

# Formalizing Probabilistic Safety Claims

Heber Herencia-Zapana<sup>1\*</sup>, George Hagen<sup>2</sup>, and Anthony Narkawicz<sup>2</sup>

<sup>1</sup> National Institute of Aerospace, Hampton, VA

<sup>2</sup> NASA Langley Research Center, Hampton, VA

**Abstract.** A safety claim for a system is a statement that the system, which is subject to hazardous conditions, satisfies a given set of properties. Following work by John Rushby and Bev Littlewood, this paper presents a mathematical framework that can be used to state and formally prove probabilistic safety claims. It also enables hazardous conditions, their uncertainties, and their interactions to be integrated into the safety claim. This framework provides a formal description of the probabilistic composition of an arbitrary number of hazardous conditions and their effects on system behavior. An example is given of a probabilistic safety claim for a conflict detection algorithm for aircraft in a 2D airspace. The motivation for developing this mathematical framework is that it can be used in an automated theorem prover to formally verify safety claims.

## 1 Introduction

In [9, 5], Rushby and Littlewood present a framework for formalizing safety claims for systems, which is illustrated with probabilistic safety claims in an automated theorem prover. In this paper, the mathematics behind their ideas is formalized. The mathematical framework presented will equip the reader to formalize a probabilistic safety claim about a system with an *arbitrary number* of hazardous conditions in a precise mathematical formula that can be proved in a theorem prover. One advantage that this adds to Rushby’s approach is that it provides a formal way for new hazardous conditions to be considered without changing the overall structure of the safety argument.

A *safety claim* is a statement that a system will behave in a desired manner with an acceptable probability. A *hazard* is a state or set of conditions that, together with other conditions in the environment, will cause a system to enter an undesirable state. For more on terminology related to safety analyses and system hazards, see [4]. In this paper, a *potentially hazardous condition*, referred to hereafter simply as a *hazardous condition*, is anything that may cause a system to behave in an unexpected or undesired manner. Examples of hazardous conditions may include such things as signal noise, timing delays, or interruptions of service. The number of hazardous conditions in a safety argument typically depends on the available expertise in analyzing the system, and it is important to

---

\* This work was supported in part by the National Aeronautics and Space Administration under NASA Cooperative Agreement NCC-1-02043.

allow the safety claim to evolve as new factors are uncovered. Hazardous Conditions typically have uncertainties associated with them, and they can therefore be modeled as random variables. This paper proposes a formal mathematical framework for modeling hazardous conditions as random variables in a way that makes it possible to also model interactions between different hazardous conditions. The underlying concepts are due to Rushby [9], but this paper gives precise mathematical definitions of probabilistic safety claims and provides a concrete example of such a claim. The example presented is for a state based conflict detection system.

In general, a probabilistic safety claim can be expressed as a mathematical formula stating that the probability of a certain event occurring is bounded in a specific range. Since new factors affecting system behavior may become known in the future, is desirable for the safety argument to be easily updated without reconstructing the entire argument. The mathematical formalism presented in this paper allows hazardous conditions to be modeled in a way that is modular and can handle the addition of new hazardous conditions.

The interdependency between random variables, e.g., hazardous conditions, is modeled by *probabilistic kernels*, which uses the fact that the set of all hazardous conditions can be modeled via a concatenation of  $\sigma$ -algebras, as seen in [10]. A  $\sigma$ -algebra is a set of sets where it is possible to assign probabilities to elements in a consistent way, and is often used to model events. See Section 2.5 for more complete discussion of probabilistic kernels.

The composition of hazardous conditions is formalized through the concatenations of Lebesgue integrals. This allows hazardous conditions and assumptions to be incorporated into the formula in a modular fashion. The majority of the complexity is encapsulated in sub-formulas specific to the assumption or hazardous condition in question, while the main safety claim formula need only be modified in a limited and systematic fashion. The mathematics behind this formalization is presented in following sections.

## 2 Systems

Systems of interest are those that can modeled as well-defined functions with inputs and outputs. In this formalization, a system is a function  $S$  with  $n$  *parameters* and  $m$  *variables*:

$$S : (K_1 \times \dots \times K_n; L_1 \times L_2 \times \dots \times L_m) \rightarrow \mathcal{T}_0,$$

where  $K_1, \dots, K_n$  and  $L_1, \dots, L_m$  are the *types* of the  $n$  parameters and  $m$  variables of  $S$ , respectively. The type  $\mathcal{T}_0$  consists of the possible outputs of  $S$ , and if  $k_i \in K_i$  and  $l_j \in L_j$ , then  $S(k_1, \dots, k_n; l_1, \dots, l_m)$  is an element of  $\mathcal{T}_0$ . It will sometimes be useful to view the system  $S$  as only a function on its  $m$  variables  $l_1, \dots, l_m$ , where the  $n$  parameters  $k_1, \dots, k_n$  are fixed, the notation  $S_{k_1, \dots, k_n}(l_1, \dots, l_m)$  is used in place of  $S(k_1, \dots, k_n; l_1, \dots, l_m)$ . Because the system  $S$  will be modeled as a random variable in order to reason about it probabilistically, it is assumed that  $\mathcal{T}_0$  is a measure space with  $\sigma$ -algebra  $\sigma(\mathcal{T}_0)$ .

The values  $k_1, \dots, k_n$  of the parameters of the system are predetermined and their values, without any errors, are known to the system. In a real system, the values of the input variables  $l_1, \dots, l_m$  are measured by the system, and the measurements can have errors. These errors may be due to either expected accuracy problems with instruments or faulty components in other systems from which the instruments receive data. In either case, events that can cause such measurement errors in the system are referred to as *hazardous conditions*, which are formally modeled in this context in Section 2.2.

For a system described in this way, a probabilistic safety claim is a statement that, given some set of possible hazardous conditions, the probability that the value of the system  $S$  lies in a predetermined subtype  $Z_0$  of  $\mathcal{T}_0$  is contained in particular range  $[p_0, p_1]$ .

## 2.1 Modeling Uncertainty in System Variables

As noted above, the values of the  $n$  parameters  $k_1, \dots, k_n$  of the system  $S$  are known to the system without errors. The errors in the measurements of the input variables  $l_1, \dots, l_m$  can be modeled as random variables

$$\mathbf{l}_i : \Omega \rightarrow L_i$$

where  $(\Omega, \sigma(\Omega))$  is a probability space ( $\sigma(\Omega)$  is a  $\sigma$ -algebra on the set  $\Omega$ ). Thus, given a fixed value  $\kappa = \{k_1, \dots, k_m\}$  for the set of parameters, the system  $S$  becomes a random variable as well:

$$\begin{aligned} S_\kappa : (\Omega, \sigma(\Omega)) &\rightarrow (\mathcal{T}_0, \sigma(\mathcal{T}_0)) \\ \chi &\mapsto S(\kappa, \mathbf{l}_1(\chi), \mathbf{l}_2(\chi), \dots, \mathbf{l}_m(\chi)) \in \mathcal{T}_0. \end{aligned}$$

Thus, if  $Z_0$  is any measurable subset of  $\mathcal{T}_0$  (i.e. an element of  $\sigma(\mathcal{T}_0)$ ), and if the distributions of the random variables  $\mathbf{l}_i$  are known, then the probability that the output of  $S_\kappa$  lies in  $Z_0$  can be computed.

## 2.2 Modeling Hazardous Conditions

As noted in Section 2, the errors in the variables  $l_1, \dots, l_m$  of the system  $S$  may be due to either expected accuracy problems with instruments or faulty components in other systems from which the instruments receive data. Conditions in the environment of a system that can cause such measurement errors in the system are referred to as *hazardous conditions*.

In a model of the environment of the system  $S$ , which includes the output of possible hazardous conditions, these conditions can be modeled as random variables

$$H_i : (\Omega, \sigma(\Omega)) \rightarrow (\mathcal{T}_i, \sigma(\mathcal{T}_i)),$$

where  $i \geq 1$ ,  $\mathcal{T}_i$  is an arbitrary type, and  $\sigma(\mathcal{T}_i)$  is a  $\sigma$ -algebra on  $\mathcal{T}_i$ . This modeling framework allows for the computation of the probability that a hazardous condition  $H_i$  takes values in a particular subtype of  $\mathcal{T}_i$ .

### 2.3 Modeling All Possible Hazardous Conditions

It is possible that the environment of a system  $S$  has an arbitrary number of hazardous conditions. Further, it may be the case that when developing a model of system behavior, only a few of these possible hazardous conditions are understood. Even in this case, the environment of the system can be modeled as an infinite product

$$\mathcal{T} = \prod_{i=0}^{\infty} \mathcal{T}_i,$$

where  $\mathcal{T}_0$  is the type of the output values of  $S$ , and for  $i \geq 1$ ,  $\mathcal{T}_i$  is the type of the output of the  $i$ -th hazardous condition  $H_i$ . This is a measure space with  $\sigma$ -algebra  $\sigma(\mathcal{T}) = \prod_{i=0}^{\infty} \sigma(\mathcal{T}_i)$ . This type of model is possible even though there are only finitely many hazardous conditions, because for  $i$  large enough,  $\mathcal{T}_i$  can be defined to be a singleton set, and  $H_i: \Omega \rightarrow \mathcal{T}_i$  as the trivial function.

In general, for any choice  $\kappa = \{k_1, \dots, k_m\}$  of system parameters, there is a random variable

$$S_{\kappa} \times H_1 \times H_2 \times \dots : \Omega \rightarrow \mathcal{T} \quad (1)$$

given by  $\chi \mapsto (S \circ (\kappa \times \mathbf{l}_1 \times \mathbf{l}_2 \times \dots \times \mathbf{l}_m))(\chi), \times H_1(\chi), H_2(\chi), \dots)$ . Thus, the type  $\mathcal{T}$  inherits the structure of a probability space from  $(\Omega, \sigma(\Omega))$  and from the random variable (1).

**Definition 1** *Since the random variable (1) depends on the choice  $\kappa$  of parameters for the system  $S$ , the probability distribution of  $\mathcal{T}$  depends on  $\kappa$  as well. Thus, the probability function on  $\mathcal{T}$  induced by  $S$  and  $\kappa$  will be denoted  $P_{\kappa}$ .*

If  $\beta$  is a subtype of  $\mathcal{T}$ , then the probability  $P_{\kappa}[\beta]$  can be defined and possibly computed.

### 2.4 Probabilistic Safety Claims

Suppose that the  $r$  hazardous conditions  $H_1, \dots, H_r$ , the corresponding types  $\mathcal{T}_1, \dots, \mathcal{T}_r$ , and the probability distributions of the random variables  $H_i$  are all known. Let  $\beta_i \in \sigma(\mathcal{T}_i)$  be events in  $\mathcal{T}_i$ . That is, each  $\beta_i$  is a subtype of  $\mathcal{T}_i$ , and the probability that the value of  $H_i$  is an element of  $\beta_i$  can be computed.

In general, the probability that the value of every  $H_i$  (for  $i = 1, \dots, r$ ) is in  $\beta_i$  and that the system  $S$  takes a value in  $\beta$  is given by

$$P_{\kappa}[\sigma(\beta_0, \beta_1, \dots, \beta_r)],$$

where  $\sigma(\beta_0, \beta_1, \dots, \beta_r)$  is the concatenation of  $\sigma$ -algebras given by

$$\sigma(\beta_0, \beta_1, \dots, \beta_r) = \{\omega \in T \mid \omega_0 \in \beta_0, \omega_1 \in \beta_1, \dots, \text{ and } \omega_r \in \beta_r\}.$$

An introduction to concatenations of  $\sigma$ -algebras can be found in [10]. As more sigma algebras are concatenated, the concatenation becomes smaller:

$$\sigma(\beta_0) \supseteq \sigma(\beta_0, \beta_1) \supseteq \sigma(\beta_0, \beta_1, \beta_2) \supseteq \sigma(\beta_0, \beta_1, \beta_2, \beta_3) \supseteq \dots,$$

and the sequence of associated probabilities is decreasing:

$$P_\kappa[\sigma(\beta_0)] \geq P_\kappa[\sigma(\beta_0, \beta_1)] \geq P_\kappa[\sigma(\beta_0, \beta_1, \beta_2)] \geq P_\kappa[\sigma(\beta_0, \beta_1, \beta_2, \beta_3)] \geq \dots$$

With this formalism, it is possible to formally state a safety claim in a way that can be specified in an automated theorem prover. Let  $p_0$  and  $p_1$  be any two probabilities, and let  $\beta_0$  and  $\alpha_0$  be two subtypes of  $\mathcal{T}_0$ .

**Definition 2** *A probabilistic safety claim on the system  $S$  is a statement of the following form: If  $l_1^{meas}, \dots, l_m^{meas}$  are measured values for the variables of the system  $S$  such that the system output value  $S(\kappa'; l_1^{meas}, \dots, l_m^{meas})$  is an element of  $\alpha_0$ , then the probability that the system  $S$ , with parameter set  $\kappa$ , takes values in  $\beta_0$  is between  $p_0$  and  $p_1$ . i.e.*

$$P_\kappa[\sigma(\beta_0)] \in [p_0, p_1]. \quad (2)$$

It should be noted that the hypothesis that  $S(\kappa; l_1^{meas}, \dots, l_m^{meas})$  is an element of  $\alpha_0$  is not needed to formally state a safety claim in a theorem prover. However, such a hypothesis will often be required to prove such a safety claim, because the expected values of the random variables  $\mathbf{l}_1, \dots, \mathbf{l}_m$  are often equal to  $l_1^{meas}, \dots, l_m^{meas}$ , respectively. Thus, the computation of the probability (2) often depends on these measured values.

Another important property of this definition is that the set of system parameters  $\kappa'$  is different than the set  $\kappa$ . In practice, the parameter set  $\kappa'$  may be chosen so that if  $S(\kappa; l_1^{meas}, \dots, l_m^{meas})$  is an element of  $\mathcal{T}_0$ , then the probability  $P_\kappa[\sigma(\beta_0)]$  is more likely to be between  $p_0$  and  $p_1$ . An example of this is given below in Section 3.2, where the radius of the protected zone around an aircraft and the lookahead time for conflict detection are artificially increased to ensure that if a conflict detection probe returns **False**, then the probability that the two aircraft are actually in conflict (using the correct radius and lookahead time) is reduced.

It is also important to note that neither the infinite product  $T$  nor concatenations of sigma algebras are required to make a safety claim on a system. However, as illustrated in Section 2.4, both of these concepts are necessary when developing a formal proof of such a safety claim.

An example of such a safety claim, for a conflict detection probe, is presented in Section 3.

## 2.5 Dependence of System Variables on Hazardous Conditions

In general, the hazardous conditions  $H_i$  for the system  $S$  may have an impact on the accuracy of the variables of  $S$ , which are modeled as random variables  $\mathbf{l}_1, \dots, \mathbf{l}_m$ , as in Section 2.1. It is possible to model the dependence of the random variables  $\mathbf{l}_i$  on the random variables  $H_i$  using probabilistic kernels. This section provides a brief introduction to probabilistic kernels, and the construction follows that in [10].

**Probabilistic Kernels** Suppose that the distribution of the random variable  $S_\kappa: \Omega \rightarrow \mathcal{T}_0$  (the output of the system  $S$ ) depends on the value of  $H_1: \Omega \rightarrow \mathcal{T}_1$ . That is, if  $\omega_1 \in \mathcal{T}_1$ , then there is an associated random variable  $\Omega \rightarrow \mathcal{T}_0 \times \mathcal{T}_1$  given by

$$\chi \mapsto (S_\kappa(\chi), \omega_1), \quad (3)$$

for  $\chi \in \Omega$  and the distribution of this random variable depends on the choice of  $\omega_1$ . If this is the case, then there is an induced probability function

$$p: \mathcal{T}_1 \times \sigma(\mathcal{T}_0) \rightarrow [0, 1].$$

Since this function depends on the parameter  $\kappa$  of the system  $S$ , it will be written as  $p_\kappa$ . Given  $\omega_1 \in \mathcal{T}_1$  and  $\beta_0 \in \mathcal{T}_0$ , the corresponding output of  $p_\kappa$  is written  $p_\kappa(\omega_1; \beta_0)$ , which is the probability that the random variable (3) takes a value in  $\beta_0 \times \{\omega_1\}$ . If  $\beta_0$  and  $\beta_1$  are elements of  $\sigma(\mathcal{T}_0)$  and  $\sigma(\mathcal{T}_1)$ , respectively, then the probability  $P_\kappa[\sigma(\beta_0, \beta_1)]$ , defined in Section 2.4, is given by the Lebesgue integral

$$P_\kappa[\sigma(\beta_0, \beta_1)] = \int_{\omega_1 \in \beta_1} \int_{\omega_0 \in \beta_0} p_\kappa(\omega_1; d\omega_0) p(d\omega_1).$$

It is important to note that there is no assumption of independence required for this equation. In order to compute this integral, it is necessary to know how the random variable  $S_\kappa$  depends on the random variable  $H_1$ .

**Probabilistic Kernels with Several Variables** The construction of this probabilistic kernel can be generalized to handle multiple hazardous conditions as follows. Suppose as above that the random variable  $S_\kappa: \Omega \rightarrow \mathcal{T}_0$  depends on the random variables  $H_1, \dots, H_r$ . Suppose further that for all  $i = 1, \dots, r$ , the random variable  $H_i: \Omega \rightarrow \mathcal{T}_i$  depends on the values of the random variables  $H_{i+1}, \dots, H_r$ . That is,  $S_\kappa$  depends on  $H_1, \dots, H_r$ ;  $H_1$  depends on  $H_2, \dots, H_r$ ;  $H_2$  depends on  $H_3, \dots, H_r$ ; etc. As above, this means that if  $i \geq 0$ , then for  $\omega_{i+1} \in \mathcal{T}_{i+1}, \dots, \omega_r \in \mathcal{T}_r$ , the distribution of the random variable  $\Omega \rightarrow \mathcal{T}_i \times \dots \times \mathcal{T}_r$ , given by

$$\chi \mapsto (H_i(\chi), \omega_{i+1}, \dots, \omega_r), \quad (4)$$

depends on the values of  $\omega_{i+1}, \dots, \omega_r$  (by abuse of notation,  $H_0 = S_\kappa$  in this equation). Further, there is an induced probability function

$$p: \mathcal{T}_r \times \dots \times \mathcal{T}_{i+1} \times \sigma(\mathcal{T}_i) \rightarrow [0, 1]$$

given by  $(\omega_r, \dots, \omega_{i+1}; \beta_i) \mapsto p(\omega_r, \dots, \omega_{i+1}; \beta_i)$ , which is the probability that the random variable (4) takes a value in  $\beta_i \times \{\omega_1\} \times \dots \times \{\omega_r\}$ . This probability is written with a subscript of  $\kappa$  if  $i = 0$  to indicate the dependence on the system parameter  $\kappa$ . If  $\beta_i$  is an element of the  $\sigma$ -algebra  $\sigma(\mathcal{T}_i)$  for  $i = 0, \dots, r$ , then the probability  $P_\kappa[\sigma(\beta_0, \dots, \beta_r)]$  (cf. Section 2.4) is given by the Lebesgue integral

$$\int_{\omega_r \in \beta_r} \dots \int_{\omega_0 \in \beta_0} p_\kappa(\omega_r, \dots, \omega_1; d\omega_0) p(\omega_r, \dots, \omega_1; d\omega_1) \dots p(\omega_r; d\omega_{r-1}) p(d\omega_r).$$

An example of such an integral is given in Section 3.2, where this integral is explicitly computed to prove a safety claim for a conflict detection system.

### 3 A Proved Safety Claim for Conflict Detection

This section illustrates the framework presented in the previous sections with an example of a safety claim for a conflict detection probe in a 2D airspace. This is an algorithm that detects conflicts between two aircraft, referred to here as the *ownship* and the *intruder*. Its variables include the *state* information of the aircraft, which consists of their current positions and velocities, which are represented by points and vectors in  $\mathbb{R}^2$ , respectively.

Aircraft trajectories are represented by a point moving at constant linear speed, i.e., if the current state of an aircraft is given by the position  $\mathbf{s}$  and the velocity vector  $\mathbf{v}$ , then its predicted position at time  $t$  is  $\mathbf{s} + t\mathbf{v}$ . In this paper, the vectors  $\mathbf{s}_o, \mathbf{v}_o, \mathbf{s}_i$ , and  $\mathbf{v}_i$  represent the ownship's position and velocity and the intruder's position and velocity, respectively. The formalization presented here usually considers a relative view where the intruder is fixed at the origin of the coordinate system. The vectors  $\mathbf{s}$  and  $\mathbf{v}$  will denote the relative position  $\mathbf{s}_o - \mathbf{s}_i$  and the relative velocity  $\mathbf{v}_o - \mathbf{v}_i$ , respectively.

In the airspace, it is required that aircraft maintain a certain horizontal separation, specified by a minimum horizontal distance  $D$ . Typically,  $D$  is 5 nautical miles. A conflict detection probe detects conflicts between the aircraft over some given lookahead time  $T$ , usually less than five minutes. A *conflict* between the ownship and the intruder aircraft occurs when there is a time  $t \in [0, T]$  at which the horizontal distance between the aircraft is projected to be less than  $D$ , i.e.,

$$\|(\mathbf{s}_o + t\mathbf{v}_o) - (\mathbf{s}_i + t\mathbf{v}_i)\| < D.$$

Since  $(\mathbf{s}_o + t\mathbf{v}_o) - (\mathbf{s}_i + t\mathbf{v}_i) = (\mathbf{s}_o - \mathbf{s}_i) + t(\mathbf{v}_o - \mathbf{v}_i)$ , the predicate that characterizes conflicts can be defined in terms of the relative vectors  $\mathbf{s} = \mathbf{s}_o - \mathbf{s}_i$  and  $\mathbf{v} = \mathbf{v}_o - \mathbf{v}_i$ , i.e., the relative position and velocity vectors, respectively, of the ownship with respect to the intruder. The predicate *horizontal\_conflict?*, parametric on the lookahead time  $T$  and the horizontal distance  $D$ , is formally defined as follows.

$$\text{horizontal\_conflict?}(D, T, \mathbf{s}, \mathbf{v}) \equiv \exists t \in [0, T] : \|\mathbf{s} + t\mathbf{v}\| < D.$$

A conflict detection probe is an algorithm that computes whether the predicate *horizontal\_conflict?* holds for the current states of two aircraft. One example of such an algorithm is *cd2d*, developed at NASA Langley [6]. Formally, a conflict detection probe is defined as a function

$$\text{cd}: \mathbb{R}^+ \times \mathbb{R}^+; \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \{\text{True}, \text{False}\}.$$

It is designed so that  $\text{cd}(D, T; \mathbf{s}, \mathbf{v}) \iff \text{horizontal\_conflict?}(D, T, \mathbf{s}, \mathbf{v})$ , for all  $D, T \in \mathbb{R}^+$  and  $\mathbf{s}, \mathbf{v} \in \mathbb{R}^2$ . Such a conflict detection probe is a system, as described above. The distance  $D$  and time  $T$  are parameters of *cd* because their values are typically known to the aircraft without error. For instance, the airspace may have a 5 nautical mile minimum horizontal separation, and a standards document may define the lookahead time  $T$  to be 3 minutes.

### 3.1 GPS and ADS-B Hazardous Conditions

If the ownship is using the conflict probe **cd** to detect conflicts, it must depend on broadcast signals from the intruder to determine the intruder's position and velocity vectors. In this example, the aircraft use Automatic Dependent Surveillance Broadcast (ADS-B)[8] messages to communicate their positions and velocities, and it is assumed that ADS-B messages with state information are sent by each aircraft once per second. When the ownship uses the algorithm **cd**, it is possible that several consecutive position and velocity updates from the intruder have been dropped due to signal attenuation, which results in greater uncertainty in the values of  $\mathbf{s}_i$  and  $\mathbf{v}_i$ . Thus, ADS-B message loss due to signal attenuation can be modeled as a hazardous condition:

$$H_{2,adsb}: \Omega \rightarrow \mathcal{T}_{2,adsb} \quad \mathcal{T}_{2,adsb} = \{0, 1, 2, 3, \dots\}.$$

The random variable  $H_{2,adsb}$  returns the number of consecutive ADS-B messages from the intruder that were not received by the ownship, since the last received message from the intruder. At a given instant of time when a conflict detection probe is used,  $\tau$  will be used to represent this number of consecutive dropped messages. The number  $\tau$  is easy for the ownship to compute, since it just has to know when the last ADS-B update from the intruder was received. The number  $\tau$  is an integer, and  $\tau_s$  will be used to represent the time period  $\tau$  seconds.

In addition, if the conflict detection probe **cd** is being used by the ownship, then the position and velocity vectors  $\mathbf{s}_o$ ,  $\mathbf{s}_i$ ,  $\mathbf{v}_o$ , and  $\mathbf{v}_i$  will be estimated using instruments such as GPS. These instruments can be faulty or have expected errors. For instance, there may be some error in the position predicted by a GPS device. The effects of uncertainty in positions and velocities of aircraft on conflict detection have been studied before [3].

Error in GPS is modeled as a hazardous condition as follows. The vectors  $\mathbf{s}_i^m$  and  $\mathbf{v}_i^m$  represent the intruder's reported position and velocity vectors, respectively, from the last ADS-B signal that was received by the ownship, and the vectors  $\mathbf{s}_o^m$  and  $\mathbf{v}_o^m$  represent the ownship's measured position and velocity at that time. The relative vectors  $\mathbf{s}^m$  and  $\mathbf{v}^m$  are defined by  $\mathbf{s}^m = \mathbf{s}_o^m - \mathbf{s}_i^m$  and  $\mathbf{v}^m = \mathbf{v}_o^m - \mathbf{v}_i^m$ . The true positions of the ownship and the intruder at the time when the vectors  $\mathbf{s}^m$  and  $\mathbf{v}^m$  were measured ( $\tau$  seconds ago) are given by  $\mathbf{s}_o - \tau_s \mathbf{v}_o$  and  $\mathbf{s}_i - \tau_s \mathbf{v}_i$ , respectively. It is clear that if the measured vectors  $\mathbf{s}_o^m, \mathbf{v}_o^m, \mathbf{s}_i^m$ , and  $\mathbf{v}_i^m$  have no error, then  $\mathbf{s}^m = \mathbf{s} - \tau_s \mathbf{v}$  and  $\mathbf{v}^m = \mathbf{v}$ . In this case, if  $\text{cd}(D, T + \tau_s; \mathbf{s}^m, \mathbf{v}^m) = \text{False}$ , then  $\text{cd}(D, T; \mathbf{s}, \mathbf{v}) = \text{False}$  as well. Thus, the symbol **e** (called *GPS error*) denotes the fact that one of the following inequalities is satisfied.

$$\begin{aligned} (i) \quad & \|(\mathbf{s}_o - \tau_s \mathbf{v}_o) - \mathbf{s}_o^m\| \geq a_o & \|(\mathbf{s}_i - \tau_s \mathbf{v}_i) - \mathbf{s}_i^m\| \geq a_i & \text{(iii)} \\ (ii) \quad & \|\mathbf{v}_o - \mathbf{v}_o^m\| \geq b_o & \|\mathbf{v}_i - \mathbf{v}_i^m\| \geq b_i & \text{(iv)} \end{aligned}$$

Here, the distances  $a_o$  and  $a_i$  and the speeds  $b_o$  and  $b_i$  are predetermined parameters. For instance, one set of these parameters that is used in the proof of a safety claim in Section 3.3 is  $a_o, a_i = 30$  m and  $b_o, b_i = 0.3$  m/s, which correspond to



certain navigation accuracy categories ( $\text{NAC}_P$  9 and  $\text{NAC}_V$  4, respectively), as specified by RTCA, Inc. in DO-242A for precision in ADS-B messages [8]. This specification is for 95 percent confidence intervals on the position and velocity vectors of aircraft, within the given ranges. Other choices for  $a_o, a_i, b_o$ , and  $b_i$  may be considered, and thus in the next few sections they are simply treated as variables.

With this construction, GPS error is modeled as a hazardous condition

$$H_{1,gps} : \Omega \rightarrow \mathcal{T}_{1,gps} \quad (\text{where } \mathcal{T}_{1,gps} = \{\mathbf{e}, \neg\mathbf{e}\}).$$

The return type  $\mathcal{T}_{2,adsb}$  of the second hazardous condition  $H_{2,adsb}$  represents the number of seconds since the last ADS-B update from the intruder aircraft. If  $d$  is any non-negative integer, it is possible to formally define the probability that the most recent ADS-B message that was sent by the intruder and detected/decoded by the ownship occurred within the last  $d$  seconds.

As noted above, inaccuracies in the measurements of the positions  $\mathbf{s}_o$  and  $\mathbf{s}_i$  and the velocities  $\mathbf{v}_o$  and  $\mathbf{v}_i$  imply that the conflict detection probe  $\mathbf{cd}$  can be modeled as a random variable:

$$\begin{aligned} \mathbf{cd}_{D,T} : \Omega &\rightarrow \mathcal{T}_0 = \{\mathbf{True}, \mathbf{False}\} \\ \chi &\mapsto \mathbf{cd}(D, T; \mathbf{s}(\chi), \mathbf{v}(\chi)) \end{aligned}$$

This random variable depends on the hazardous conditions  $H_{1,gps}$  and  $H_{2,adsb}$ .

### 3.2 Probabilistic Kernels in Conflict Detection

It is clear that the random variable  $S_{D,T}$ , which takes values in  $\{\mathbf{False}, \mathbf{True}\}$ , depends on the hazardous conditions  $H_{1,gps}$  and  $H_{2,adsb}$ . Thus, as in Section 2.5, if  $\beta_2 \subset \mathcal{T}_{2,adsb}$ ,  $\beta_1 \subset \mathcal{T}_{1,gps}$ , and  $\beta_0 \subset \mathcal{T}_0 = \{\mathbf{False}, \mathbf{True}\}$ , then the probability that  $H_{2,adsb}$  and  $H_{1,gps}$  take values in  $\beta_2$  and  $\beta_1$ , respectively, and that  $\mathbf{cd}_{D,T}$  takes a value in  $\beta_0$ , is given by

$$P_{D,T}[\sigma(\beta_0, \beta_1, \beta_2)] = \int_{\omega_2 \in \beta_2} \int_{\omega_1 \in \beta_1} \int_{\omega_0 \in \beta_0} p_{D,T}(\omega_1, \omega_2; d\omega_0) p(\omega_2; d\omega_1) p(d\omega_2).$$

As a simple example of this, if  $i \in \mathcal{T}_{2,adsb}$ , then the probability that the random variable (conflict probe)  $\mathbf{cd}_{D,T}$  returns  $\mathbf{True}$ , that there is no error in GPS, and that the last ADS-B signal from the intruder aircraft was exactly  $i$  seconds ago is given by

$$\begin{aligned} &P_{D,T}[\sigma(\{\mathbf{True}\}, \{\neg\mathbf{e}\}, \{i\})] \\ &= \int_{\omega_2 \in \{i\}} \int_{\omega_1 \in \{\neg\mathbf{e}\}} \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\omega_1, \omega_2; d\omega_0) p(\omega_2; d\omega_1) p(d\omega_2) \\ &= \int_{\omega_1 \in \{\neg\mathbf{e}\}} \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\omega_1, i; d\omega_0) p(i; d\omega_1) p(\{i\}) \\ &= \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\neg\mathbf{e}, i; d\omega_0) p(i; \{\neg\mathbf{e}\}) p(\{i\}) \\ &= p_{D,T}(\neg\mathbf{e}, i; \{\mathbf{True}\}) p(i; \{\neg\mathbf{e}\}) p(\{i\}) \end{aligned}$$

The random variables  $\mathbf{cd}_{D,T}$ ,  $H_{1,gps}$ , and  $H_{2,adsb}$  are all discrete, so the probability that  $\mathbf{cd}_{D,T}$  returns **True**, which is given by  $P_{D,T}[\sigma(\{\mathbf{True}\})]$ , can be computed as an infinite sum as follows.

$$\begin{aligned}
P_{D,T}[\sigma(\{\mathbf{True}\})] &= \int_{\omega_2 \in \{0,1,2,\dots\}} \int_{\omega_1 \in \{\mathbf{e}, \neg\mathbf{e}\}} \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\omega_1, \omega_2; d\omega_0) p(\omega_2; d\omega_1) p(d\omega_2) \\
&= \sum_{i=0}^{\infty} \int_{\omega_1 \in \{\mathbf{e}, \neg\mathbf{e}\}} \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\omega_1, i; d\omega_0) p(i; d\omega_1) p(i) \\
&= \sum_{i=0}^{\infty} \left( \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\mathbf{e}, i; d\omega_0) p(i; \{\mathbf{e}\}) p(\{i\}) \right. \\
&\quad \left. + \int_{\omega_0 \in \{\mathbf{True}\}} p_{D,T}(\neg\mathbf{e}, i; d\omega_0) p(i; \{\neg\mathbf{e}\}) p(\{i\}) \right) \\
&= \sum_{i=0}^{\infty} (p_{D,T}(\mathbf{e}, i; \{\mathbf{True}\}) p(i; \{\mathbf{e}\}) p(\{i\}) \\
&\quad + p_{D,T}(\neg\mathbf{e}, i; \{\mathbf{True}\}) p(i; \{\neg\mathbf{e}\}) p(\{i\}))
\end{aligned} \tag{5}$$

**Distribution of the ADS-B Hazardous Condition** Under the assumption that there is no ADS-B signal interference due to multiple intruder aircraft, the distribution of the hazardous condition  $H_{2,adsb}$  follows a Poisson distribution, as discussed in [2]. In that paper, the probability that a given ADS-B message from the intruder aircraft will not be detected and decoded by the ownship, which is equal to  $p(\{0\})$ , is (approximately) given by  $p(\{0\}) = 1 - \left(\frac{r}{r_0}\right)^k$  with  $r \leq r_0$ , where  $k = 6.4314$  and  $r_0 = 96.6$  nmi [2]. The number  $r$  is the current distance between the two aircraft. Thus, if it is known that the ownship and the intruder are no greater than 60 nmi apart, a reasonable distance for most commercial aircraft given short lookahead times such as 3 minutes, then  $p(\{0\}) \geq \eta$ , where

$$\eta = 0.953.$$

The key assumption that can be used to deduce that  $H_{2,adsb}$  follows a Poisson distribution is that whether any particular ADS-B message from the intruder aircraft is received by the ownship is independent from whether any other, different, ADS-B message from the intruder is received. Under this assumption,

$$p(\{i\}) = \eta(1 - \eta)^i \quad \text{for } i \geq 0.$$

This is because the last  $i$  messages (sent 0, 1,  $\dots$  and  $i-1$  seconds ago) have been dropped, which has a probability of  $(1 - \eta)^i$  of occurring, and the message sent exactly  $i$ -seconds ago was not dropped, which has a probability of  $\eta$  of occurring. The equation above can be used to replace  $p(\{i\})$  in Equation (5).

**Probability of GPS Error** A key assumption in this example is that probabilities  $p_{so}$ ,  $p_{si}$ ,  $p_{vo}$  and  $p_{vi}$  are known that satisfy the following properties.

- At any given time, the probability, that the distance between the ownship's predicted position (by GPS) and its actual position is at least  $a_o$ , is bounded above by  $p_{so}$ .
- At any given time, the probability, that the difference (speed) between the ownship's predicted velocity (by GPS) and its actual velocity is at least  $b_o$ , is bounded above by  $p_{vo}$ .
- At any given time, the probability, that the distance between the intruder's predicted position (by GPS) and its actual position is at least  $a_i$ , is bounded above by  $p_{si}$ .
- At any given time, the probability, that the difference (speed) between the intruder's predicted velocity (by GPS) and its actual velocity is at least  $b_i$ , is bounded above by  $p_{vi}$ .

Specific examples of such numbers can be found in the RTCA, Inc. document DO-242A [8], which provides examples for the analyses in Section 3.3.

At a given instant of time, the actual positions of the ownship and the intruder  $\tau$  seconds ago were given by  $\mathbf{s}_o - \tau_s \mathbf{v}_o$  and  $\mathbf{s}_i - \tau_s \mathbf{v}_i$ , respectively. The positions at that time, as predicted by GPS, are by definition given by  $\mathbf{s}_o^m$  and  $\mathbf{s}_i^m$ , respectively. Thus, the following four equations hold.

$$\begin{aligned} P[||(\mathbf{s}_o - \tau^m \mathbf{v}_o) - \mathbf{s}_o^m|| \geq a_o] &\leq p_{so} & P[||(\mathbf{s}_i - \tau^m \mathbf{v}_i) - \mathbf{s}_i^m|| \geq a_i] &\leq p_{si} \\ P[||\mathbf{v}_o - \mathbf{v}_o^m|| \geq b_o] &\leq p_{vo} & P[||\mathbf{v}_i - \mathbf{v}_i^m|| \geq b_i] &\leq p_{vi} \end{aligned}$$

By the definition of the error  $\mathbf{e}$  in Section 3.1,  $p(i; \{\mathbf{e}\}) \leq p_{so} + p_{vo} + p_{si} + p_{vi}$ . Set  $p_{error} = p_{so} + p_{vo} + p_{si} + p_{vi}$ . Equation (5) implies that if  $d$  is any integer (a specific number of seconds), then

$$\begin{aligned} &P_{D,T}[\sigma(\{\mathbf{True}\})] \\ &= \sum_{i=0}^{\infty} (p_{D,T}(\mathbf{e}, i; \{\mathbf{True}\})p(i; \{\mathbf{e}\})p(\{i\}) \\ &\quad + p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})p(i; \{\neg \mathbf{e}\})p(\{i\})) \\ &\leq \sum_{i=0}^{\infty} (P_{error} \eta(1-\eta)^i + p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})p(i; \{\neg \mathbf{e}\})p(\{i\})) \\ &= P_{error} + \sum_{i=0}^{\infty} p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})p(i; \{\neg \mathbf{e}\})p(\{i\}) \\ &\leq P_{error} + \sum_{i=0}^{\infty} p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})\eta(1-\eta)^i \\ &\leq P_{error} + \sum_{i=d+1}^{\infty} \eta(1-\eta)^i + \sum_{i=0}^d p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})\eta(1-\eta)^i \\ &= P_{error} + (1-\eta)^{d+1} + \sum_{i=0}^d p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\})\eta(1-\eta)^i \end{aligned} \tag{6}$$

The number  $d$ , which is an element of  $\mathcal{T}_{2,adsb}$  can be chosen so that the finite sum is a good approximation to the infinite sum (since  $(1 - \eta)^{d+1}$  is quite small). This equation is true for any choice of  $d$ .

**An Upper Bound on the Probability of Failure** Equation (6) implies that if  $p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\}) = 0$  for  $i \in \{0, \dots, d\}$ , then the probability that  $\mathbf{cd}(D, T; \mathbf{s}, \mathbf{v}) = \mathbf{True}$ , which is given by  $P_{D,T}[\sigma(\{\mathbf{True}\})]$ , is bounded above by  $P_{error} + (1 - \eta)^{d+1}$ . As noted in Section 2.4, to mitigate the effect of measurement errors on the conflict detection probe  $\mathbf{cd}$ , a positive distance  $\psi$  and a positive time  $\lambda$  can be artificially added to the distance  $D$  and the time  $T$  when they are used as parameters in  $\mathbf{cd}$ . The important question here is how large do  $\psi$  and  $\lambda$  need to be so that if  $\mathbf{cd}(D + \psi, T + \lambda; \mathbf{s}^m, \mathbf{v}^m) = \mathbf{False}$ , then  $p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\}) = 0$  for  $i \in \{0, \dots, d\}$ . This question is answered by the following lemma. It refers to the distances  $a_o$  and  $a_i$  and the speeds  $b_o$  and  $b_i$  that define the probabilities  $p_{so}, p_{vo}, p_{si}, p_{vi}$  (cf. Section 3.1).

**Lemma 1.** *If  $\lambda = d$  seconds,  $\psi = a_o + a_i + (T + \lambda)(b_o + b_i)$ , and  $\mathbf{cd}(D + \psi, T + \lambda; \mathbf{s}^m, \mathbf{v}^m) = \mathbf{False}$ , then  $p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\}) = 0$  for  $i \in \{0, \dots, d\}$ .*

*Proof.* Suppose that  $\neg \mathbf{e}$  holds, and recall from Section 3.1 that  $\tau$  denotes the number of seconds since the ownship successfully received position and velocity updates from the intruder aircraft's ADS-B device. Suppose that  $\tau = i$ , where  $i \leq d$ . Then in order to show that  $p_{D,T}(\neg \mathbf{e}, i; \{\mathbf{True}\}) = 0$ , it suffices to prove that  $\mathbf{cd}(D, T; \mathbf{s}, \mathbf{v}) = \mathbf{False}$ . Since  $\tau \leq d$ , it follows from the hypotheses of the lemma that  $\mathbf{cd}(D + \psi, T + \tau_s; \mathbf{s}^m, \mathbf{v}^m) = \mathbf{False}$ . Further, since  $\neg \mathbf{e}$  holds, the equations  $\|(\mathbf{s}_o - (i \text{ sec})\mathbf{v}_o) - \mathbf{s}_o^m\| < a_o$  and  $\|(\mathbf{s}_i - (i \text{ sec})\mathbf{v}_i) - \mathbf{s}_i^m\| < a_i$  and  $\|\mathbf{v}_o - \mathbf{v}_o^m\| < b_o$  and  $\|\mathbf{v}_i - \mathbf{v}_i^m\| < b_i$  are all satisfied.

By contradiction, suppose that  $\mathbf{cd}(D, T; \mathbf{s}, \mathbf{v}) = \mathbf{True}$ , and choose  $t^* \in [0, T]$  such that  $\|\mathbf{s} + t^*\mathbf{v}\| < D$ . Then  $t^* + \tau_s \in [0, T + \lambda]$  and since  $\mathbf{s} = \mathbf{s}_o - \mathbf{s}_i$  and  $\mathbf{v} = \mathbf{v}_o - \mathbf{v}_i$ , it follows that

$$\begin{aligned}
& \|\mathbf{s}^m + (t^* + \tau_s)\mathbf{v}^m\| \\
&= \|(\mathbf{s}_o^m - \mathbf{s}_i^m) + (t^* + (i \text{ sec}))(\mathbf{v}_o^m - \mathbf{v}_i^m)\| \\
&= \|(\mathbf{s}_o^m - \mathbf{s}_i^m) + (t^* + (i \text{ sec}))(\mathbf{v}_o^m - \mathbf{v}_i^m) - (\mathbf{s} + t^*\mathbf{v}) + (\mathbf{s} + t^*\mathbf{v})\| \\
&= \|(\mathbf{s}_o^m - (\mathbf{s}_o - (i \text{ sec})\mathbf{v}_o)) - (\mathbf{s}_i^m - (\mathbf{s}_i - (i \text{ sec})\mathbf{v}_i)) + (t^* + (i \text{ sec}))(\mathbf{v}_o^m - \mathbf{v}_o) \\
&\quad - (t^* + (i \text{ sec}))(\mathbf{v}_i^m - \mathbf{v}_i) + (\mathbf{s} + t^*\mathbf{v})\| \\
&\leq \|\mathbf{s}_o^m - (\mathbf{s}_o - (i \text{ sec})\mathbf{v}_o)\| + \|\mathbf{s}_i^m - (\mathbf{s}_i - (i \text{ sec})\mathbf{v}_i)\| + (t^* + (i \text{ sec}))\|\mathbf{v}_o^m - \mathbf{v}_o\| \\
&\quad + (t^* + (i \text{ sec}))\|\mathbf{v}_i^m - \mathbf{v}_i\| + \|\mathbf{s} + t^*\mathbf{v}\| \\
&< a_o + a_i + (t^* + \lambda)b_o + (t^* + \lambda)b_i + D \\
&\leq a + (t^* + (i \text{ sec}))b + D \\
&\leq \psi + D.
\end{aligned}$$

This is a contradiction, since  $\mathbf{cd}(D + \psi, T + \lambda; \mathbf{s}^m, \mathbf{v}^m) = \mathbf{False}$  and  $\lambda = d$  seconds. This completes the proof.  $\square$

### 3.3 The Safety Claim for Conflict Detection

The safety claim that can be proved by using Lemma 1 is stated below. It has not been formally proved in a theorem prover, but the formal mathematics has been developed in this paper that enables a standard mathematical proof. It follows trivially from that Lemma and from Equation 6 in Section 3.2.

**Proved Safety Claim for the Conflict Probe cd:** *Let  $\lambda = d$  seconds,  $\psi = a_o + a_i + (T + \lambda)(b_o + b_i)$ . Suppose that  $cd(D + \psi, T + \lambda; \mathbf{s}^m, \mathbf{v}^m) = \text{False}$  and that the ownship and the intruder aircraft are no greater than 60 nmi apart. Then the probability that the aircraft are in conflict, i.e. that  $cd(D, T; \mathbf{s}, \mathbf{v}) = \text{True}$ , is no greater than  $p_{so} + p_{vo} + p_{si} + p_{vi} + (1 - \eta)^{d+1}$ .*

A *missed alert* is a conflict that is not detected. Artificially increasing the distance  $D$  and the lookahead time  $T$  in the conflict probe *cd* will make missed alerts less likely. The proved safety claim above gives a formula that returns the amount that  $D$  and  $T$  must be increased, as well as an upper bound on the probability of a missed alert if  $D$  is increased in this way, assuming that the ownship and the intruder aircraft are within 60 nmi of each other. The inputs to these formulas are the distances  $a_o$  and  $a_i$ , the speeds  $b_o$  and  $b_i$ , the probabilities  $p_{so}, p_{vo}, p_{si}$  and  $p_{vi}$ , and the number of seconds  $d$  that  $T$  is to be increased in the conflict probe *cd*. Equation (6) expresses the relationships between  $a_o, a_i, b_o, b_i, p_{so}, p_{vo}, p_{si}$  and  $p_{vi}$ . Given these inputs, the associated upper bound for the probability of a missed alert is

$$p_{\text{missed-alert}} = p_{so} + p_{vo} + p_{si} + p_{vi} + (1 - \eta)^{d+1}, \quad (7)$$

where, as in Section 3.2,  $\eta$  is a lower bound for the probability that a given ADS-B message from the intruder aircraft will not be detected and decoded by the ownship, and in this example  $\eta = 0.953$ .

In the equation above, the amount  $\psi$  that  $D$  should be artificially increased to ensure that the probability of a missed alert is less than  $p_{\text{missed-alert}}$  is given by

$$\psi = a_o + a_i + (T + \lambda)(b_o + b_i), \quad (8)$$

where  $\lambda = d$  second. It should be noted that Equations (8) and (7) imply that if the velocity  $b$  dominates the calculation of  $\psi$ , then as  $\psi$  increases,  $d$  increases as well, and so the probability of a missed alert decreases.

**Computing Actual Probabilities** DO-242A [8] specifies several system performance confidence-levels that are to be included in ADS-B messages detailing how precise and trusted the contained state information is. The relevant ones here are the navigation accuracy categories for position and velocity ( $\text{NAC}_P$  and  $\text{NAC}_V$ ).  $\text{NAC}_P$  is a maximum distance for errors in position; similarly  $\text{NAC}_V$  is a maximum velocity error. That is, these numbers specify the parameters  $a_o, a_i$  and  $b_o, b_i$ , respectively. Both  $\text{NAC}_P$  and  $\text{NAC}_V$  specify that the stated values will fall within a 95% confidence interval, which is equivalent to saying that  $p_{so}, p_{vo}, p_{si}$  and  $p_{vi}$  are all equal to 0.05. Table 1 uses these numbers along with

Equations (7) and (8) to compute the amount the distance that  $D$  needs to be increased, as well the associated upper bounds on the probabilities of missed alerts for different choices of the number of seconds  $d$ .

Position Error	Velocity Error	Time $+\lambda$	Buffer $\psi$	$p_{\text{missed-alert}}$
$< 30$ m	$< 0.3$ m/s	180+0 sec	+0.09 nmi (168 m)	0.24700
$< 30$ m	$< 0.3$ m/s	180+1 sec	+0.09 nmi (169 m)	0.20221
$< 30$ m	$< 0.3$ m/s	180+2 sec	+0.09 nmi (169 m)	0.20010
$< 30$ m	$< 0.3$ m/s	180+3 sec	+0.09 nmi (170 m)	0.20000
$< 185.2$ m	$< 1.0$ m/s	180+0 sec	+0.39 nmi (730 m)	0.24700
$< 185.2$ m	$< 1.0$ m/s	180+3 sec	+0.40 nmi (736 m)	0.20000

**Table 1.** Horizontal uncertainty, lookahead, and buffer sizes. The  $< 30$  m position error corresponds to the NAC<sub>P</sub> 9 error category (NAC<sub>P</sub> 11 is the most accurate) and the  $< 0.3$  m/s velocity error corresponds to the NAC<sub>V</sub> 4 (most accurate) error category. The velocity error dominates in calculating  $\psi$  these cases. When the position error is  $< 185.2$  m (NAC<sub>P</sub> 7) and the velocity error is  $< 1.0$  m/s (NAC<sub>V</sub> 3) the position error dominates the calculation of  $\psi$  for lookahead times less than 186 seconds.

It should be noted that the upper bounds on the probabilities of missed alerts in this table are quite high, but that this is not due to imprecision in the presented methods. This is mostly due to the fact that the confidence intervals specified in DO-242A are for 95% confidence and provide little knowledge of what is happening the other 5% of the time. It is quite possible that these formulas could calculate the probability of missed alerts to be less than  $4 \times 10^{-9}$ , if  $1 - (10^{-9})$ -confidence intervals were available for the positions and velocities of the aircraft.

## 4 Conclusion and Future Work

This paper has built on Rushby and Littlewood’s framework [9, 5] for formalizing safety claims, specifically providing a mathematical basis for dealing with certain probabilistic safety claims. The mathematics behind this is based on the notion of probabilistic kernels, which were illustrated in a safety claim for a conflict detection system for aircraft. The framework presented allows for an arbitrary number of potentially hazardous conditions. Future work in this area will include formalizing the mathematics presented here in a theorem prover such as PVS [7]. Many of the tools needed for this task already exist, including PVS libraries for Riemann integration [1] and Riemann-Stieltjes integration, as well as a Lebesgue measure and integration library developed by David Lester. Some additions are needed to these libraries to facilitate manipulations of multiple integrals.

An additional area for future work would be to incorporate a degree of assumption checking into the framework. This may include formally capturing the assumptions of independence between hazardous conditions, which could

be formed into a verification condition that can be automatically checked for inconsistencies by a satisfiability checker (a SAT-solver).

## References

1. Butler, R.: Formalization of the integral calculus in the PVS theorem prover. *Journal of Formalized Reasoning* 2(1) (2009)
2. Chung, W.W., Staab, R.: A 1090 extended squitter automatic dependent surveillance broadcast (ADS-B) reception model for air-traffic-management simulations. In: *AIAA Modeling and Simulation Technologies Conference and Exhibit* (2006)
3. Herencia-Zapana, H., Jeannin, J.B., Muñoz, C.: Formal verification of safety buffers for state-based conflict detection and resolution. In: *Proceedings of 27th International Congress of the Aeronautical Sciences, ICAS 2010. Nice, France* (2010)
4. Holloway, C.M.: Safety case notations: alternatives for the non-graphically inclined? In: *3rd IET International Conference on System Safety* (2008)
5. Littlewood, B., Rushby, J.: Reasoning about the reliability of diverse two channel systems in which one channel is possibly perfect. In: *Tech report SRI-CSL-09-02* (2010)
6. NASA Langley Formal Methods Team: Airborne coordinated conflict resolution and detection (2010), <http://shemesh.larc.nasa.gov/people/cam/ACCoRD/>
7. Owre, S., Rushby, J.M., , Shankar, N.: PVS: A prototype verification system. In: Kapur, D. (ed.) *11th International Conference on Automated Deduction (CADE)*. *Lecture Notes in Artificial Intelligence*, vol. 607, pp. 748–752. Springer-Verlag, Saratoga, NY (Jun 1992), <http://www.csl.sri.com/papers/cade92-pvs/>
8. Minimum aviation system performance standards for automatic dependent surveillance broadcast (ADS-B). DO-242A, RTCA (June 2002), section 2.1.2.12–2.1.2.15
9. Rushby, J.: Formalism in safety cases. In: *Proceedings of the eighteenth safety-critical systems symposium* (2010)
10. Shiryayev, A.N.: *Probability*. Springer-Verlag (1995)