12TH ICCRTS
"Adapting C2 to the 21st Century"

Title of Paper: The Impact of Entropic Drag on Command and Control

Paper ID: I-020

Track(s): Network Centric Experimentation and Applications, C2 Concepts Theory and Policy, Organizational Issues

Authors: David Scheidt, Michael Pekala

Point of Contact: David Scheidt
Name of Organization: Johns Hopkins University Applied Physics Laboratory
Complete Address:
Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Road,
Laurel, MD 20723

Telephone: (240) 228-7301
E-mail: david.scheidt@jhuapl.edu

Abstract: A commonly believed axiom in signal detection theory is that "more information is good" [1]. That is, when attempting to determine the state of a partially observable system the addition of correct information monotonically improves the correctness of the state assessment. When diagnosing static systems the assertion that "the effect of information is to increase the likelihood of getting correct diagnosis, while reducing the likelihood of incorrect diagnosis" holds. However, when diagnosing dynamic systems, the improvement in diagnosis is offset by increases in uncertainty generated by the dynamic forces within the system being observed. A similar effect occurs when *controlling* dynamic, stochastic systems. The act of exercising control requires a finite amount of time, during which uncertainty enters into the system reducing the efficacy of the control policy. The impact of information processing delays increases in relevance as the complexity and pace of both the system and control apparatus increase. This paper defines a mathematical framework describing the effect of information and information processing on the diagnosis of dynamic systems.

Introduction

Regardless of the command and control structure, the quality and speed of information flow between team members directly impacts the performance of a team. In all structures, latencies are caused by each hop in the communications chain. Communication limitations have been a topic of growing popularity since Shannon [2]. The importance Shannon's information entropy has on command and control has been studied by Perry [3], Moffett [4] and Caves [5]. However, these studies are time insensitive and do not attempt to measure and account for the dynamic forces associated with complex, changing environments. In this paper we take a step toward filling that gap after initially laying the technical and conceptual foundations of abstract systems and information theory.

Systems and Worlds

An understanding of the information required to represent a system begins with a formal definition of that system. We refer to a system being diagnosed or controlled as a *world* (**W**). Worlds are defined in terms of elements, attributes, events, and state transitions, as follows.

Definition: A *world* **W** = (**K, E, A, Att**, **Val**, f), where

K = $\{k_1, \dots k_n\}$ is a finite set of *elements*

E = $\{e_1, \dots e_r\}$ is a set of *events*

A = $\{a_1, \dots a_p\}$ is a finite set of *attributes*, each applicable to one or more elements.

Att $\subseteq$ **K** $\times$ **A** $\cup$ $\mathbf{A}^2 \cup \cdots \cup \mathbf{A}^P$ is an element-attribute relation in which $(k_i, (a_{ki_1}, \dots, a_{ki_m})) \in Att$ if and only if $\forall i (1 \le i \le m) a_{k_{ij}}$ are those and only those attributes that "apply to" element k_i.

[Notation: $\mathbf{A}_k$ denotes the set of k's attributes; i.e., $\mathbf{A}_k = \{a_{ki_1}, \dots, a_{ki_m}\} \leftrightarrow (k_i, (a_{ki_1}, \dots, a_{ki_m}))$]

[Notation: $\forall a_{ki_j} \in \mathbf{A}_{k_i}$ $\mathbf{V}_{a_{ki_j}}$ is the range of values for attribute a_{ki_j}]

$\forall k_i \in \mathrm{K}$, $\mathrm{Val}_{k_i} \subseteq \mathbf{V}_{a_{ki_1}} \times \dots \times \mathbf{V}_{a_{ki_m}}$ is a valuation relation for k_i's attributes.

Val = $\bigcup_{k_i \in K} \mathrm{Val}_{k_i}$ is the collection of valuation relations for all elements in **W**.

f : **K** $\times$ **Val** $\times$ **E** $\rightarrow$ **Val** is a state transition function.

■

Elements represent real or simulated objects, e.g., baffles, thermostats, switches, and ducts—or their representations—in a real or simulated HVAC world. *Attributes* provide descriptions of elements, such as location and orientation. These are finite, constrained values which collectively define the scope of the world.

The *state* of an element is an assignment of values to all its attributes; i.e., the state space ($\mathbf{P}_k$) of element k is $\mathbf{Val}_k$.

$$\mathrm{P}_k = \mathrm{Val}_k \quad (1)$$

An individual state of k is represented by $\mathbf{x}_k$ or equivalently by the list $(x_{k_1}, \dots, x_{k_m})$ of values of k's attributes.

Elements change states via *events*. Laws that govern state transitions are mapped through the function f .

$\mathbf{P}_\mathrm{W}$ denotes the state space of **W** and is defined as the power set of the state spaces for all $k \in \mathbf{K}$, i.e.;

$$\mathrm{P_w} = \mathrm{P}_{k_1} \times \mathrm{P}_{k_2} \times \ldots \times \mathrm{P}_{k_n} \qquad (2)$$

Each element of $\mathbf{P}_W$ is called a *state of the world*, or *world state*, some of which are feasible ($\mathbf{P}_f$) and others are infeasible ($\mathbf{P}_i$). $\mathbf{P}_f$ and $\mathbf{P}_i$ are mutually exclusive and comprehensive:

$$\mathrm{P}_f \cap \mathrm{P}_i = \emptyset \qquad (3)$$

$$\mathrm{P}_w = \mathrm{P}_f \cup \mathrm{P}_i \qquad (4)$$

Information and Entropy

The information content (I) of message is the amount of information required to uniquely identify a state. At this point we only consider messages whose information content is ideally coded for the set of possible values that could be stored in the message. That is, the representation scheme uses the minimum number of bits to completely represent the set of possible values. Information content relative to **state space P** is defined by Brillouin [6] as:

$$\mathrm{I} = c \ln |P| \qquad (5)$$

Here, c is a constant associated with the size of the language used to represent a world state. In modern computers and communications this language is binary allowing us to define information content as:

$$\mathrm{I} = \log_2 |P| \qquad (6)$$

Through Brillouin the number of feasible states that exist for a world after message m describing that world has been received (and processed) is

$$\left| P_{f_{\mathrm{I}_m}} \right| = 2^{(\log_2 |P_w| - 2^{\mathrm{I}_m})} \qquad (7)$$

or, alternatively,

$$\left| P_{f_{\mathrm{I}_m}} \right| = \frac{|P_w|}{2^{\mathrm{I}_m}} \qquad (8)$$

The feasible space represents disorder, lack of knowledge, or *information entropy* (h) [7]. Shannon defines information entropy of a piece of information, a, as

$$h(a) = \log_2 \frac{1}{\Pr(a)} \qquad (9)$$

in which $\Pr(a)$ is the general probability of a. In a world of fixed size, the post-message entropy of that world is:

$$H = \log_2 |P_f| = \log_2 |P_w| - I_m \qquad (10)$$

Information entropy is a useful construct that we will extend to collections of information and information exchanges.

Information entropy and uncertainty are frequently used interchangeably. While both are used to measure disorder in a system we distinguish them by the units in which they are measured. Information entropy, being a form of Shannon information, is measured in bits. Uncertainty is measured in natural units for the system being described (e.g., meters).

Equation (10) shows how new information can be used to reduce entropy within the representation of a world. In static worlds the information content of a message that contains positive information ($I_m > 0$) decreases information entropy. For ideally coded messages, providing information about a previously unknown world (worlds in which $H = \log_2 |P_w|$) the decrease in information entropy is equal to the size of the message.

$$\Delta H = \begin{cases} -I_m; |P_w| \geq 2^{I_m} \\ \log_2 P_w; |P_w| < 2^{I_m} \end{cases} \qquad (11)$$

If some knowledge of the world exists prior to receipt of a message, a portion of the message may be redundant, reducing the message's information content. The information content of a message about a world for which a priori information is available is expressed as:

$$\Delta H = \log_2 | P_f \cap p(I_m) | - I_m \qquad (12)$$

where $p(x)$ is a decoding function that produces a set of *feasible* states from x.

Uncertainty Due to Fidelity

In practical terms worlds are represented as abstractions. In modern communications, semantics for encoding attributes (and elements, in this paper) are defined and encoded in bits. Some real-world information is naturally discrete. Some, though, is naturally continuous and can only be approximated with binary representations. The amount of uncertainty associated with a discrete representation of continuous attributes (ε_r) is limited (i.e., bounded from below) by the number of bits used to encode the data and the dimensionality of the data. For example, when communicating attributes describing positions in Cartesian space, a message must be transmitted digitally at some point, limiting the fidelity by the least precise representation used in the communications path, which is the unit distance ($\hat{r}$) of the representation. In this discussion we assume that a normalized representation scheme is used and all states are equally probable. MacKay [8] showed that a normalized representation scheme maximizes the information content per transmitted bit. Since we are primarily interested in expressing the limits of communication we can assume optimal, normalized encoding is used and one bit of transmitted data equals one bit of un-decayed Shannon information. The unit distance defines the minimum uncertainty of our knowledge such that the minimum amount of error is:

$$\varepsilon_r = \frac{\sqrt{n}}{2}\hat{r}; r \in Z^n \quad (13)$$

where r is the unit size of representation and n is the dimensionality of the space.

For a world of fixed size, the size of the state space is derived from the range of attributes within the system and the unit distance of each attribute. Specifically,

$$|P_k| = \prod_{x \in X} \left| \frac{x_i - x_0}{\hat{r}_x} \right| \quad (14)$$

where x_0 and x_i are the minimum and maximum values of an ordered attribute and $\hat{r}_a$ is the unit distance of the attribute. When attribute dimensions are equivalent (e.g. Cartesian coordinates), then

$$|P_k| = \left| \frac{x_i - x_0}{\hat{r}} \right|^n \quad (15)$$

and the size of the state space of the world is

$$|P_w| = \left| \frac{x_i - x_0}{\hat{r}} \right|^{|K| \cdot n} \quad (16)$$

The information content of a regular world (i.e. a world in which transitions between states are Markovian) can now be shown as a function of dimensions of the world and the unit distance:

$$I_w = |K| \cdot n\left(\log_2\left(x_i - x_0\right) - \log_2 \hat{r}\right) \quad (17)$$

Entropic Drag

So far we have limited our discussion to static worlds, in which information entropy is decreased by each successive message that describes the world. On the other hand, in dynamic worlds the decrease in entropy achieved by a message can be offset by an increase in entropy due to the world's dynamic forces. We call this *entropic drag* (Γ), as the increase in entropy creates a drag on an observer's ability to understand the world. Entropic drag is derived from the rate at which the dynamic forces create unpredictable change. To define these dynamic forces recall that a world's state space P_w is divided into feasible and infeasible states and that the transition between these states is Markovian, defined by the transition function f, defined above. For any set of infeasible states there exists a boundary layer between it and feasible states. The rate at which infeasible states transition to feasible states is the driving force behind entropic drag. We formally define entropic drag as the ***log₂*** of the rate at which previously infeasible world states become feasible:

$$\Gamma_w(t) = H_w(I)\frac{dI}{dt} = \log_2 |P_f|(x)\frac{dx}{dt} \quad (18)$$

The impact of entropic drag on information is shown as a decrease in the information content on a world after an observation such that:

$$I_w = (I_0 + I_m)\cdot\left(1 - \int_0^{t_m} \Gamma(t)dt\right) \quad (19)$$

where t_m is the time required to process the message, I_0 is the information content on the world prior to processing the message and I_m is the information content of the message. The forces of entropic drag continuously degrade information on a world whether messages continue to arrive or not. Only through the continual acquisition of new information can a state of knowledge of a dynamic world be maintained. If a commander's objective is to make a decision with maximal information he must be cognizant of the rate of entropy and the value of new information. An important feature of equation (19) is that, if the time required to process a message is sufficiently long, the loss of information associated with entropic drag will exceed the information content of the message and the net result of processing the message will be an information loss!

Obtaining and using information takes some finite amount of time. The process of obtaining and using information was described by Boyd as the Observe, Orient, Decide and Act (OODA) loop. The first part of this loop consists of a sensor observing the environment and transmitting a message to an information sink where it is fused with additional information to create a consistent model of the world. The delay (δ) between an observation (beginning) at time t_o and the incorporation of the observed data at t' is the sum of the sensor processing (δ_s), communications (δ_c) and data fusion (δ_f) delays.

$$t' = t_0 + \delta = t_0 + \delta_s + \delta_c + \delta_f \quad (20)$$

We assume there are no additional delays, for instance placing on hold a message between the completion of processing and the onset of communications.

The sensor processing and communications channels are assumed to process information at a fixed rate measured in bits per second (β). The data processing profile for data fusion varies by technique, with some techniques, such as nearest-time replacement, operating at a fixed rate, and others, such as search-based techniques, incurring exponential increases in latency as the number of observations increase. For the purpose of this paper we will assume that each element in the observe-orient chain operates at a fixed rate in bits per second.

$$\beta = \frac{I_m}{\delta} \quad (21a)$$

or

$$\delta = t_m = \frac{I_m}{\beta} \quad (21b)$$

By applying this definition to (19) we can express information content in terms of message size and bandwidth.

$$I_w = (I_0 + I_m)\cdot\left(1 - \int_0^{\frac{I_m}{\beta}} \Gamma(t)dt\right) \quad (22)$$

and

$$\Delta I_w = I_m - (I_0 + I_m)\int_0^{\frac{I_m}{\beta}} \Gamma(t)dt \quad (23)$$

From (18) we can see that information is only gained from a single message if the rate of information entropy is less than the bandwidth.

$$\Delta I_w > 0 \text{ iff } \beta > \Gamma \quad (24)$$

Information entropy progresses along a frontier between the feasible state space and the infeasible state space. The frontier (**S**) is set of piecewise smooth surfaces ($\mathbf{s}_i \in \mathbf{S}$) within P_w that are defined as the boundaries between infeasible states $x` \in \mathbf{P}_i$ that are reachable from a feasible state $x`` \in \mathbf{P}_f$ through an event in accordance to the transition function. Each piecewise smooth surface is either a member of a set of surfaces that form a closed n-dimensional surface that encapsulates positive or negative information, or a member of a semi-closed surface whose edges abut limits of $\mathbf{P}_w$ space. **S** changes at a rate that is defined by the entropic drag with the boundary perpetually growing the amount of feasible space and shrinking the infeasible space.

The entropic drag for a surface $\mathbf{s_i}$ is:

$$\Gamma_{s_i}(t) = H_{s_i}(t)\frac{dh}{dt} = \log_2 s_i \cdot \vec{v}\frac{dx}{dt} \quad (25)$$

where $\vec{v}$ is the n-dimensional vector of unpredictable change normalized to s_i. The entropic drag for the entire world is the sum of the entropic drag for each surface.

$$\Gamma_w(t) = \sum_{\forall s_i} H_{s_i}(t)\frac{dh}{dt} \quad (26)$$

Positive and Negative Information

Positive information (P^+) is a set of assertions that one or more states are true.

$$P^+ = \{x_1, x_2, .., x_n\} \quad (27)$$

Positive information usually represents highly improbable states, so messages containing positive information provide a large amount of information content. For example, a message stating that an ant was observed at a given location at time t_0 is positive information that, through inference, ensures that that specific ant was not at any other location at t_0.

Negative information (P^-) is a set of assertions that one or more states are false; i.e., that the negation of each state holds.

$$P^- = \{\neg x_1, \neg x_2, .., \neg x_n\} \quad (28)$$

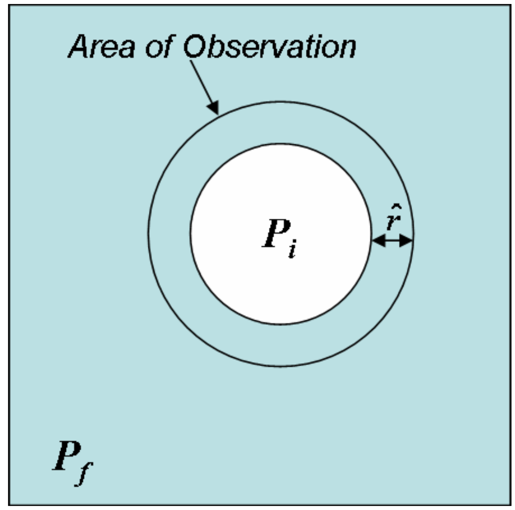


Figure 1 – A negative observation in $\mathbf{P}_w$ projected onto $\mathbf{Z}^2$ space

For example, a message stating that an observation at time t_0 found that an element did not exist at a specific location is negative information. Information partially describing a world state can be expressed either through positive information or negative information. In practice, observations frequently produce a mixture of positive and negative information. Positive and negative information both create infeasible regions in the state space and the information gain from a message is proportional to the reduction in feasible states.

This can be visualized as If message *m* encodes negative information, the information gain from *m* is

$$I_m^- = \log_2 \left| P^- \right| \quad (29)$$

information describing a hole of infeasibility in the state space [Fig. 1].

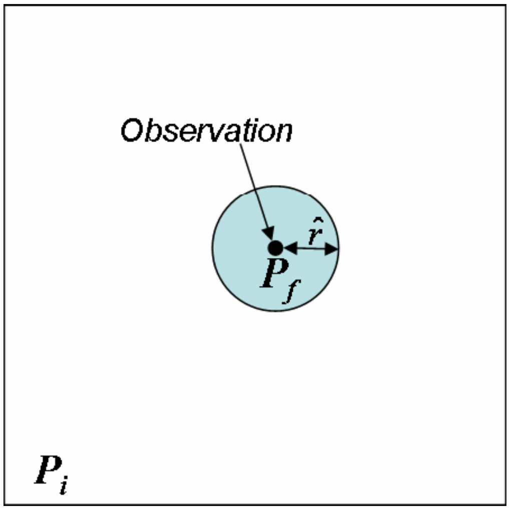


Figure 2 – A positive observation of a unique object in P_w projected onto Z^2 space

The information gained from positive information given in message *m* is equal to the size of the space outside of the positive information set:

$$I_m^+ = \log_2\left|P_w\right| - \log_2\left|P^+\right| \quad (30)$$

This can be visualized as a feasibility volume surrounded by infeasibility [Fig. 2].

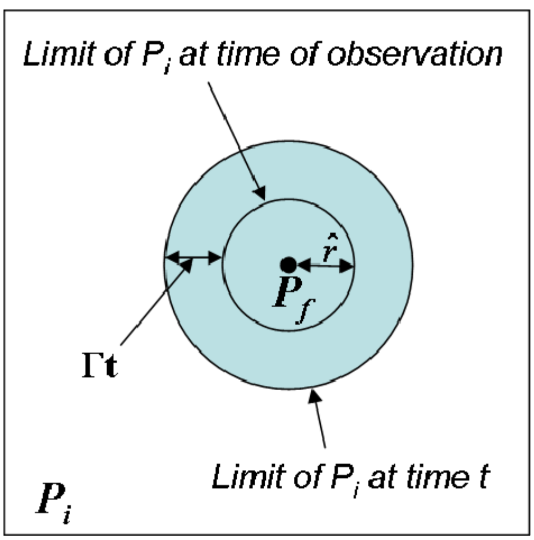


Figure 3 – Reducing infeasible state space after a positive observation

As information from positive observations entropies, information is increasingly lost at a rate that increases in proportion to the dimensionality of the original messages. The loss of positive information can be envisioned as the feasibility frontier growing outward from a clustered set of feasible states [Fig. 3]. Information content decreases until eventually all states become feasible.

As negative information entropies, information is decreasingly lost in proportion to the dimensionality of the world. This can be envisioned as the feasibility frontier growing inward, removing a diminishing set of clustered infeasible states [Fig. 4].

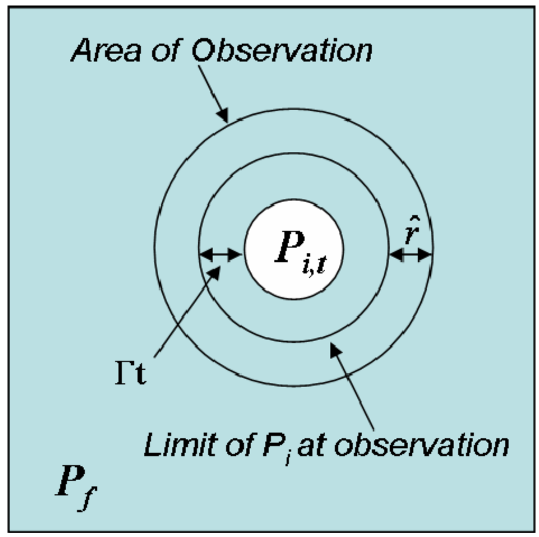


Figure 4 – Reducing infeasible state space after a negative observation

Each closed surface either grows (I^+) or shrinks (I^-) due to entropic drag until it: (i) terminates in a singularity; (ii) terminates against other surfaces; or (iii) terminates against the boundaries of the state space. The information loss for the world is a piecewise linear surface that is the sum of the information loss for each surface [Fig. 5]. In Fig. 5, $\mathbf{s_i}$ is the border between the shaded feasibility state space and the unshaded infeasible state space.

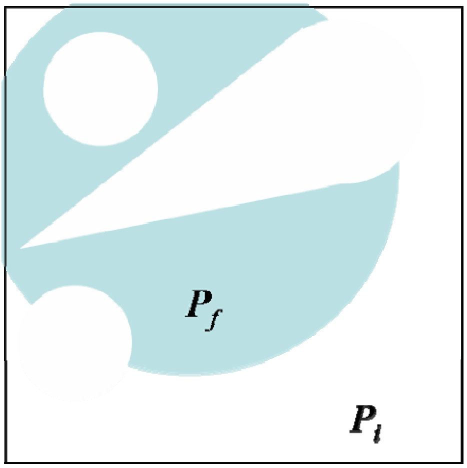


Figure 5 – Piecewise linear state space resulting from I^+ and I^- messages

Information Channels

Information channels (C(t)) are streams of information, with information flow—measured in bits per second—being the bandwidth of a channel (β). We use "information channel" as an abstraction representing the aggregate ability of a system to acquire and process new information although information can be, and often is, a traditional communications channel. The maximum capacity of an information channel defines the rate at which a system is capable of acquiring new information. In ideal conditions where C(t) = max β, information on a static world increases linearly until the information capacity of the world is reached [Fig. 6].

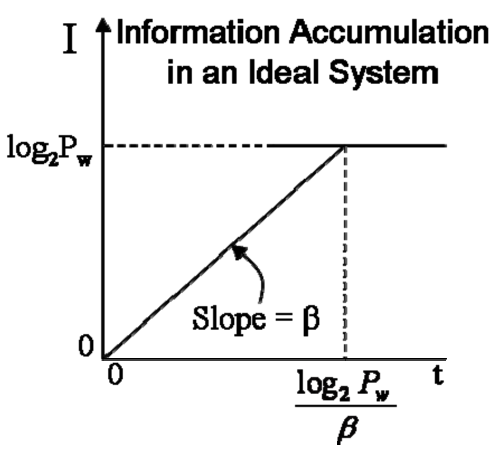


Figure 6 – Information accumulation from ideal communications

When the information in a channel describes a dynamic world, the rate of information gained from a constant information channel is reduced by the entropic drag on information within the channel *and* the entropic drag on previously acquired information. The amount of information accumulation on a world from an information channel is:

$$I = I_0 + \int_0^{\frac{I_m}{\beta}} \big(C(t) - \Gamma(t)\big)dt \qquad (31)$$

As mentioned briefly above, edge effects can play an important role in the rate of entropic drag. We highlight two basic cases here. In the first case the information gathering bandwidth is sufficient to exhaustively describe the world before the information content of the first message has vanished in a singularity or against an edge. In this case the information bandwidth exceeds the maximum entropic drag and the information content will become maximized when the information space has been transmitted.

$$t_{\max} \geq \frac{\log_2(P_w)}{\beta} \qquad (32)$$

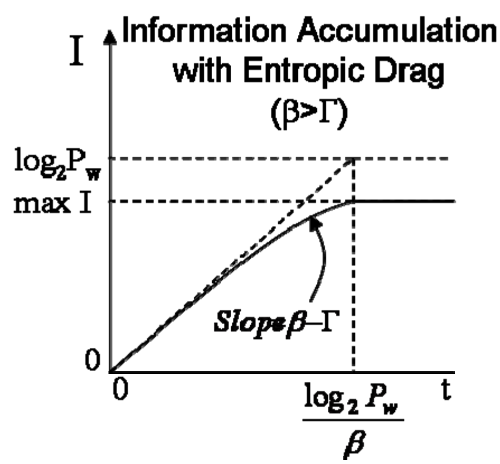


Figure 7 - Information accumulation with entropic drag

Once again assuming our ideal continuous channel where C(t) = max β, the information accumulation for a dynamic system with sufficient bandwidth is shown in Fig. 7. In this figure the maximum information is

$$\max I = \log_2 P_w - \Gamma\left(\frac{\log_2(P_w)}{\beta}\right) \quad (33)$$

In this case information monotonically increases until time t_{max} after which further messages can only maintain the current amount of information, as the information gained by successive messages cannot exceed the combination of information lost due to entropic drag and the information loss through redundancy with prior information.

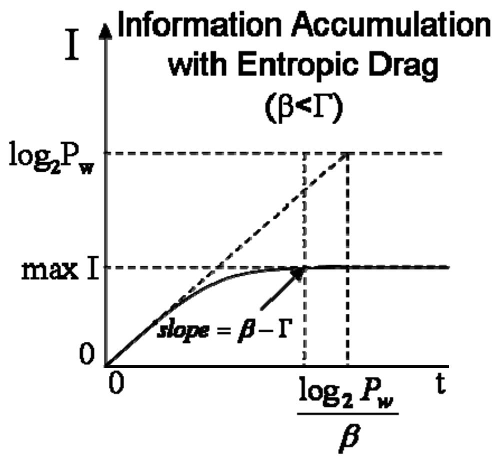


Figure 8 – Information accumulation

In the second case the bandwidth is less than the maximum entropic drag on the entire world and, as the entropic drag increases in proportion to the information known about the world, a stable point will be reached when the entropic drag equals the bandwidth. This point of informational stability is the maximum feasible information for the system [Fig. 8].

Note that in both cases the maximum information reached is less than the information space of the world. This gap is the minimum information loss associated with describing a dynamic world.

Uncertainty Redux

A key design decision for the construction of command and control systems is the selection of the fidelity of information being communicated. Fidelity is adjustable as the designer can arbitrarily decide the unit measure for each dimension in the world within sensor limitations. For example, a designer can choose to send positions in which the lowest bit of information represents a millimeter, meter or kilometer. Typically the fidelity, or bit accuracy, of a measurement is set to be slightly less than the resolving power of the sensor. This is a useful heuristic when the uncertainty associated with sensor fidelity is much larger than the uncertainty associated with information entropy. However, when the speed of that which is being observed increases, or the complexity of the world being observed increases, or the number of recipients of the observed information increases the uncertainty related to information can become an important design criterion for the command and control system. It is these cases that we are most interested in understanding.

Earlier we showed how a minimum amount of uncertainty (ε) is associated with the choice of representation. The minimum uncertainty of a system (U) at some time *t* is the aggregation of the system's representational uncertainty (U_r) and the feasible state space translated from bits to real world units.

$$U = U_r + \hat{r} \cdot |P_f| \quad (34)$$

If a system designer wanted to design a system to obtain the maximum amount of information possible, and an ideal sensor capable of observing the entire world was available, what level of fidelity should be used? At observation time our ideal sensor would reduce the feasible state space to a single state: $|P_f| = 1$. The representational uncertainty is the product of the size of the information frontier and the informational error:

$$U_r = |S_t| \cdot \varepsilon_r \quad (35)$$

where $s_i \in \mathbf{S_t}$ is the set of piecewise smooth surfaces that define the boundary between P_f and P_i at time *t*. By substitution in (34), uncertainty becomes

$$U = \left|S_t\right| \cdot \varepsilon_r + \hat{r} \cdot \left|P_f\right| \quad (36)$$

Both terms of the uncertainty equation (34) vary as a function of size unit vector. However, the rate at which they vary differs, with the uncertainty due to representational error becoming predominant as the unit representation grows and the uncertainty due to entropy becoming predominant as the unit representation shrinks. This duality allows us to identify the optimal unit representation for maximizing uncertainty.

To explore this relationship let us examine the simple case in which a C2 designer has an omniscient camera capable of completely viewing the entire world at any level of fidelity. From (16), the amount of time required to communicate the camera's observation is

$$\delta = \frac{\log_2\left|P_w\right|}{\beta} = \frac{|K|}{\beta} \cdot \log_2\left(\prod_{x \in X}\left|\frac{x_i - x_0}{\hat{r}_x}\right|\right) \quad (37)$$

For a regular world in which all n dimensions are equal, this evolves to

$$\delta = \frac{|K| \cdot n}{\beta} \cdot \left(\log_2 a - \log_2 r\right) \quad (38)$$

in which a is a constant proportional to the size of a dimension. If we assume independent entities, the feasibility space will be

$$P_f = |K| \int_0^{\delta} \vec{v}(t)dt \quad (39)$$

where $\vec{v}$ is again the n-dimensional vector of unpredictable change normalized to $\mathbf{S_t}$. In this case $\mathbf{S_t}$ is a set of K unit possibilities. The perimeter of this space is

$$S_t = |K| \oint_{\delta} \vec{v}(t)dt \quad (40)$$

By applying (39) to (36) we obtain a representational uncertainty of

$$U_r = \frac{|K|\hat{r}\sqrt{n}}{2}\left|\oint_{\delta} \vec{v}(t)dt\right| \quad (41)$$

and by (36), the overall uncertainty is

$$U = \frac{|K|\hat{r}\sqrt{n}}{2}\oint_{\delta} \vec{v}(t)dt + \hat{r} \cdot |K|\left|\int_0^{\delta} \vec{v}(t)dt\right| \quad (42)$$

This is a somewhat messy equation that expresses minimal uncertainty as a function of the unit representation (because the time δ over which the functions are now evaluated in terms of the unit representation). This equation can be evaluated when the exact dimensions of the world are known and a model of the entropic forces is available. When evaluated it shows minimal uncertainty for a given unit length. An example evaluation for a simple two-dimensional world is shown in Fig. 9, in which the graph of uncertainty diminishes rapidly as fidelity increases (moving from right to left in the figure) until it reaches the optimal resolution at the local minimum after which uncertainty begins to increase as entropic drag takes hold.

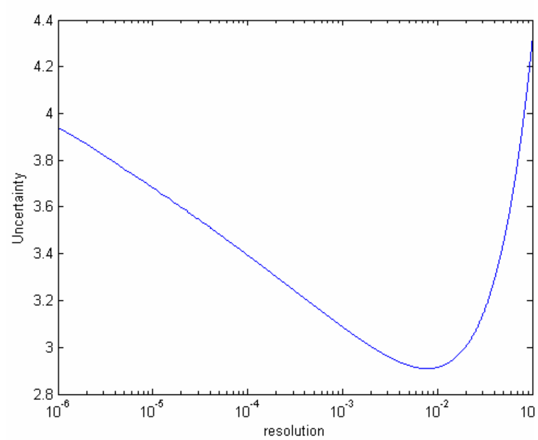


Figure 9 – Uncertainty as a Function of Unit Resolution

Observation Inefficiencies

We have been assuming that messages are ideal, with each observation bit translating to a single bit of information. This is true if and only if no a priori knowledge exists. This includes infeasible states derived from prior observations as well as probability distributions across the state space. When a priori knowledge exists, the actual information gained is the product of the efficiency (γ) of an observation and the maximum capacity of the communications channel.

$$I_{actual} = \gamma \cdot C(t) \qquad (43)$$

$$I = I_0 + \int_0^{\frac{I_m}{\beta}} \left(\gamma \cdot C(t) - \Gamma(t)\right) dt \qquad (44)$$

Koopman showed that optimally efficient search patterns can exist for regular worlds [9], yet the majority of real-world intelligence, surveillance and reconnaissance (ISR) systems are fundamentally inefficient because, over time, sensors will traverse an area over which some a priori knowledge exists. The rate of observational inefficiencies over time can be highly nonlinear as moving sensors cross feasibility frontiers. However, inefficiency rates for observing a regular environment can be approximated by decaying the information achieved with each successive observation such that

$$\dot{I} = I(1-\lambda) \qquad (45)$$

where the rate of decay (λ) is proportional to the probability that a state deemed infeasible by the observation was previously known as infeasible times information bandwidth.

$$\lambda = \beta \cdot P(p_{i_o} \mid p_{i_a}) \qquad (46)$$

In order to explore entropic relationships within inefficient systems we notionally describe this efficiency loss as a decay function. The loss in efficiency is highly dependent upon the system and its current use the use of a decay function to represent loss of efficiency is a useful approximation of real-world ISR [10]. Inefficiency results in a reduction in channel capacity, C(t).

Because it reduces the amount of knowledge accumulated over time, the forces of entropic drag increase the utility of re-searching previously observed areas. We measure this effect by replacing the constant information value in equations (19) with a decay function $\lambda(I,t) \rightarrow I$ to show the relationships between information variant messages and information entropy.

$$I = \int_0^t \left(\lambda(I_m, t) - \Gamma(\lambda(I_m, t))\right) dt \qquad (47)$$

When deploying a sensor motion strategy that searches the complete range of the world the entropic drag can become so large that it overcomes information gain. This effect is shown in Fig. 10 which shows the effect of a linear entropic drag on the observation environment shown in Fig. 9. Fig. 9 shows that the maximum amount of information is found after an ideal number of observations have occurred. If the goal of a command and control system is to make decisions based upon the most complete information possible, decisions would be made immediately after these observations as the entropic drag effect causes degradation beyond this point.

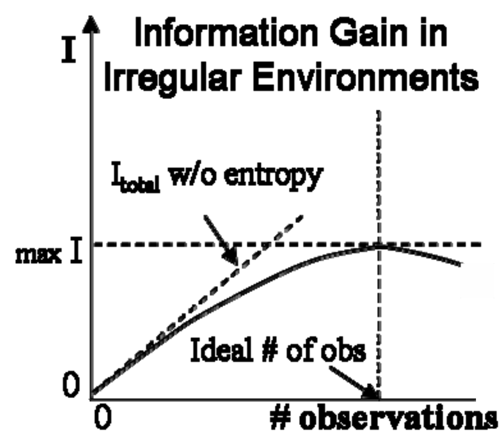


Figure 10

While we use a linear function in our example, we recognize that the information function may be highly non-linear. However, non-linearity does not negate the principle shown here.

Entropic Drag and Network Topologies

Simulation experiments were conducted to examine the effect of entropic drag on a simple command and control network infrastructure. These experiments examined the flow and relevance of *information* throughout the network. The bulk of the simulation's requirements lie in the terms

"C2 network infrastructure" and "information."

A command and control network infrastructure consists of a set of communicating entities (or nodes) whose communications topology forms an acyclic tree. The lines of communication between nodes (edges) also correspond to the chain of command. Thus, a node's authority in the network is inversely proportional to its depth in the communications tree (i.e., the root node of the tree has the highest authority and leaf nodes have the least). An additional property of the communications/command tree is that all nodes at the same depth in the tree have the same number of directly reporting subordinates (child nodes). This mapping of rank to number of immediate subordinates is the property that distinguishes alternate C2 infrastructures in the simulator.

Within the simulator, information consists of data gathered from the environment by leaf nodes in the C2 infrastructure. Each piece of information is an abstract quantity that is independent of other pieces of information (e.g., information does not overlap or correspond to multiple measurements of a known target in the environment as might be the case in a filtering problem). Information can be generated by leaf nodes periodically or stochastically, depending on the simulator configuration. The value of each piece of information is a number in the interval [0,1], where 1 corresponds to maximum value and 0 corresponds to no value.

Another property of information is that it only becomes useful to an entity after the information has been fused into the local world model. Thus, information is subject to two primary sources of latency before it can increase the knowledge of a network entity: latency due to network communications and latency due to the local information fusion algorithm.

Simulation Structure

The C2 network information simulation is a modular, discrete time-step-based simulation whose primary components are network entities, communications links, and information processing algorithms. In addition, the simulation tracks a number of metrics for each piece of information. Two key metrics are *information area* (the number of nodes that finished fusing the information in the current time step) and *information volume* (the number of nodes that have fused the information at or before the current time step). These metrics are taken from the literature on the performance of real-world networks. The overall value of a piece of information is derived by multiplying the information volume by the associated entropic drag. Since these area and volume metrics are affected by the number of nodes in the network, when comparing different C2 topologies we typically consider topologies with the same number of leaf nodes and then restrict the results to leaf node areas and volumes.

The simulation has a number of variable input parameters including simulation duration (time steps), *C2* topology, entropic drag (value lost/time step), latency along a network communications link (time steps/observation), and latency due to fusion. Unlike other simulation parameters which are constant, fusion latency can optionally be a function of the number of prior fusion operations (allowing for nonlinear fusion complexity).

Simulation Outputs

Fig. 11 shows a sample information volume plot for a single simulation run. These plots generally contain three pieces of information: the ideal information volume (red line), the information volume without accounting for entropic drag (blue line) and the information volume with entropic drag (green line). The ideal information volume depicts the spread of information throughout the network assuming no latency due to communications or fusion and hence no entropic drag. Thus, it provides an (admittedly unrealistic) upper bound on the amount of information in the network. Nonetheless, it provides a strong indication that an upper bound exists. The un-decayed information volume depicts the network information volume that would result when considering latency but not entropic drag. The decayed information volume takes into account both the effects of latency and entropic drag, and represents the main result of interest. One expects both the ideal and un-decayed information volumes to monotonically increase, with the entropic drag volume trailing behind the ideal volume. The entropic drag volume shows that there exists a maximum information volume (time=175) for the system. Further, it shows that this peak occurs prior to the complete distribution of information across the network across the network (time=240).

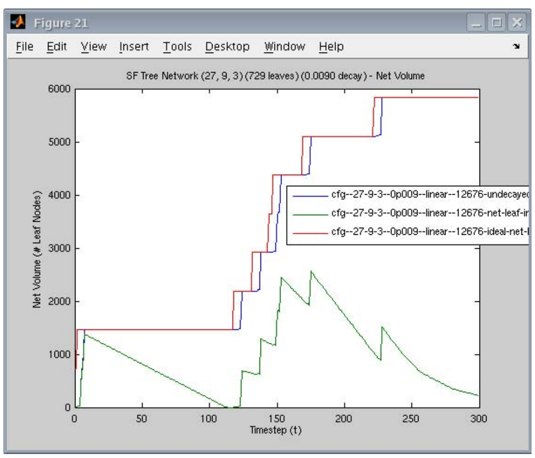


Figure 11 – Information volume in ideal static worlds, latent static worlds and dynamic worlds

A key feature of the simulation output is the ability to compare one or more simulation runs. Fig. 12 is a plot comparing multiple runs of the simulation where the rate of entropic drag was varied while the other simulation properties were held constant. The upper lines in the plot have progressively lower decays. In Fig. 12 one the existence of maximum information volume prior to full dissemination of information across the network is again in evidence. We can also see that the time at which the information maximum occurs varies with respect to the entropic drag. The information maximum occurs at time step 230 for the runs with lower information drag (lines at the top of the graph), at time step 175 for the runs with higher entropic drag (lines at the bottom) and equally across time step 230 and 175 for the lines in the middle.

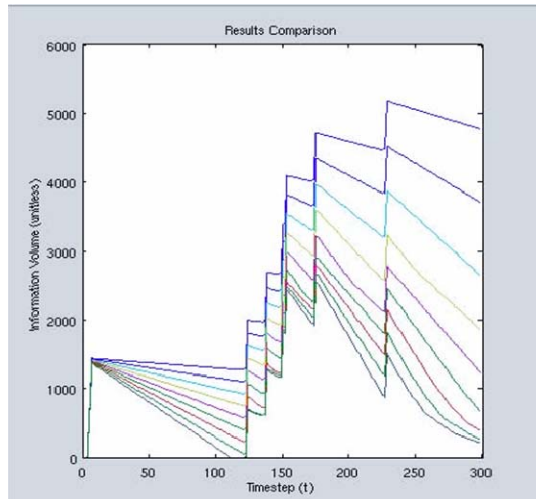


Figure 12 – Information volume as Γ varies

Future Efforts

This paper is a first step in the application of information theoretic entropy to command and control. Large bodies of work in information management, particularly data fusion, networking, and the pantheon of group control strategies need to be looked at though the lens of entropic drag. Our forward looking hypothesis is that an understanding of the entropic effects of information will allow C2 designers and in-the-field decision makers to employ command and control strategies that are optimized for any given situation.

The next significant step will be to extend the study of entropic drag into the latter portion of the OODA loop, that portion which involves decisions and actions; Fry [11] makes the case that the information entropy impact on the "DA" portion of the loop is the dual of the "OO" portion of the loop. This view is also reflected by Alberts & Hayes [12]. We concur that the impact of entropic drag on control, when viewed independently of observation, is equivalent to the impact of entropic drag on observation. However, control is dependent upon observation and the impact of entropic drag is likely to be exacerbated in the control portion of the OODA loop.

Another area of investigation involves an application of these formalisms to real-world scenarios. The state of a dynamical system changes according to the rules specified by the transition function f. The nature of this transition function can have a dominant effect on what set of states can become feasible in the future. The laws that govern the transition function are a priori information, reducing the state space of the system. In this paper we assumed ideal coding of information describing the state space. In practice it is more common (and practical) to encode information generally and to use inference based upon a priori information to reduce the working feasibility space. The investigation of these relationships will be a key part of the scenario-specific investigation.

Another important aspect of this investigation must be the entropic effects of the communications infrastructure. We divide communications infrastructures into channel communications and topologies.

Further work will also be required to improve the metrics that are used to measure the dynamism within an environment. Environmental dynamics are driven by complexity and the pace of environmental change; however, is not well understood how complexity and pace should be measured in real-world environments. These further advancements should enable the pursuit of our long-range objective, namely the construction of an adaptive command and control system that autonomously observes the environment and changes the network topology and information and decision-making strategies to optimize C2 performance.

Conclusion

We have shown that the loss of information due to dynamic forces within an environment can have a substantial impact upon the information content of one or more messages about the environment. This effect, called entropic drag, fundamentally impacts the effectiveness of command and control systems. The principles outlined in this paper can be used provide a better understanding of the utility of existing command and control systems and to improve the design of future ones.

Entropic drag impacts C2 systems in several important ways. First, entropic drag enforces a fundamental limit to the amount of information that can be known about a dynamic system. This limit can be used by C2 designers to identify the maximum useful fidelity of C2 semantics. Second, by expressing the relationship between information and time entropic drag allows a decision maker to identify when the optimal amount of information has been acquired. Third, entropic drag provides a framework for understanding the flow of information across a networked community, providing insight into the utility (or lack thereof) of sharing information with each member of the community as well as providing insight into the utility of alternative network topologies. Fourth, entropic drag provides tools to better understand the tradeoff between information and latency in control, allowing decision makers to select the optimal amount of information to use for performing a specific task or set of tasks. Fifth, entropic drag provides tools to understand the utility of collaboration in shared decision making, allowing decision makers to correctly scope the degree of collaboration for optimal performance of a task. Finally, entropic drag provides a framework for the development of next generation, adaptive C2 infrastructures,

infrastructures that autonomously adapt information and decision sharing strategies and networking topology at run time in response to an environment's changing dynamic forces.

References

[1] Heeger, D., Signal Detection Theory, http://www.cns.nyu.edu/~david/sdt/sdt.html, New York University.

[2] Shannon, C. E., A Mathematical Theory of Communications, *The Bell Systems Technical Journal*, Vol 27, pp. 379-423, 623-656, July, Oct; 1948.

[3] Perry, Burton et al, *Measures of Effectiveness for the Information Age Navy: The Effects of Net-Centric Operations on Combat Outcomes*. (2002) RAND Santa Monica, CA, USA.

[4] Moffat, *Complexity Theory and Network Centric Warfare*. DoD Command and Control Research Program, 2003.

[5] Caves, C., Entropy and Information: How Much Information is Needed to Assign a Probability?, *Complexity, Entropy and the Physics of Information*, Zuruk, W. (ed.), Westview Press, 1989.

[6] Brillouin, L., *Science and Information Theory* 2nd Ed., Dover Publications Inc., Mineola, NY, 2004.

[7] Smith, *Effects Based Operations*, DoD Command and Control Research Program, 2002.

[8] MacKay, D., *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, pg 67, 2005.

[9] Koopman, B., *Search and Screening*. Operations Evaluations Group Report No. 56, Center for Naval Analyses, Alexandria, VA, 1946.

[10] Stone, L., *Theory of Optimal Search*. Academic Press, New York, 1975.

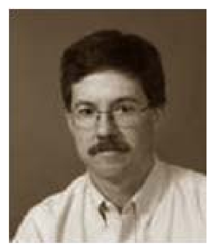
[11] Fry, R. An information-theoretic formulation of intelligent weapon systems. Technical Memo A1F(1)06-U-053, Johns Hopkins University Applied Physics laboratory, Dec., 2006.

[12] Alberts, D., Hayes, R., *Planning Complex Endeavors* [Community Review version], CCRP, 2006, pg. 67.

Acknowledgements

This research effort was supported by Johns Hopkins University Applied Physics Laboratory internal research grant X9ANXC2M. Thanks to Mike Miller, Joe Peri and David Viel for their excellent critique and advice.

Biographies



David H. Scheidt received a B.S. in computer engineering from Case Western Reserve University in 1985. He is a branch scientist at the Johns Hopkins University Applied Physics Laboratory where he chairs the laboratory's Autonomy Research Thrust. His current research concentrations are: distributed robotics, command and control of decentralized systems, distributed intelligent control and cooperative systems. Mr. Scheidt has twice received the National Performance Review's Hammer Award.

Michael J. Pekala is a Computer Scientist for the National Security Technology Department of the Johns Hopkins University Applied Physics Laboratory. He received B.S. and M.S. degrees in Computer Science from the Johns Hopkins University in 1998 and 2001 respectively. Mr. Pekala joined the Laboratory in 1998, and his current focus is on the research and development of autonomous systems. His email address is mike.pekala@jhuapl.edu.