Appendix A

Planck Radiation Law and Radiative Transfer

A.1 Planck radiation law

Blackbody objects (i.e., perfect emitters of energy) whose temperatures are greater than absolute zero emit energy E per unit volume and per unit frequency at all wavelengths according to the Planck radiation law

$$E = \frac{8\pi h f^3}{c^3} \frac{1}{\exp\left(\frac{hf}{k_B T}\right) - 1} \frac{\text{J}}{\text{m}^3\text{Hz}}, \tag{A-1}$$

where

h = Planck's constant = 6.6256×10^{-34} J-s,

k_B = Boltzmann's constant = 1.380662×10^{-23} J/K,

c = speed of light = 3×10^8 m/s,

T = physical temperature of the emitting object in degrees K,

f = frequency at which the energy is measured in Hz.

Upon expanding the exponential term in the denominator, Eq. (A-1) may be rewritten as

$$E = \frac{8\pi h f^3}{c^3} \frac{1}{\frac{hf}{k_B T} + \left(\frac{hf}{k_B T}\right)^2 + \ldots} \frac{\text{J}}{\text{m}^3\text{Hz}}. \tag{A-2}$$

For frequencies f less than $k_B T/h$ ($\approx 6 \times 10^{12}$ Hz at 300 K), only the linear term in temperature is retained and Planck's radiation law reduces to the Rayleigh-Jeans law given by

$$E = \frac{8\pi f^2 k_B T}{c^3} \frac{\text{J}}{\text{m}^3\text{Hz}}. \tag{A-3}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

In the Rayleigh-Jeans approximation, temperature is directly proportional to the energy of the radiating object, making calibration of a radiometer simpler. With perfect emitters or blackbodies, the physical temperature of the object T is equal to the brightness temperature T_B that is detected by a radiometer. However, the surfaces of real objects do not normally radiate as blackbodies (i.e., they are not 100 percent efficient in emitting the energy predicted by the Planck radiation law). To account for this nonideal emission, a multiplicative emissivity factor is added to represent the amount of energy that is radiated by the object, now referred to as a graybody. The emissivity is equal to the ratio of T_B to T.

When microwave radiometers are used in space applications, the first three terms of the exponential series [up to and including the second-order term containing $(hf/k_BT)^2$] in the denominator of Eq. (A-2) are retained, because the background temperature of space is small compared to the background temperatures on Earth. Including the quadratic term in temperature minimizes the error that would otherwise occur when converting the measured energy into atmospheric temperature profiles used in weather forecasting. The magnitude of the error introduced when the quadratic term is neglected is shown in Table A.1 as a function of frequency.

Table A.1 Effect of quadratic correction term on emitted energy calculated from Planck radiation law (T = 300 K).

f (GHz)	hf/k_BT	$(hf/k_BT)^2$	% change in E
2	0.0003199	1.0233×10^{-07}	0.03198812
6	0.0009598	9.2122×10^{-07}	0.09598041
22	0.0035192	1.2385×10^{-05}	0.35192657
60	0.0095977	9.2116×10^{-05}	0.95977161
118	0.0188755	0.00035628	1.88752616
183	0.0292730	0.00085691	2.92730503
320	0.0511878	0.00262019	5.11877830

Because of emission from molecules not at absolute zero, the atmosphere emits energy that is detected by passive sensors that directly or indirectly (such as by reflection of energy from surfaces whose emissivity is not unity) view the atmosphere. The atmospheric emission modifies and may prevent the detection of ground-based and space-based objects of interest by masking the energy emitted by low-temperature or low-emissivity objects. In contrast, radiometers used in weather forecasting applications operate at atmospheric absorption bands in order to measure the quantity of atmospheric constituents such as oxygen and water vapor. In both cases, radiative transfer theory is utilized to calculate the effects of the atmosphere on the energy measured by the radiometer.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

A.2 Radiative transfer theory

Radiative transfer theory describes the contribution of cosmic, galactic, atmospheric, and ground-based emission sources to the passive signature of objects in a sensor's field of view.[1,2] In Figure A.1, a radiometer is shown flying in a missile or gun-fired round at a height h above the ground and is pointed toward the ground. If the application was weather forecasting, the radiometer would be located in a satellite.

The quantity T_C represents the sum of the cosmic brightness temperature and the galactic brightness temperature. The cosmic temperature is independent of frequency and zenith angle and is equal to 2.735 K. Its origin is attributed to the background radiation produced when the universe was originally formed. The galactic temperature is due to radiation from the Milky Way galaxy and is a function of the viewing direction and frequency. Above about 10 GHz, the galactic contribution may be neglected in comparison with the downward emission from the atmosphere.

The variable T_D represents the total downward atmospheric emission, including the variation of temperature $T(z)$ and absorption coefficient $\kappa_\alpha(z)$ with height, and is equal to

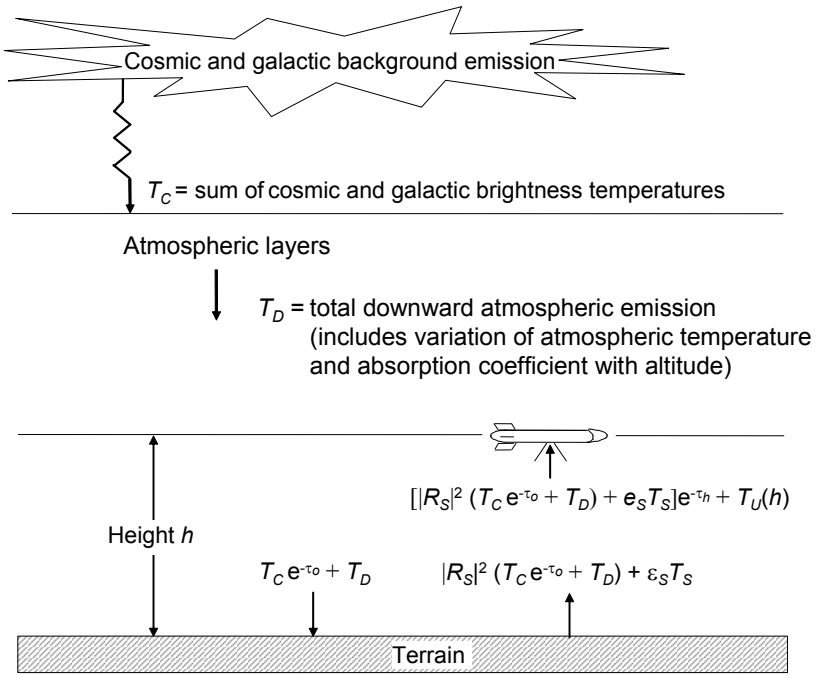

Figure A.1 Radiative transfer in an Earth-looking radiometer sensor.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$T_D = \int_0^{\infty} \{T(z)\exp[-\sec\theta \int_0^z \kappa_\alpha(z')\,dz']\kappa_\alpha(z)\}\,dz\sec\theta\,, \tag{A-4}$$

where

$\kappa_\alpha(z)$ = absorption coefficient of the atmosphere at an altitude z and

θ = incidence angle with respect to nadir as defined in Figure A.2.

The dependence of the absorption coefficient on altitude accounts for the energy emitted by the atmosphere through its constituent molecules such as water vapor, oxygen, ozone, carbon dioxide, and nitrous oxide. Since emission occurs throughout the entire atmospheric height profile, the integration limits for T_D are from 0 to infinity.

The quantity T_U is the upward atmospheric emission in the region from the ground to the height h at which the sensor is located. It is given by

$$T_U = \int_0^{h} \{T(z)\exp[-\sec\theta \int_0^z \kappa_\alpha(z')\,dz']\kappa_\alpha(z)\}\,dz\sec\theta\,. \tag{A-5}$$

The quantity τ_o is the total one-way opacity (integrated attenuation) through the atmosphere. When $\theta < 70$ deg and the atmosphere is spherically stratified, τ_o may be expressed as

$$\tau_o = \int_0^{\infty} \kappa_\alpha(z)\,dz\sec\theta\,. \tag{A-6}$$

The quantity τ_h represents the one-way opacity from ground to height h equal to

$$\tau_h = \int_0^{h} \kappa_\alpha(z)\,dz\sec\theta\,. \tag{A-7}$$

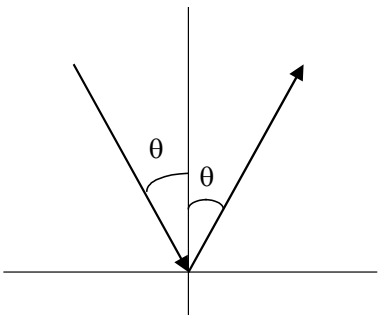

Figure A.2 Definition of incidence angle θ.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The variable R_S in Figure A.1 is the Fresnel reflection coefficient at the atmosphere-ground interface. The square of its magnitude is the reflectivity, which is expressed as

$$|R_S|^2 = 1 - \varepsilon_S, \tag{A-8}$$

where ε_S is the emissivity of the Earth's surface in the field of view of the sensor. The emissivity is a function of the operating frequency, polarization, and incidence angle of the sensor. Perfect emitters of energy, i.e., blackbody radiators, have an emissivity of one. Perfect conductors, such as shiny metal objects, have an emissivity of zero. Most objects are graybodies and have emissivities inbetween these limits.

Total energy detected by the radiometer is described by the equation in Figure A.1 shown entering the radiometer antenna. The cosmic and galactic brightness temperatures produce the first term. Both are attenuated by τ_o and, along with the total downward atmospheric emission, are reflected from the ground upward toward the radiometer. The brightness temperature T_B emitted by the ground (or a target if one is present within the footprint of the radiometer) produces the second term. It is equal to the product $\varepsilon_S T_S$, where T_S is the absolute temperature of the ground surface. The opacity of the intervening atmosphere τ_h between the ground and the radiometer at height h attenuates these two energy sources before they reach the radiometer. The third component of the total detected energy is produced by upward emission $T_U(h)$ due to atmospheric absorption phenomena that exist between the ground and height h.

Therefore, the total energy E received by the radiometer at a height h above the ground surface is found by adding the above energy sources as

$$E = [|R_S|^2 (T_C e^{-\tau_o} + T_D) + \varepsilon_S T_S] e^{-\tau_h} + T_U(h). \tag{A-9}$$

Two simplifications to the general radiative transfer equation of (A-9) can be made when the radiometer is deployed at low altitudes. First the cosmic, galactic, and downwelling atmospheric emission terms can be combined into one term called the sky radiometric temperature denoted by T_{sky}. The temperature T_{sky} is still a function of atmospheric water content, cloud cover, and radiometer operating frequency. A summary of the downwelling atmospheric temperature and atmospheric attenuation is given in Table A.2 for a zenith-looking radiometer under clear air and 11 types of cloud conditions at S-band, X-band, and Ka-band frequencies.[3] As discussed in Chapter 2, both the downwelling atmospheric temperature and atmospheric attenuation increase with increasing frequency and increasing water content of the clouds. The cosmic and galactic temperatures are not included in the atmospheric temperature shown in the table. The second simplification that occurs when a radiometer is deployed at low altitudes (e.g., as

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

part of a suite of sensors in a surface-to-surface missile) is made possible by neglecting the small upwelling atmospheric contribution.

Table A.2 Downwelling atmospheric temperature T_D and atmospheric attenuation A for a zenith-looking radiometer. (S.D. Slobin, *Microwave Noise Temperature and Attenuation of Clouds at Frequencies Below 50 GHz*, JPL Publication 81-46, Jet Propulsion Laboratory, Pasadena, CA [July 1, 1981].)

Case	Lower Cloud				Upper Cloud					S-Band (2.3 GHz) Zenith		X-Band (8.5 GHz) Zenith		Ka-Band (32 GHz) Zenith	
	Density g/m^3	Base km	Top km	Thickness km	Density g/m^3	Base km	Top km	Thickness km	Remarks	T_D (K)	A (dB)	T_D (K)	A (dB)	T_D (K)	A (dB)
1	–	–	–	–	–	–	–	–	Clear Air	2.15	0.035	2.78	0.045	14.29	0.228
2	0.2	1.0	1.2	0.2	–	–	–	–	Light, Thin Clouds	2.16	0.036	2.90	0.047	15.92	0.255
3	–	–	–	–	0.2	3.0	3.2	0.2		2.16	0.036	2.94	0.048	16.51	0.266
4	0.5	1.0	1.5	0.5	–	–	–	–		2.20	0.036	3.55	0.057	24.56	0.397
5	–	–	–	–	0.5	3.0	3.5	0.5		2.22	0.037	3.83	0.062	28.14	0.468
6	0.5	1.0	2.0	1.0	–	–	–	–	Medium Clouds	2.27	0.037	4.38	0.070	35.22	0.581
7	–	–	–	–	0.5	3.0	4.0	1.0		2.31	0.038	4.96	0.081	42.25	0.731
8	0.5	1.0	2.0	1.0	0.5	3.0	4.0	1.0		2.43	0.040	6.55	0.105	61.00	1.083
9	0.7	1.0	2.0	1.0	0.7	3.0	4.0	1.0		2.54	0.042	8.04	0.130	77.16	1.425
10	1.0	1.0	2.0	1.0	1.0	3.0	4.0	1.0	Heavy Clouds	2.70	0.044	10.27	0.166	99.05	1.939
11	1.0	1.0	2.5	1.5	1.0	3.5	5.0	1.5		3.06	0.050	14.89	0.245	137.50	3.060
12	1.0	1.0	3.0	2.0	1.0	4.0	6.0	2.0	Very Heavy Clouds	3.47	0.057	20.20	0.340	171.38	4.407

Cases 2-12 are clear air and clouds combined.

Antenna located at sea level, heights are measured from ground level.

Cosmic and galactic brightness temperatures and ground contributions are not included in the downwelling temperature T_D.

Attenuation A is measured along a vertical path from ground to 30-km altitude.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

References

1. F.T. Ulaby, R.K. Moore, and A.K. Fung, *Microwave Remote Sensing: Active and Passive, Vol. I, Microwave Remote Sensing Fundamentals and Radiometry*, Artech House, Norwood, MA (1981).
2. G.M. Hidy, W.F. Hall, W.N. Hardy, W.W. Ho, A.C. Jones, A.W. Love, M.J. Van Melle, H.H. Wang, and A.E. Wheeler, *Development of a Satellite Microwave Radiometer to Sense the Surface Temperature of the World Oceans*, NASA Contractor Rpt. CR-1960, National Aeronautics and Space Administration, Washington, DC (February 1972).
3. S.D. Slobin, *Microwave Noise Temperature and Attenuation of Clouds at Frequencies Below 50 GHz*, JPL Publication 81-46, Jet Propulsion Laboratory, Pasadena, CA (July 1, 1981).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix B

Bayesian Inference Applied to Freeway Incident Detection

B.1 Problem statement and development

Incident detection may be enhanced by fusing data from more than one information source if each produces a signature or data generated by independent phenomena, that is, the information sources are uncorrelated. Suppose a scenario exists where traffic flow data and incident reports are available from roadway sensors, cellular telephone calls from travelers, and radio reports from commercial truck drivers.[1] Furthermore, suppose that the roadway sensor spacing, elapsed time from the start of road sensor data transmission, or false alarm history is not adequate to detect or confirm an incident with a sufficiently high probability (>80 percent) in a timely manner. The cellular calls are known to contain inaccurate incident location data and the radio reports are too infrequent to confirm the incident by themselves.

Using historical data, traffic management personnel serving the affected area have constructed *a priori* probabilities for the likelihood that roadway sensor data are reporting a true incident based on the length of time lane occupancy (i.e., percent of selected time interval that vehicles are detected in the detection area of a sensor) and traffic volume are above preset thresholds and speed is below some other threshold. *A priori* probabilities also are assumed available to describe the accuracy of the cellular telephone and radio incident reports as a function of the number of calls and the variance of the reported incident locations.

We wish to apply Bayesian inference to compute the *a posteriori* probabilities that the detected event belongs to one of three types:

H_1 = one or more vehicles on right shoulder of highway,

H_2 = traffic in right-most lane slower than normal,

H_3 = traffic is flowing normally in all lanes.

The Bayesian approach to data fusion is based on updating probabilities as illustrated in the influence diagram shown in Figure B.1. The probability of the road sensors (*RS*) reporting data conditioned on event type *j* is given by

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$P_{RS}(\text{data}|\text{Event}_j) = P_{RS}(\text{data}|H_1)\,P(H_1|\text{Event}_j) + P_{RS}(\text{data}|H_2)\,P(H_2|\text{Event}_j) + P_{RS}(\text{data}|H_3)\,P(H_3|\text{Event}_j), \quad \text{(B-1)}$$

the probability of the cellular telephone (*CT*) calls reporting data conditioned on event type *j* is given by

$$P_{CT}(\text{data}|\text{Event}_j) = P_{CT}(\text{data}|H_1)\,P(H_1|\text{Event}_j) + P_{CT}(\text{data}|H_2)\,P(H_2|\text{Event}_j) + P_{CT}(\text{data}|H_3)\,P(H_3|\text{Event}_j), \quad \text{(B-2)}$$

and the probability of the radio (*RA*) reporting data conditioned on event type *j* is given by

$$P_{RA}(\text{data}|\text{Event}_j) = P_{RA}(\text{data}|H_1)\,P(H_1|\text{Event}_j) + P_{RA}(\text{data}|H_2)\,P(H_2|\text{Event}_j) + P_{RA}(\text{data}|H_3)\,P(H_3|\text{Event}_j), \quad \text{(B-3)}$$

where Event_j is one of the three events H_1, H_2, H_3. The set of arrows from "Information source report" to "Type ID given report" in Figure B.1 represents the probability calculations defined by Eqs. (B-1) through (B-3).

The values of the likelihood functions for the roadway sensors, $P_{RS}(\text{data}|H_1)$, $P_{RS}(\text{data}|H_2)$, and $P_{RS}(\text{data}|H_3)$; cellular telephone, $P_{CT}(\text{data}|H_1)$, $P_{CT}(\text{data}|H_2)$, and $P_{CT}(\text{data}|H_3)$; and the radio, $P_{RA}(\text{data}|H_1)$, $P_{RA}(\text{data}|H_2)$, and $P_{RA}(\text{data}|H_3)$ are found through *a priori* measurements and data collection and analysis activities.

Road sensor lane occupancy, traffic volume, and speed data are compared with predetermined or real-time calculated thresholds, depending on the incident detection algorithm, distance between sensors, and the data-reporting interval

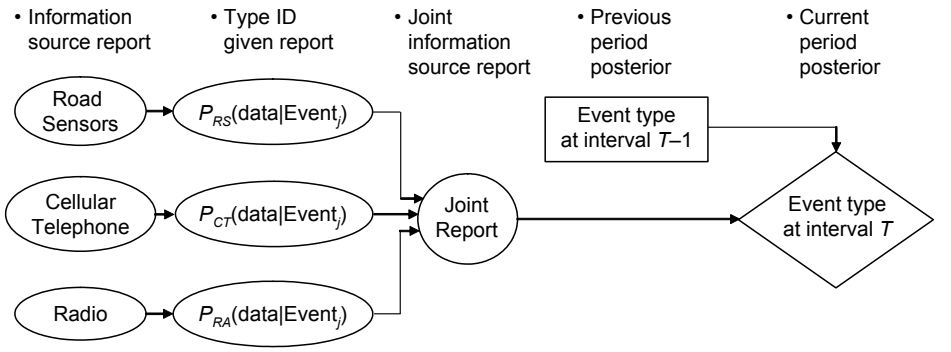

Figure B.1 Influence diagram for freeway event detection using data from three uncorrelated information sources.

characteristic of declaring an event of type *j*. Thus, offline analysis of the values and duration of real-time data determines the value of the likelihood function that expresses the probability that the data represent hypothesis H_i.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

The cellular telephone data are the number of calls that report the same event and the variance of the reported event location. The value of the likelihood function $P_{CT}(\text{data}|\text{Event}_j)$ is equal to the probability of receiving a predetermined number of calls with a predetermined event location variance, given the event is of type j. These probabilities are found from historical data collected as a function of event type, number of lanes affected, road configuration, traffic volume, weather, time-of-day, day-of-week, season, lighting, etc. Similar data are used to define the likelihood functions for the radio reports.

Quantitative values for the *a priori* probabilities $P(H_1|\text{Event}_j)$, $P(H_2|\text{Event}_j)$, and $P(H_3|\text{Event}_j)$ are determined from offline analysis of the types and numbers of events in the monitored area.

Next, the joint information source report shown in Figure B.1 is computed for a given time interval as the product of Eqs. (B-1) through (B-3), since the information sources are presumed to generate data from independent phenomena. Thus, the joint information source report is

$$P(\text{data}|\text{Event}_j) = \prod_k P_k(\text{data}|\text{Event}_j), \tag{B-4}$$

where k is the information source index, here equal to 1, 2 and 3 for road sensor, cellular telephone, and radio, respectively.

Finally, Bayes' rule is applied to calculate the current period *a posteriori* probability $P(\text{Event}_j|\text{data})$ that the detected event is of type j based on the values of the posterior probabilities evaluated during the previous period and the joint information source report. Accordingly,

$$P(\text{Event}_j \mid \text{data}) = \frac{P(\text{data} \mid \text{Event}_j)\, P(\text{Event}_j)}{P(\text{data})}, \tag{B-5}$$

where

$P(\text{data}|\text{Event}_j)$ = value from Eq. (B-4),

$P(\text{Event}_j)$ = value of $P(\text{Event}_j|\text{data})$ from previous period, and

$$P(\text{data}) = \sum_j P(\text{data} \mid \text{Event}_j)\, P(\text{Event}_j) \tag{B-6}$$

is the preposterior or probability of observing the data collected during the previous period given that events denoted by Event_j are present. Larger values of $P(\text{data})$ imply that the previous period values are more predictive of the situation as it evolves, i.e., the change in $P(\text{Event}_j|\text{data})$ from previous to current period is

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

smaller. When the information sources do not report an event type for the current time interval, updating is not performed and the values of $P(\text{Event}_j|\text{data})$ for the current interval are set equal to those from the previous period.

If the primary task is to detect abnormal traffic flow or an incident, the first and second terms in Eqs. (B-1) through (B-3) can be combined into a single declaration INCIDENT. The problem is further simplified by choosing Event_1 = INCIDENT and $\text{Event}_2 = \overline{\text{INCIDENT}}$, where the bar denotes negation. Therefore, the required probabilities are only dependent on $P(\text{data}|\text{Event}_1)$ since

$$P(\text{data}|\text{Event}_2) = 1 - P(\text{data}|\text{Event}_1) \tag{B-7}$$

and

$$P(\text{Event}_2) = 1 - P(\text{Event}_1). \tag{B-8}$$

Returning to the three-hypothesis problem, an initial value for $P(\text{Event}_j)$ and lower and upper bounds inside the interval (0, 1) for admissible values of $P(\text{Event}_j|\text{data})$ are needed to evaluate Eq. (B-5). When information concerning the initial values of $P(H_1)$, $P(H_2)$, and $P(H_3)$ is lacking, the initial values are set equal to one another with the value of 1/3 (i.e., the insufficient reason principle is applied). The boundaries for $P(H_1|\text{data})$, $P(H_2|\text{data})$, and $P(H_3|\text{data})$ are limited to (0.01, 0.99) to prevent the process that computes the *a posteriori* probability from terminating prematurely at the limiting endpoint values of 0 and 1.

The posterior probabilities for events H_1, H_2, and H_3 are computed from the posterior probabilities for event type and the scenario defined values for $P(H_1|\text{Event}_j)$, etc., as

$$P(H_1 \mid \text{data}) = \sum_j [P(\text{Event}_j \mid \text{data})\ P(H_1 \mid \text{Event}_j), \tag{B-9}$$

$$P(H_2 \mid \text{data}) = \sum_j [P(\text{Event}_j \mid \text{data})\ P(H_2 \mid \text{Event}_j), \tag{B-10}$$

and

$$P(H_3 \mid \text{data}) = \sum_j [P(\text{Event}_j \mid \text{data})\ P(H_3 \mid \text{Event}_j). \tag{B-11}$$

Thus, the probability of determining whether an incident has occurred is the sum of individual probabilities that are dependent on the identification of various features. The term $P(H_1|\text{Event}_j)$ expresses the *a priori* probability of finding event H_1 conditioned on event type j being present. Similar interpretations for $P(H_2|\text{Event}_j)$ and $P(H_3|\text{Event}_j)$ apply. Practical applications require the data fusion algorithms to be robust in their ability to identify the obvious events and those that are unexpected. It is also beneficial to have information sources at your

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

disposal that can assist in the detection and identification of more than one type of event.

B.2 Numerical example

Assume the likelihood functions $P(\text{data}/H_i)$ are given by the matrix in Table B.1.

Table B.1 Likelihood functions for road sensors, cellular telephones, and radios for the three-hypothesis freeway incident detection problem.

	E^{RS}: Detection of event by road sensors	E^{CT}: Detection of event by cellular telephone	E^{RA}: Detection of event by radio
H_1	0.02	0.70	0.65
H_2	0.70	0.08	0.15
H_3	0.95	0.01	0.02

Assume the prior probabilities are known and given by

$$P(H_i) = (0.5, 0.3, 0.2). \tag{B-12}$$

Then according to Eqs. (5-37) and (5-38),

$$\begin{aligned}\Lambda = \lambda^1\,\lambda^2\,\lambda^3 &= (0.02, 0.70, 0.95) \bullet (0.70, 0.08, 0.01) \bullet (0.65, 0.15, 0.02)\\ &= (0.0091, 0.00084, 0.00152)\end{aligned} \tag{B-13}$$

The posterior probability becomes [from Eq. (5-39)]

$$\begin{aligned}P(H_i/E^{RS}, E^{CT}, E^{RA}) = \alpha\, P(H_i)\, \Lambda_i = \alpha\, (0.5, 0.3, 0.2) \bullet (0.0091, 0.00084,\\ 0.00152) = \alpha\, (0.00455, 0.00024, 0.0003)\end{aligned} \tag{B-14}$$

where $\alpha = 196.4637$.

Thus,

$$P(H_i/E^{RS}, E^{CT}, E^{RA}) = (0.8939, 0.0472, 0.0589). \tag{B-15}$$

The output of the data fusion process is to declare H_1 the most likely hypothesis, namely one or more vehicles are on the right shoulder of the roadway.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

REFERENCES

1. L.A. Klein, *Sensor Technologies and Data Requirements for ITS*, Artech House, Norwood, MA (June 2001).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix C

Plausible and Paradoxical Reasoning

C.1 Example of unexpected results arising from conflicting information sources

The following example is attributed to Lotfi Zadeh.[1] Suppose two doctors examine a patient and agree the patient suffers from either meningitis (M), concussion (C), or brain tumor (T). The frame of discernment for these propositions is given by

$$\Theta = \{M, C, T\}. \tag{C-1}$$

Assume the doctors agree on their low expectation of a tumor, but disagree as to the other likely cause and provide diagnoses as follows:

$$m_1(M) = 0.99 \qquad m_1(T) = 0.01 \tag{C-2}$$

$$m_2(C) = 0.99 \qquad m_2(T) = 0.01, \tag{C-3}$$

where the subscript 1 indicates the diagnosis of the first doctor and the subscript 2 the diagnosis of the second doctor.

The belief functions can be combined by using Dempster's rule to calculate the orthogonal sum as shown in Table C.1.

Table C.1 Orthogonal sum calculation for conflicting medical diagnosis example (step 1).

$m_1(M) = 0.99$	$m(\phi) = 0.9801$	$m(\phi) = 0.0099$
$m_1(T) = 0.01$	$m(\phi) = 0.0099$	$m(T) = 0.0001$
	$m_2(C) = 0.99$	$m_2(T) = 0.01$

The normalization factor K equal to

$$K = \frac{1}{1 - 0.9801 - 0.0099 - 0.0099} = 10{,}000 \tag{C-4}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

reassigns the probability mass of the empty set matrix elements to the nonempty set element (2, 3) as shown in Table C.2.

Thus, application of Dempster's rules gives the unexpected result that

$$m(T) = 1, \tag{C-5}$$

which arises from the bodies of evidence (the doctors) agreeing that patient does not suffer from a tumor, but being in almost full contradiction about the other causes of the disease.

Table C.2 Normalization of nonempty set matrix element for conflicting medical diagnosis example (step 2).

$m_1(M) = 0.99$	$m(\phi) = 0$	$m(\phi) = 0$
$m_1(T) = 0.01$	$m(\phi) = 0$	$m(T) = 1$
	$m_2(C) = 0.99$	$m_2(T) = 0.01$

Such an example provides a negative implication for the use of Dempster-Shafer in automated reasoning processes when a large amount of conflict can potentially exist in the information sources. Therefore, in most practical applications of Dempster-Shafer theory, some ad-hoc or heuristic approach must be added to the fusion process to correctly account for the possibility of a large degree of conflict between the information sources.

C.2 Plausible and paradoxical reasoning

A solution to this dilemma has been proposed by Dezert, whereby he modifies the Dempster-Shafer requirements that bodies of evidence be independent (i.e., each information source does not take into account the knowledge of the other sources) and provide a belief function based on the power set 2^{Θ}, which is defined as the set of all proper subsets of Θ when all elements θ_i, $i = 1, n$ are disjoint.[2] His formulation allows admission of evidence from the conjunction (AND) operator $\cap$ as well as the disjunction (OR) operator $\cup$. The broadened permissible types of evidence form a hyper-power set D^{Θ} as the set of composite possibilities built from Θ with $\cup$ and $\cap$ operators $\forall A \in D^{\Theta}$, $B \in D^{\Theta}$, $(A \cup B) \in D^{\Theta}$, and $(A \cap B) \in D^{\Theta}$.

Plausible and paradoxical reasoning may be viewed as an extension of probability theory and Dempster-Shafer theory. For example, let $\Theta = \{\theta_1, \theta_2\}$ be the simplest frame of discernment involving only two elementary hypotheses with no additional assumptions on θ_1, θ_2. Probability theory deals with basic probability assignments $m(\bullet) \in [0, 1]$ such that

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

$$m(\theta_1) + m(\theta_2) = 1 \tag{C-6}$$

Dempster-Shafer theory extends probability theory by dealing with basic belief assignments $m(\bullet) \in [0, 1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1 \tag{C-7}$$

Plausible and paradoxical theory extends the two previous theories by accepting the possibility of paradoxical information and deals with new basic belief assignments $m(\bullet) \in [0, 1]$ such that

$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) + m(\theta_1 \cap \theta_2) = 1 \tag{C-8}$$

To explore how plausible and paradoxical theory functions, consider the paradoxical information basic probability assignments for $\Theta = \{\theta_1, \theta_2\}$ from two information sources given by

$$m_1(\theta_1) = 0.80 \quad m_1(\theta_2) = 0.15 \quad m_1(\theta_1 \cup \theta_2) = 0 \quad m_1(\theta_1 \cap \theta_2) = 0.05 \tag{C-9}$$

$$m_2(\theta_1) = 0.90 \quad m_2(\theta_2) = 0.05 \quad m_2(\theta_1 \cup \theta_2) = 0 \quad m_2(\theta_1 \cap \theta_2) = 0.05 \tag{C-10}$$

As shown in Table C.3,

$$m(\theta_1) = 0.72 \quad m(\theta_2) = 0.0075 \quad m(\theta_1 \cup \theta_2) = 0 \quad m(\theta_1 \cap \theta_2) = 0.2725, \tag{C-11}$$

where the result for $m(\theta_1 \cap \theta_2)$ is calculated as the sum of the matrix elements marked with an asterisk. Accordingly,

$$\begin{aligned} m(\theta_1 \cap \theta_2) &= 0.135 + 0.045 + 0.04 + 0.0025 + 0.04 + 0.0075 + 0.0025 \\ &= 0.2725 \end{aligned} \tag{C-12}$$

Table C.3 Two information source, two hypothesis application of plausible and paradoxical theory.

$m_1(\theta_1)$ = 0.80	$m(\theta_1)$=0.72	$m_1(\theta_1 \cap \theta_2)$ =0.04*	$m(\theta_1)$=0	$M[\theta_1 \cap (\theta_1 \cap \theta_2)]$ =0.04*
$m_1(\theta_2)$ = 0.15	$m(\theta_1 \cap \theta_2)$ =0.135*	$m(\theta_2)$=0.0075	$m[(\theta_2) \cap (\theta_1 \cup \theta_2)]$ =0	$m[\theta_2 \cap (\theta_1 \cap \theta_2)]$ =0.0075*
$m_1(\theta_1 \cup \theta_2)$ = 0	$m(\theta_1)$=0	$m[(\theta_1 \cup \theta_2) \cap \theta_2]$ =0	$m(\theta_1 \cup \theta_2)$=0	$m[(\theta_1 \cup \theta_2)(\theta_1 \cap \theta_2)]$ =0
$m_1(\theta_1 \cap \theta_2)$ = 0.05	$m[(\theta_1 \cap \theta_2) \cap \theta_1]$ =0.045*	$m[(\theta_1 \cap \theta_2) \cap \theta_2]$ =0.0025*	$m[(\theta_1 \cap \theta_2)(\theta_1 \cup \theta_2)]$ =0	$m[(\theta_1 \cap \theta_2)(\theta_1 \cap \theta_2)]$ =0.0025*
	$m_2(\theta_1)$ = 0.90	$m_2(\theta_2)$ = 0.05	$m_2(\theta_1 \cup \theta_2)$ = 0	$m_2(\theta_1 \cap \theta_2)$ = 0.05

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

C.3 Resolution of medical diagnosis dilemma

Returning to the medical diagnosis problem and applying plausible and paradoxical theory to the diagnoses given by Eqs. (C-2) and (C-3) gives

$$m(M \cap C) = 0.9801 \quad m(M \cap T) = 0.0099 \quad m(T \cap C) = 0.0099$$
$$m(T) = 0.0001 \tag{C-13}$$

as shown by the entries in Table C.4.

Table C.4 Resolution of medical diagnosis example through plausible and paradoxical reasoning.

$m_1(M) = 0.99$	$m(M \cap C) = 0.9801$	$m(M \cap T) = 0.0099$
$m_1(T) = 0.01$	$m(T \cap C) = 0.0099$	$m(T) = 0.0001$
	$m_2(C) = 0.99$	$m_2(T) = 0.01$

The belief assignments become

$$bel(M) = m(M \cap C) + m(M \cap T) = 0.9801 + 0.0099 = 0.99 \tag{C-14}$$

$$bel(C) = m(M \cap C) + m(T \cap C) = 0.9801 + 0.0099 = 0.99 \tag{C-15}$$

$$bel(T) = m(T) + m(M \cap T) + m(T \cap C) = 0.0001 + 0.0099 + 0.0099 = 0.0199 \tag{C-16}$$

If both doctors can be considered equally reliable, the combined information granule $m(\bullet)$ focuses the weight of evidence on the paradoxical proposition $M \cap C$, which means the patient suffers from both meningitis and concussion, but almost assuredly not from a brain tumor. This conclusion is one common sense would support and rules out an evasive surgical procedure to remove a nonexistent tumor. Further medical evaluation is called for before treatment for meningitis or concussion is administered.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

REFERENCES

1. L.A. Zadeh, *On the Validity of Dempster's Rule of Combination of Evidence*, Memo M 79/24, Univ. of California, Berkeley (1979).
2. J. Dezert, "Foundations for a new theory of plausible and paradoxical reasoning," *Information and Security*, Vol. 9, 1-45 (2002).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix D

Summary of Results from Alternative Dempster-Shafer Models Applied to the Godfather-Mr. Jones Saga

This appendix summarizes the results from the methods used to identify the assassin of Mr. Jones based on coin toss only evidence and on coin toss plus Peter's alibi evidence.

D.1 Coin toss (E_1) only evidence

Table D.1 shows the results of applying evidence set E_1 to the transferable belief, Bayes, or plausibility inference model.

Table D.1 Probability summary using evidence set E_1 only.

Assassin Set	TBM*	Bayes	Plausibility
P_1({Mary})	0.5	0.5	1/3
P_1({Peter, Paul})	0.5	—	—
P_1({Peter})	—	0.25	1/3
P_1({Paul})	—	0.25	1/3

* TBM = transferable belief model

D.2 Coin toss (E_1) and Peter's alibi (E_2) evidence

The entries in columns 2–4 of Table D.2 reflect the use of evidence sets E_1 and E_2 in the same inference model, either transferable belief, Bayes, or plausibility, as indicated by subscripts 1 and 2 after the model designation. In columns 5 and 6, subscript 1 indicates that E_1 is input to the transferable belief or plausibility model, respectively, while subscript 2 indicates that E_2 is input and processed with a Bayesian probability model.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Table D.2 Probability summary using evidence sets E_1 and E_2.

Assassin Set	$TBM_{1,2}$	$Bayes_{1,2}$	$Plausibility_{1,2}$	TBM_1Bayes_2	Pl_1Bayes_2
$P_1(\{Mary\})$	0.5	2/3	0.5	2/3	0.5
$P_1(\{Paul\})$	0.5	1/3	0.5	1/3	0.5

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix E

Training-Set Size for Valid Generalization

When the fraction of errors made on the training set is less than ε/2, where ε is the fraction of errors permitted on the test of the network, the number of training examples N_p is

$$N_p \geq (32N_w/\varepsilon) \ln(32M/\varepsilon), \tag{E-1}$$

where

N_w = number of synaptic weights in the network and

M = total number of hidden computation nodes.

This is a worst case formula for estimating training set size for a single layer neural network that is sufficient for good generalization.[1] On average, a smaller number of training samples will suffice, such as

$$N_p > N_w/\varepsilon. \tag{E-2}$$

Thus, for an error of 10 percent, the number of training examples is approximately 10 times the number of synaptic weights in network (N_w).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

REFERENCES

1. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Chap. 6, Macmillan College Publishing Company, NY (1994).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix F

Voting Fusion with Nested Confidence Levels

The key to deriving Eq. (8-6) or (8-7) in Chapter 8 is the creation of nonnested confidence levels for each sensor as was illustrated in Figure 8.3. Nonnested confidence levels allow a unique value to be selected for the inherent sensor detection probability when different signal-to-interference ratios are postulated and implemented at each confidence level. In fact, the ability to specify and then implement unique detection probabilities for each confidence level is one of the considerations that make this voting fusion technique practical.

Alternatively, a Venn diagram such as the one in Figure F.1 with nested confidence levels implies that the detection probabilities at each confidence level are not independent. Here, the confidence levels A_1, A_2, and A_3 of Sensor A and the confidence levels in the other sensors are not independent of each other. Confidence level A_3 is a subset of level A_2, which is a subset of level A_1. Discriminants other than signal-to-interference ratio are used in this case to differentiate among the confidence levels. For example, target-like features that are present in the signal can be exploited by algorithms to increase the confidence that the signal belongs to a bonafide target. This model is more restrictive and may not depict the way the sensors are actually operating in a particular application.

A different Boolean expression is also needed to compute the detection probability of the three-sensor suite when nested confidence levels are postulated. Since the confidence levels for each sensor are not independent, the simplifying assumptions of Eqs. (8-4) and (8-5) no longer apply. The Boolean

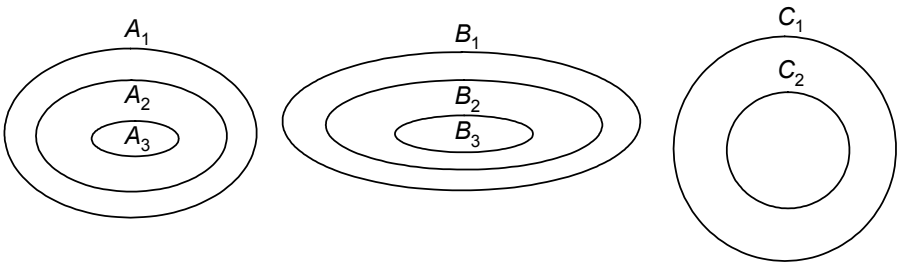

Figure F.1 Nested sensor confidence levels.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

equation for the sensor system detection probability with nested confidence levels and the detection modes defined in Table 8.1 takes the form

$$\text{System } P_d = P_d\{A_1\ B_1\ C_1 \text{ or } A_2\ C_2 \text{ or } B_2\ C_2 \text{ or } A_3\ B_3\} \tag{F-1}$$

or

$$\begin{aligned}\text{System } P_d = {} & P_d\{A_1\ B_1\ C_1\} + P_d\{A_2\ C_2\} + P_d\{B_2\ C_2\} + P_d\{A_3\ B_3\} \\ & - P_d\{A_2\ B_1\ C_2\} - P_d\{A_1\ B_2\ C_2\} - P_d\{A_3\ B_3\ C_1\}.\end{aligned} \tag{F-2}$$

If the sensors respond to independent signature-generation phenomena such that the likelihood of detection by one sensor is independent of that of another, then

$$\begin{aligned}\text{System } P_d = {} & P_d\{A_1\}\ P_d\{B_1\}\ P_d\{C_1\} + P_d\{A_2\}\ P_d\{C_2\} + P_d\{B_2\}\ P_d\{C_2\} \\ & + P_d\{A_3\}\ P_d\{B_3\} - P_d\{A_2\}\ P_d\{B_1\}\ P_d\{C_2\} \\ & - P_d\{A_1\}\ P_d\{B_2\}\ P_d\{C_2\} - P_d\{A_3\}\ P_d\{B_3\}\ P_d\{C_1\}.\end{aligned} \tag{F-3}$$

The difference terms represent areas of overlap that are accounted for more than once in the sum terms.

The false alarm probability of the three-sensor system is also in the form of (F-3) with P_d replaced by P_{fa}. Thus, with nested confidence levels, the system false alarm probability is

$$\begin{aligned}\text{System } P_{fa} = {} & P_{fa}\{A_1\}\ P_{fa}\{B_1\}\ P_{fa}\{C_1\} + P_{fa}\{A_2\}\ P_{fa}\{C_2\} + P_{fa}\{B_2\}\ P_{fa}\{C_2\} \\ & + P_{fa}\{A_3\}\ P_{fa}\{B_3\} - P_{fa}\{A_2\}\ P_{fa}\{B_1\}\ P_{fa}\{C_2\} \\ & - P_{fa}\{A_1\}\ P_{fa}\{B_2\}\ P_{fa}\{C_2\} - P_{fa}\{A_3\}\ P_{fa}\{B_3\}\ P_{fa}\{C_1\}.\end{aligned} \tag{F-4}$$

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Appendix G

Fuzzy Set Boundaries, Fuzzy Set Widths, and Defuzzification Methods

G.1 Fuzzy-set boundaries

Fuzzy logic relaxes constraints imposed by conventional set theory and thereby enables it to better model vagueness or imprecision that occurs in many real-world processes. This concept is illustrated in Figure G.1, which shows short, medium, and tall sets as depicted by conventional and fuzzy set theory. In conventional set theory, the set boundaries for each member of the set are precise, whereas in fuzzy logic they are defined by a membership function. A particular quantity of the variable, in this case height, has membership in a fuzzy set between the limits of 0 and 1. For example, if the height of a person or object is 4½ feet, this particular quantity would have partial membership in both the short and medium fuzzy sets with values determined by where a vertical line drawn through 4½ feet on the height axis (i.e., the abscissa) intersects the corresponding membership functions.

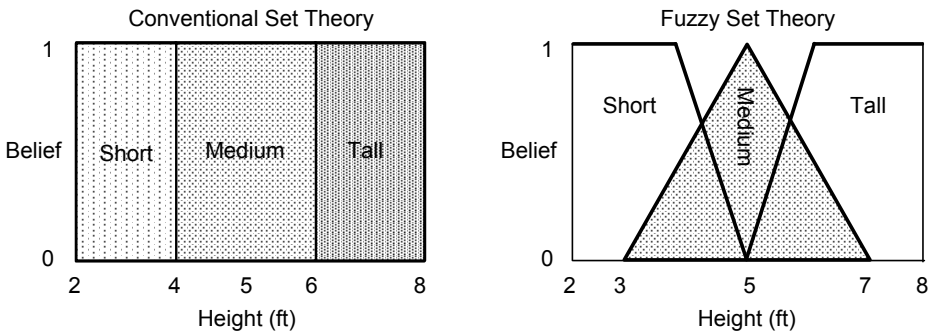

Figure G.1 Short, medium, and tall sets as depicted in conventional and fuzzy-set theory.

G.2 Effects of fuzzy-set widths on control

Figure G.2 shows the effect of varying the width of fuzzy sets on their overlap and, hence, the type of control that is achieved.[1] Small fuzzy set widths (0.2 and 1) that are completely separated produce bad control and do not converge on the set point. Large fuzzy set widths with too much overlap [8 (not shown) and 10]

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

produce satisfactory control, but overshoot is large. Large fuzzy set widths can require a larger number of fuzzy control rules and the convergence to a set point is slow. Fuzzy sets that are not isolated and do not have too much overlap (4 and 6) produce good control.

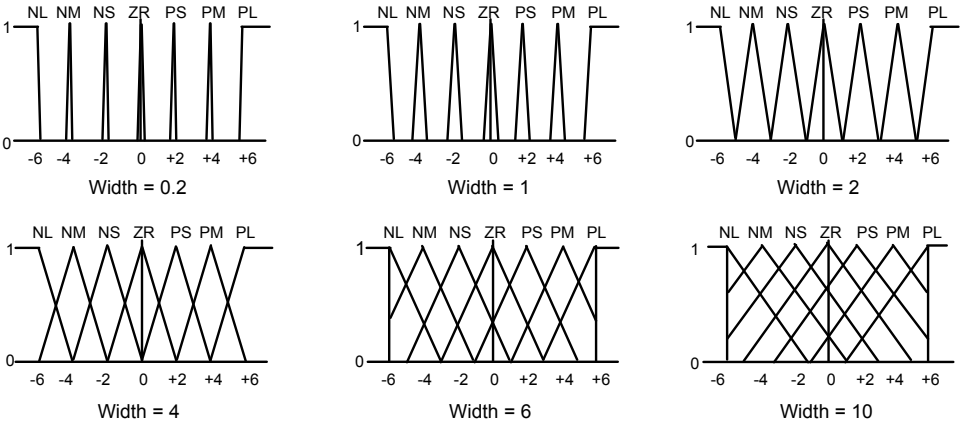

Figure G.2 Influence of fuzzy-set widths on fuzzy set overlap.

G.3 Defuzzification methods

Examples of defuzzification methods used to obtain a crisp output are illustrated in Figure G.3.[1]

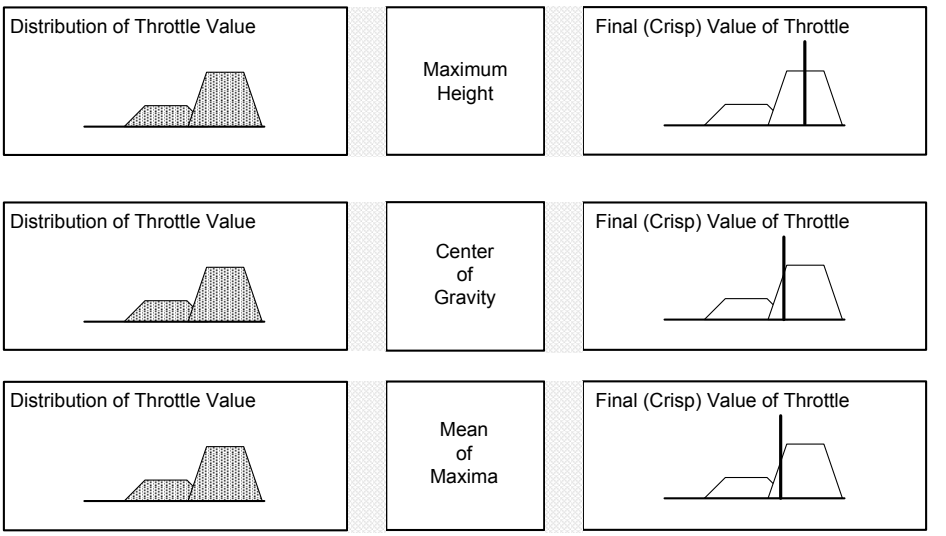

Figure G.3 Defuzzification methods.

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

REFERENCES

1. M. Mizumoto, "Fuzzy controls under various fuzzy reasoning methods," *Information Sciences*, (45), 129-151 (1988).

Downloaded From: https://www.spiedigitallibrary.org/ebooks/ on 18 Oct 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use