005), weakest sufficient condition (Lin 2001), reasoning about knowledge (Zhang and Zhou 2009) and so on.

In this paper, we consider forgetting in classical first-order logic as it plays a central role in the logic family and is technically representative. Although only focusing on first-order logic, we believe that the basic ideas and technical results raised in this paper can also be employed by forgetting in other rich knowledge paradigms, such as description logics (Baader et al. 2003), logics of knowledge (Fagin et al. 1995) and logic programming (Gelfond and Lifschitz 1988).

An early attempt to characterize forgetting in first-order logic is due to Lin and Reiter (1994), who proposed a semantic definition of forgetting a set of predicates in a first-order *theory* (also called *strong forgetting*). However, as shown in their paper, the result of strong forgetting is not always first-order expressible. That is, there exists a first-order theory $T$ and a set $\Omega$ of predicates such that there is no first-order theory $T'$, which is the result of strongly forgetting $\Omega$ in $T$.

To tackle this issue, Zhang and Zhou (2010) proposed an alternative characterization, called *weak forgetting*, which is weaker than strong forgetting but guarantees that the result of weakly forgetting in theories is always first-order expressible. That is, given a first-order theory $T$ and a set $\Omega$ of predicates, there always exists a theory $T'$, which is the result of weakly forgetting $\Omega$ in $T$.

However, from a knowledge representation point of view, we are more interested in *sentences* (i.e. finite theories) rather than arbitrary theories as infinite theories are essentially not representable. Hence, a natural question arises whether the result of weak forgetting in first-order sentences can always be represented by a sentence as well. Unfortunately, as shown in (Zhang and Zhou 2010), the answer is negative. This, indeed, severely restricts the usage of forgetting in applications, as it becomes infeasible to represent the result of (weak, strong) forgetting.

It can be observed that, in most applications, we are only interested in those relatively simple (first-order) formulas. One natural measurement of the simplicity of first-order formulas is their quantifier rank, i.e. the maximal number of nested quantifiers in a formula.

This motivates us to consider a notion of *bounded forgetting* in this paper. The idea is quite simple. That is, we only take into account those sentences whose quantifier ranks are less or equal than a fixed number $k$. We present several equivalent characterizations of bounded forgetting and show that the result of bounded forgetting in sentences can always be represented by a single first-order sentence. Also, we discuss some related properties of bounded forgetting. In particular, we show that bounded forgetting, although weaker than strong and weak forgetting, preserves the information that we are concerned with.

## Strong and Weak Forgetting

In this section, we briefly review the notions of strong forgetting (Lin and Reiter 1994) and weak forgetting (Zhang and Zhou 2010) in first-order logic. We assume that the readers are familiar with some basic notions and notations of classic first-order logic with equality.

In this paper, we are focused on forgetting in sentences (i.e. finite theories) rather than arbitrary theories. In this sense, the context of forgetting is to consider whether a first-order sentence $\psi$ is the result of forgetting a set $\Omega$ of predicates in another sentence $\phi$.

An early attempt to characterize this is so-called strong forgetting, introduced by Lin and Reiter (1994). Strong forgetting is defined based on the identicalness of structures. Two structures $\mathcal{M}$ and $\mathcal{M}'$ are said to be *identical with exception* on a set $\Omega$ of predicates, denoted by $\mathcal{M} \sim_\Omega \mathcal{M}'$, if they agree on everything except the interpretations of the predicates in $\Omega$.

**Definition 1 (Strong forgetting)** *Let $\phi$ and $\psi$ be two sentences and $\Omega$ a set of predicates. Then, $\psi$ is said to be the result of* strongly forgetting *$\Omega$ in $\phi$ if* [1]

$$Mod(\psi) = \{\mathcal{M}' \mid \exists \mathcal{M} \in Mod(\phi), \mathcal{M} \sim_\Omega \mathcal{M}'\}.$$

As shown in (Lin and Reiter 1994), the result of strong forgetting can always be represented as an existential second-order sentence. More precisely, the class of models defined in Definition 1 is exactly captured by the second-order sentence

$$\exists \Omega' \phi(\Omega/\Omega'),$$

where $\Omega'$ is a set of new predicates corresponding to $\Omega$, and $\phi(\Omega/\Omega')$ is the formula obtained from $\phi$ by simultaneously replacing each predicate $P$ in $\Omega$ by its corresponding predicate in $\Omega'$. For convenience, we also use $SF(\phi, \Omega)$ to denote the second-order sentence.

However, $SF(\phi, \Omega)$ is not always first-order expressible (Lin and Reiter 1994). That is, in some cases, there is no first-order sentence $\psi$, which is equivalent to $SF(\phi, \Omega)$. In this case, we also say that the result of strongly forgetting $\Omega$ in $\phi$ does not *exist*.

An alternative notion, called weak forgetting, is introduced by Zhang and Zhou (2010) based on the notion of irrelevance. A sentence $\phi$ is said to be *irrelevant* to a set $\Omega$ of predicates, denoted by $IR(\phi, \Omega)$, iff there exists another sentence, which is equivalent to $\phi$ and contains no predicates from $\Omega$.

**Definition 2 (Weak forgetting)** *Let $\phi$ and $\psi$ be two sentences and $\Omega$ a set of predicates. Then, $\psi$ is said to be the result of* weakly forgetting *$\Omega$ in $\phi$ if it is equivalent to*

$$\{\phi' \mid \phi \models \phi', IR(\phi', \Omega)\}.$$

According to its definition, the result of weak forgetting can always be represented as a (possibly infinite) theory. For convenience, we also use $WF(\phi, \Omega)$ to denote this theory.

However, $WF(\phi, \Omega)$ cannot always be rewritten as a single first-order sentence (i.e. finite theory). That is, in some cases, there is no first-order sentence $\psi$, which is equivalent to $WF(\phi, \Omega)$. In this case, we also say that the result of weakly forgetting $\Omega$ in $\phi$ does not *exist*.

The relationships between strong forgetting and weak forgetting are discussed in detail by Zhang and Zhou (2010). One major difference between weak forgetting and strong forgetting is that, the result of weak forgetting can be captured by a first-order (possibly infinite) theory, while this is even not the case for strong forgetting. In addition, weak forgetting is weaker than strong forgetting, i.e. $SF(\phi, \Omega) \models WF(\phi, \Omega)$, but not the other way around. Interestingly, if the result of strong forgetting is first-order expressible, then it coincides with the result of weak forgetting.

As infinite theories are essentially not representable from a knowledge representation point of view, it arises a problem whether we can define a new notion of forgetting in sentences such that the result of forgetting can always be captured by a single first-order sentence. To address this issue, we introduce bounded forgetting in the next section, and show that the result of bounded forgetting in first-order sentences always exists.

## Bounded Forgetting

In order to capture the notion of bounded forgetting comprehensively, we present several equivalent characterizations of it, including a direct definition to enumerate all consequences, an indirect definition as the strongest necessary condition, an axiomatic definition by five postulates and a model theoretical definition based on the notion of modified elementary equivalence.

Like strong and weak forgetting, bounded forgetting is defined for forgetting a set of predicates in a first-order sentence. However, unlike the other two, it is defined with respect to a fixed natural number $k$, called the *bound*. We use the term *bounded forgetting* for all $k$ if it is clear from the context.

First of all, we recall the notion of quantifier rank for first-order formulas, which plays a vital role in bounded forgetting. The *quantifier rank* of a first-order formula $\phi$, denoted by $qr(\phi)$, is defined recursive as follows:

- $qr(\phi) = 0$ if $\phi$ is an atomic formula;
- $qr(\neg\phi) = qr(\phi)$;
- $qr(\phi \wedge \psi) = qr(\phi \vee \psi) = max\{qr(\phi), qr(\psi)\}$.
- $qr(\exists x \phi) = qr(\forall x \phi) = qr(\phi) + 1$.

Perhaps the most direct way to define bounded forgetting is to enumerate all the possible formulas in the result of bounded forgetting.

**Definition 3 (Bounded Forgetting: Form I)** *Let $\phi$ and $\psi$ be two sentences, $\Omega$ a set of predicates and $k$ a natural number. Then, $\psi$ is said to be the result of $k$-*forgetting *$\Omega$ in $\phi$ if $qr(\psi) \leq k$ and it is equivalent to*

$$\{\gamma \mid IR(\gamma, \Omega), qr(\gamma) \leq k, \phi \models \gamma\}.$$

An alternative way to defined bounded forgetting is to consider the so-called strongest necessary condition.

[1] For convenience, we use $Mod(\phi)$ to denote the class of all models of a sentence $\phi$.

**Definition 4 (Bounded Forgetting: Form II)** *Let $\phi$ and $\psi$ be two sentences, $\Omega$ a set of predicates and $k$ a natural number. Then, $\psi$ is said to be the result of $k$-*forgetting *$\Omega$ in $\phi$ if*

1. *$\phi \models \psi$, $IR(\psi, \Omega)$ and $qr(\psi) \leq k$;*
2. *$\psi$ is the strongest one among all sentences satisfying the above conditions. That is, for any $\psi'$ satisfying condition 1, $\psi \models \psi'$.*

As a logical notion, it is always desirable to provide an axiomatic system for it. A possible way, particularly popular in the field of knowledge representation, is to use a postulate-based approach.

**Definition 5 (Bounded Forgetting: Form III)** *Let $\phi$ and $\psi$ be two sentences, $\Omega$ a set of predicates and $k$ a natural number. Then, $\psi$ is said to be the result of $k$-*forgetting *$\Omega$ in $\phi$ if it satisfies the following postulates:*

**(B)** *Bounded: $qr(\psi) \leq k$.*
**(W)** *Weakening: $\phi \models \psi$.*
**(IR)** *Irrelevance: $IR(\psi, \Omega)$.*
**(PP)** *Positive Persistence: for any sentence $\psi'$, if $IR(\psi', \Omega)$, $\phi \models \psi'$ and $qr(\psi') \leq k$, then $\psi \models \psi'$.*
**(NP)** *Negative Persistence: for any sentence $\psi'$, if $IR(\psi', \Omega)$, $\phi \not\models \psi'$ and $qr(\psi') \leq k$, then $\psi \not\models \psi'$.*

Although Definition 5 and Definition 4 look very similar, we treat them as two different approaches because the former is more focused on the simple underlying intuitions that bounded forgetting should obey, while the latter is more focused on how to specify the result of bounded forgetting.

Finally, we present a model theoretical characterization of bounded forgetting. We say that two structures $\mathcal{M}$ and $\mathcal{M}'$ are *elementary equivalent*, denoted by $\mathcal{M} \equiv \mathcal{M}'$, if they satisfy the same set of first-order sentences. We say that two structures $\mathcal{M}$ and $\mathcal{M}'$ are *$k$-elementary with exception on* a set $\Omega$ of predicates, denoted by $\mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'$, if they satisfy the same set of first-order sentences that are irrelevant to $\Omega$ and whose quantifier ranks are less or equal than $k$, i.e., for all $\phi$ such that $IR(\phi, \Omega)$ and $qr(\phi) \leq k$, $\mathcal{M} \models \phi$ iff $\mathcal{M}' \models \phi$. Also, we omit $\Omega$ when it is empty.

**Definition 6 (Bounded Forgetting: Form IV)** *Let $\phi$ and $\psi$ be two sentences, $\Omega$ a set of predicates and $k$ a natural number. Then, $\psi$ is said to be the result of $k$-*forgetting *$\Omega$ in $\phi$ if $qr(\psi) \leq k$ and its models are models of $\phi$ closed under $k$-elementary with exception on $\Omega$, i.e.,*

$$Mod(\psi) = \{\mathcal{M}' \mid \exists \mathcal{M} \in Mod(\phi), \mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'\}.$$

In fact, all the above definitions for bounded forgetting are equivalent.

**Theorem 1** *All the above definitions of bounded forgetting, i.e. Definitions 3-6, are equivalent.*

**Proof:**[2] From Form III to Form II is obvious. Now we prove from Form II to Form I. Suppose that $\gamma$ is a sentence such that $IR(\gamma, \Omega)$, $qr(\gamma) \leq k$ and $\phi \models \gamma$. Then, $\psi \models \gamma$ as well according to Condition 2 in Definition 4. Hence, $\psi \models \{\gamma \mid IR(\gamma, \Omega), qr(\gamma) \leq k, \phi \models \gamma\}$. On the other hand, since $\psi$ itself satisfies Condition 1 in Definition 4, $\psi \in \{\gamma \mid IR(\gamma, \Omega), qr(\gamma) \leq k, \phi \models \gamma\}$. This shows that $\psi$ is equivalent to $\{\gamma \mid IR(\gamma, \Omega), qr(\gamma) \leq k, \phi \models \gamma\}$.

Now we prove from Form I to Form III. The postulates (B) holds obviously. (W) holds as well since every $\gamma$ is entailed by $\phi$. For (IR), let $\gamma^*$ be the sentence equivalent to $\gamma$ and contains no predicates from $\Omega$. Then, $\psi$ is equivalent to the set of all $\gamma^*$ constructed above, which contains no predicates from $\Omega$. Thus, (IR) holds. For (PP), if $\psi'$ satisfies the conditions, then $\psi'$ is one of the $\gamma$, thus entailed by $\psi$. This shows that (PP) holds as well. Finally, for (NP), suppose that there exists $\psi'$ such that $\phi \not\models \psi'$, $qr(\psi') \leq k$ and $IR(\psi', \Omega)$. Then, $\psi \not\models \psi'$. Otherwise, $\phi \models \psi'$ since $\phi \models \psi$, a contradiction.

Finally, we prove the equivalence between Form I and Form IV. We show that their models are exactly the same. On one hand, suppose that $\mathcal{M}'$ and $\mathcal{M}$ are two structures such that $\mathcal{M} \models \phi$ and $\mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'$. Let $\gamma$ be a sentence such that $IR(\gamma, \Omega)$, $qr(\gamma) \leq k$ and $\phi \models \gamma$. Then, $\mathcal{M} \models \gamma$ since $\mathcal{M} \models \phi$. Thus, $\mathcal{M}' \models \gamma$ since $\mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'$. On the other hand, suppose that $\mathcal{M}'$ is a model of $\{\gamma \mid IR(\gamma, \Omega), qr(\gamma) \leq k, \phi \models \gamma\}$. Let $T_0$ be the set of sentences that are irrelevant to $\Omega$ and satisfied by $\mathcal{M}'$ and whose quantifier rank is no more than $k$, i.e., $T_0 = \{\psi' \mid IR(\psi', \Omega), \mathcal{M}' \models \psi', qr(\psi') \leq k\}$. Then, $T_0 \cup \{\phi\}$ is consistent. Otherwise, by the compactness theorem, there exists a finite subset of $T_0$, say $T_0'$, such that $T_0' \cup \{\phi\}$ is inconsistent. Let $\phi_0'$ be the conjunction of all sentences in $T_0'$. Therefore, $\phi \models \neg\phi_0'$. In addition, $IR(\neg\phi_0', \Omega)$ and $qr(\neg\phi_0') \leq k$. It follows that $\mathcal{M}'$ is a model of $\neg\phi_0'$, a contradiction. Hence, there exists a model $\mathcal{M}$ of $T_0 \cup \{\phi\}$. Clearly, for all $\gamma$ such that $IR(\gamma, \Omega)$, $qr(\gamma) \leq k$ and $\mathcal{M}' \models \gamma$, $\mathcal{M} \models \gamma$ since $\mathcal{M} \models T_0$. Conversely, for all $\gamma$ such that $IR(\gamma, \Omega)$, $qr(\gamma) \leq k$ and $\mathcal{M} \models \gamma$, $\mathcal{M}' \models \gamma$ as well. Hence, $\mathcal{M}$ and $\mathcal{M}'$ are $k$-elementarily equivalent with exception on $\Omega$. In addition, $\mathcal{M}$ is a model of $\phi$. ■

As the above definitions are essentially equivalent, we use them indistinguishably if clear from the context. Like strong and weak forgetting, we say that the result of $k$-forgetting $\Omega$ in $\phi$ *exists* if there is such a first-order sentence $\psi$. In this case, we also use $kF(\phi, \Omega)$ to denote $\psi$.

The following good news shows that, unlike strong forgetting and weak forgetting, the result of bounded forgetting for any fixed $k$ can always be represented as a single first-order sentence.

**Theorem 2 (Existence of Bounded Forgetting)** *The result of bounded forgetting always exists. That is, given a sentence $\phi$, a natural number $k$ and a set $\Omega$ of predicates, there always exists a sentence $\psi$, which is the result of $k$-forgetting $\Omega$ in $\phi$.*

**Proof:** This assertion follows from the fact that, if only considering the signature used in $\phi$, the set of sentences (up to equivalence) whose quantifier ranks are less or equal than $k$ is finite. ■

[2] Due to a space limit, we select those relatively important and difficult proofs in this paper.

## Case Studies

In this section, we consider several examples to illustrate how bounded forgetting works.

**Example 1** Let $\phi_1$ be the sentence $\forall x \exists y(P(x,y) \vee Q(y))$. Then, $SF(\phi_1, P)$ is $\exists R \forall x \exists y(R(x,y) \vee Q(y))$, which is equivalent to $\top$. In addition, $WF(\phi_1, P)$ is equivalent to $\top$ as well since $\top$ is the only sentence, up to equivalence, entailed by $\phi_1$ and irrelevant to $P$. Also, for any $k$, $kF(\phi_1, P)$ is $\top$.

In this case, all the different notions of forgetting exist, and actually coincide with each other.

**Example 2** Let $\phi_2$ be the sentence $\forall x \exists y(P(x,y) \wedge Q(y))$. Then, $SF(\phi_2, P)$ is $\exists R \forall x \exists y(R(x,y) \wedge Q(y))$, which is equivalent to $\exists y Q(y)$. It can be checked that $WF(\phi_2, P)$ is equivalent to $\exists y Q(y)$ as well since $\exists y Q(y)$ implies any consequence of $\phi_2$ that is irrelevant to $P$.

On the other hand, if we set $k = 0$, then $0F(\phi_2, P)$ is $\top$. Note here that although $\exists y Q(y)$ is irrelevant to $P$ and entailed by $\phi_2$, it is not counted in $0F(\phi_2, P)$ as its quantifier rank is 1. Nevertheless, if we set $k \geq 1$, then $kF(\phi_2, P)$ is equivalent to $\exists y Q(y)$ again.

In this case, all the different notions of forgetting exist. However, $0F(\phi_2, P)$ is $\top$, which is weaker than the rest (captured by $\exists y Q(y)$).

**Example 3** Let $\phi_3$ be the conjunction of the following sentences:

$$\begin{aligned} &\forall xyz(P(x,y) \wedge P(x,z) \rightarrow y = z), \\ &\forall xyz(P(x,z) \wedge P(y,z) \rightarrow x = y), \\ &\forall x \exists y P(x,y), \\ &\exists y \forall x \neg P(x,y). \end{aligned}$$

Roughly speaking, $P$ in $\phi_3$ associates each element in a domain to another element. The first two sentences specify that this association is one-to-one. The third one means that each element associates with one element, while the last sentence means that there exists an element not associated. A typical model of $\phi_3$ is in the infinite natural number domain, and $P$ is interpreted as the successor relation, i.e., $P(x,y)$ iff $y = x + 1$. In fact, the number 0 is not associated in this model. On the other hand, it is not difficult to see that every model of $\phi_3$ must be infinite.

Now we consider to forget $P$ in $\phi_3$. On one hand, it can be checked that both $SF(\phi_3, P)$ and $WF(\phi_3, P)$ are equivalent to the following infinite set of sentences:

$$\{\psi_2, \psi_3, \ldots, \psi_n, \ldots\},$$

where for any $n \geq 2$, $\psi_n = \exists x_1 \ldots x_n \bigwedge_{i \neq j} x_i \neq x_j$. This is because all the models of this set of sentences are exactly all infinite structures. In addition, it is well known in model theory that this set cannot be captured by a single sentence. Hence, according to our definitions, neither $SF(\phi_3, P)$ nor $WF(\phi_3, P)$ exists.

On the other hand, it can be checked that

$$kF(\phi_3, P) = \begin{cases} \bigwedge_{2 \leq i \leq k} \psi_i & k \geq 2, \\ \top & k < 2. \end{cases}$$

This is because $\{\psi_i \mid i \leq k\}$ is the set of all sentences that are entailed by $\phi_3$ and do not contain $P$ and whose quantifier ranks are no more than $k$.

In this case, $SF(\phi_3, P)$ and $WF(\phi_3, P)$ can be captured by an infinite theory but not a sentence (i.e. finite theory). On the other hand, $kF(\phi_3, P)$ can be captured by a single sentence for any $k$.

**Example 4** Let $\phi_4$ be the conjunction of the following sentences:

$$\begin{aligned} &\forall xyz(P(x,y) \wedge P(x,z) \rightarrow y = z), \\ &\forall xyz(P(x,z) \wedge P(y,z) \rightarrow x = y), \\ &\forall x \exists y P(x,y), \\ &\forall y \exists x P(x,y), \\ &\forall xy(P(x,y) \rightarrow (Q(x) \leftrightarrow \neg Q(y))). \end{aligned}$$

Roughly speaking, $Q$ in $\phi_4$ splits a domain into two groups of elements: in which those $Q$ holds and in which those $Q$ does not hold, and $P$ in $\phi_4$ represents a bijection between the two groups. In other words, the two groups should have the same cardinality and $P$ is a witness.

Now we consider to forget $P$ in $\phi_4$. Since the class of all structures, in which the above two groups have the same cardinality, is not first-order expressible, $SF(\phi_4, P)$ cannot even be represented as an infinite theory.

It can be verified that $WF(\phi_4, P)$ is equivalent to the following infinite set of sentences:

$$\{\exists x Q(x), \exists x \neg Q(x), \psi'_2, \psi'_3, \ldots, \psi'_n, \ldots\},$$

where for any $n \geq 2$,
$\psi'_n = \exists x_1 \ldots x_n Q(x_1) \wedge \cdots \wedge Q(x_n) \wedge \bigwedge_{i \neq j}(x_i \neq x_j)$
$\leftrightarrow \exists y_1 \ldots y_n \neg Q(y_1) \wedge \cdots \wedge \neg Q(y_n) \wedge \bigwedge_{i \neq j}(y_i \neq y_j)$.
Intuitively, $\psi'_n$ means that if there are at least $n$ elements that $Q$ holds for, then there are at least $n$ elements that $Q$ does not hold for, and vice versa. However, this set of sentences cannot be captured by a single first-order sentence.[3]

On the other hand,

$$kF(\phi_4, P) = \begin{cases} \bigwedge_{2 \leq i \leq k} \psi'_i \wedge \exists x Q(x) \wedge \exists x \neg Q(x) & k \geq 2, \\ \exists x Q(x) \wedge \exists x \neg Q(x) & k = 1, \\ \top & k = 0. \end{cases}$$

This is because if we consider sentences whose quantifier ranks are 0 (i.e. quantifier-free sentences), $\top$ is the only one (up to equivalence) entailed by $\phi_4$. Now consider sentences whose quantifier ranks are 1. We have that $\phi_4$ entails both $\exists x Q(x)$ and $\exists x \neg Q(x)$. If we consider more sentences, then $\phi_4$ entails $\psi'_k$ for any $k$.

In this case, $SF(\phi_4, P)$ cannot even be captured by an infinite theory; $WF(\phi_4, P)$ can be captured by an infinite theory but not a sentence. On the other hand, $kF(\phi_4, P)$ can be captured by a single sentence for any $k$.

[3] To verify this result and other negative results in our examples, some techniques in model theory, e.g. the Löwenheim-Skolem theorem, are needed. Here, we omit the detailed proofs of them as they are not the focus of this paper.

## Bounded Forgetting vs Strong/Weak Forgetting

As bounded forgetting claims to be a new notion for forgetting, in this section, we consider its relationships to both strong and weak forgetting.

In fact, the properties shown in this paper, although they look natural, are somewhat intricate as first-order logic is very sensitive to the related notions such as finite axiomatizability, restricted signature, $k$-elementary equivalence and so on. As mentioned in the introduction, this is one of the reasons why we consider first-order logic in this paper. Hence, we do need extra cautious here. As an example, it might be anticipated that if $SF(\phi, \Omega)$ exists and $qr(\phi) \leq k$, then it can be represented by a sentence $\psi$ such that $qr(\psi) \leq k$ as well. At first glance, this seems "intuitive" because the result of strong forgetting should not be more complicated (captured by the quantifier rank) than the original sentence. However, this is not true. For instance, let $\phi$ be the formula

$$\forall x P(x) \wedge \exists x Q(x) \wedge \exists x \neg Q(x).$$

Clearly, $qr(\phi) = 1$. Suppose we want to forget $Q$ in $\phi$. The result of (strongly, weakly) forgetting is

$$\forall x P(x) \wedge \exists x_1 x_2 (x_1 \neq x_2),$$

whose quantifier rank is 2, actually greater than $qr(\phi)$.

First of all, we show that bounded forgetting is weaker than weak forgetting, thus strong forgetting as well.

**Proposition 1** *Let $\phi$ be a sentence, $\Omega$ a set of predicates and $k$ a natural number. Then, $WF(\phi, \Omega) \models kF(\phi, \Omega)$. Also, $SF(\phi, \Omega) \models kF(\phi, \Omega)$.*

By observed from Example 3, the converses of Proposition 1 are not necessarily true in general.

However, the following property shows that, if the bound $k$ is big enough, then $k$-forgetting can be used to simulate both strong and weak forgetting.

**Proposition 2** *Let $\phi$ be a sentence and $\Omega$ a set of predicates. If $SF(\phi, \Omega)$ is first-order expressible, then there is a natural number $k$ such that $kF(\phi, \Omega)$ is equivalent to $SF(\phi, \Omega)$. This assertion holds for weak forgetting as well.*

More importantly, the following theorem shows that, if we only consider irrelevant sentences whose quantifier ranks are no more than $k$, then $k$-forgetting does not lose information in the sense that any consequences of (weak, strong) forgetting are consequences of $k$-forgetting as well.

**Theorem 3** *Let $\phi$ be a sentence, $\Omega$ a set of predicates and $k$ a natural number. Let $\psi$ be a sentence such that $qr(\psi) \leq k$ and $IR(\psi, \Omega)$. Then, $SF(\phi, \Omega) \models \psi$ iff $WF(\phi, \Omega) \models \psi$ iff $kF(\phi, \Omega) \models \psi$.*

**Proof:** According to the definitions, it is not difficult to see that $SF(\phi, \Omega) \models WF(\phi, \Omega)$ and $WF(\phi, \Omega) \models kF(\phi, \Omega)$. Hence, it suffices to show that, if $SF(\phi, \Omega) \models \psi$, then $kF(\phi, \Omega) \models \psi$. Now suppose that $SF(\phi, \Omega) \models \psi$ and $\mathcal{M}'$ is a model of $kF(\phi, \Omega)$. According to Definition 6, there exists $\mathcal{M}$ such that $\mathcal{M} \models \phi$ and $\mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'$. Therefore, $\mathcal{M} \models SF(\phi, \Omega)$. It follows that $\mathcal{M} \models \psi$ since $SF(\phi, \Omega) \models \psi$. Hence, $\mathcal{M}' \models \psi$ since $\mathcal{M} \equiv_{\Omega}^{k} \mathcal{M}'$. ■

Theorem 3 is another important result for bounded forgetting. In fact, in most application scenarios, we are usually interested in those relatively simple formulas. On one hand, simple first-order formulas have rich expressive power to model an application domain. On the other hand, complicated first-order formulas are barely manageable from a knowledge engineering point of view. One of the most natural measures for the simplicity of first-order formulas is its quantifier rank. Hence, Theorem 3 indicates that, bounded forgetting can be used to simulate both strong and weak forgetting when only considering those bounded formulas.

**Proposition 3** *Let $\phi$ be a sentence and $\Omega$ a set of predicates. Then, $WF(\phi, \Omega)$ is equivalent to $\{kF(\phi, \Omega) \mid k \in \mathbb{N}\}$.*

Proposition 3 shows that weak forgetting is exactly the union of all bounded forgetting.

**Proposition 4** *Let $\phi$ be a sentence and $\Omega$ a set of predicates. Let $k$ and $k'$ be two numbers such that $k \leq k'$. Then, $k'F(\phi, \Omega) \models kF(\phi, \Omega)$.*

Proposition 4 shows that, the larger the bound is, the stronger the bounded forgetting is. Together with Proposition 3, it indicates that bounded forgetting actually converges to weak forgetting.

A problem arises which bound $k$ should be chosen for forgetting in a first-order sentence $\phi$. A natural candidate is to set $k$ as the quantifier rank of $\phi$ itself, i.e. $k = qr(\phi)$. The rationale for this is to keep the forgetting result as complicated as the original formula. Nevertheless, it would be better to leave it open for different application scenarios.

## Other Semantic Properties

In this section, we discuss some basic properties which are not directly related to strong or weak forgetting.

**Proposition 5** *Let $\psi$ be the result of $k$-forgetting $\Omega$ in $\phi$. Let $\psi'$ be another sentence. Then, $\psi$ and $\psi'$ are equivalent iff $\psi'$ is also the result of $k$-forgetting $\Omega$ in $\phi$.*

Proposition 5 shows that bounded forgetting is closed under equivalence.

**Proposition 6** *Let $\phi$ be a sentence, $\Omega$ a set of predicates and $k$ a natural number. Let $\psi$ be a sentence such that $IR(\psi, \Omega)$ and $qr(\psi) \leq k$. Then, $\phi \models \psi$ iff $kF(\phi, \Omega) \models \psi$.*

Proposition 6 is a combination of (PP) and (NP) in Definition 5, stating that bounded forgetting should not affect those formulas irrelevant to the set of predicate forgotten and whose quantifier rank is no more than the bound.

**Proposition 7** *There always exists a sentence $\psi$ which is equivalent to $kF(\phi, \Omega)$ and contains no predicates from $\Omega$.*

Proposition 7 shows that we can always compute bounded forgetting within the restricted signature disjoint from the set of predicates forgotten.

**Proposition 8** *Let $\phi$ be a sentence, $\Omega_1$ and $\Omega_2$ two sets of predicates and $k$ a natural number. Then, $kF(\phi, \Omega_1 \cup \Omega_2)$ is equivalent to $kF(kF(\phi, \Omega_1), \Omega_2)$.*

Proposition 8 shows that the order does not matter when forgetting a sequence of predicate sets.

**Proposition 9** *Let $\phi_1$ and $\phi_2$ be two sentences such that $\phi_1 \models \phi_2$. Let $\Omega$ be a set of predicates and $k$ a natural number. Then, $kF(\phi_1, \Omega) \models kF(\phi_2, \Omega)$.*

Proposition 9 shows that bounded forgetting in a sentence is entailed by bounded forgetting in a stronger sentence. As a consequence, if two sentences are equivalent, then results of $k$-forgetting in them should be equivalent as well.

**Proposition 10** *Let $\phi_1$ and $\phi_2$ be two sentences, $\Omega$ a set of predicates and $k$ a natural number. Then, $kF(\phi_1 \vee \phi_2, \Omega)$ is equivalent to $kF(\phi_1, \Omega) \vee kF(\phi_2, \Omega)$.*

**Proof:** $kF(\phi_1, \Omega) \vee kF(\phi_2, \Omega) \models kF(\phi_1 \vee \phi_2, \Omega)$ follows from Proposition 9. We now show the other way around. Suppose that $\mathcal{M}$ is a model of $kF(\phi_1 \vee \phi_2, \Omega)$. Then, by Definition 6, there exists $\mathcal{M}'$ which is a model of $\phi_1 \vee \phi_2$ and $\mathcal{M}' \equiv_\Omega^k \mathcal{M}$. Hence, $\mathcal{M}' \models \phi_1$ or $\mathcal{M}' \models \phi_2$. Without loss of generality, assume that $\mathcal{M}' \models \phi_1$. Again, by Definition 6, $\mathcal{M}$ is a model of $kF(\phi_1, \Omega)$. ■

Proposition 10 shows that bounded forgetting is closed under disjunction. However, like standard forgetting in propositional logic, bounded forgetting is closed under neither conjunction nor negation.

## Conclusion

This paper introduces a new notion called bounded forgetting to address the issue that traditional forgetting in first-order sentences may not exist. Several equivalent characterizations of bounded forgetting are proposed (see Definitions 3-6 and Theorem 1) and two important results about bounded forgetting are shown. First, the result of bounded forgetting can always be captured by a single first-order sentence (see Theorem 2). Second, compared to both strong and weak forgetting, bounded forgetting with bound $k$ does not lose information if we are only concerned with sentences whose quantifier ranks are no more than $k$ (see Theorem 3), which, we argue that, is indeed the case in most application scenarios. Note that although the technical results of this paper appear natural, they are technically challenging to prove.

This paper focuses on theoretical foundations of bounded forgetting in first-order logic. We plan to consider its applications at the next step. For instance, we can apply the notion of bounded forgetting to progression in situation calculus, in which we may only take into account the post conditions whose quantifier ranks are bounded. This will guarantee the result of progression in action theories to be finitely first-order representable.

Although this paper is focused on first-order logic, we believe that the basic ideas and results of bounded forgetting can be applied to other rich knowledge representation formalisms as well, such as description logics, logics of knowledge and logic programming. For instance, both strong and weak forgetting in description logics suffer from the existence problem either. Similarly, we can introduce bounded forgetting for description logics, where the bound can be defined as the maximal number of nested roles. This points out another important direction for our future investigations.

## References


Amir, E., and McIlraith, S. A. 2005. Partition-based logical reasoning for first-order and propositional theories. *Artif. Intell.* 162(1-2):49–88.

Baader, F.; Calvanese, D.; McGuinness, D. L.; Nardi, D.; and Patel-Schneider, P. F., eds. 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press.

Eiter, T., and Wang, K. 2008. Semantic forgetting in answer set programming. *Artif. Intell.* 172(14):1644–1672.

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

Gelfond, M., and Lifschitz, V. 1988. The stable model semantics for logic programming. In *ICLP'88*, 1070–1080.

Ghilardi, S.; Lutz, C.; Wolter, F.; and Zakharyaschev, M. 2006. Conservative extensions in modal logic. In *Advances in Modal Logic*, 187–207.

Herzig, A., and Mengin, J. 2008. Uniform interpolation by resolution in modal logic. In *JELIA*, 219–231.

Konev, B.; Walther, D.; and Wolter, F. 2009. Forgetting and uniform interpolation in large-scale description logic terminologies. In *IJCAI*, 830–835.

Lang, J., and Marquis, P. 2002. Resolving inconsistencies by variable forgetting. In *KR*, 239–250.

Lang, J.; Liberatore, P.; and Marquis, P. 2003. Propositional independence: Formula-variable independence and forgetting. *J. Artif. Intell. Res. (JAIR)* 18:391–443.

Lin, F., and Reiter, R. 1994. Forget it! In *Proceedings of the AAAI Fall Symposium on Relevance*, 154–159.

Lin, F., and Reiter, R. 1997. How to progress a database. *Artif. Intell.* 92(1-2):131–167.

Lin, F. 2001. On strongest necessary and weakest sufficient conditions. *Artif. Intell.* 128(1-2):143–159.

Liu, Y., and Lakemeyer, G. 2009. On first-order definability and computability of progression for local-effect actions and beyond. In *IJCAI*, 860–866.

Lutz, C.; Walther, D.; and Wolter, F. 2007. Conservative extensions in expressive description logics. In *IJCAI*, 453–458.

van Ditmarsch, H. P.; Herzig, A.; Lang, J.; and Marquis, P. 2008. Introspective forgetting. In *AI*, 18–29.

Vassos, S., and Levesque, H. J. 2008. On the progression of situation calculus basic action theories: Resolving a 10-year-old conjecture. In *AAAI*, 1004–1009.

Wang, K.; Wang, Z.; Topor, R. W.; Pan, J. Z.; and Antoniou, G. 2009. Concept and role forgetting in ALC ontologies. In *International Semantic Web Conference*, 666–681.

Zhang, Y., and Foo, N. Y. 2006. Solving logic program conflict through strong and weak forgettings. *Artif. Intell.* 170(8-9):739–778.

Zhang, Y., and Zhou, Y. 2009. Knowledge forgetting: Properties and applications. *Artif. Intell.* 173(16-17):1525–1537.

Zhang, Y., and Zhou, Y. 2010. Forgetting revisited. In *KR*, 602–604.