Annual Reviews in Control 49 (2020) 15–26

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Autonomy and metrics of autonomy

Panos Antsaklis*

Department of Electrical Engineering, University of Notre Dame, USA

ARTICLE INFO

Article history:
Received 2 May 2020
Accepted 3 May 2020
Available online 20 May 2020

Keywords:
Autonomy
Autonomy level
Autonomous control
Intelligent control
Adaptive control
Uncertainty

ABSTRACT

The quest for autonomy has been a pervasive theme in human culture through-out history. In this paper a general definition of autonomous systems is presented and discussed that leads naturally to the establishment of metrics to measure the level of autonomy of a system. This definition is based on the systems ability to achieve goals under uncertainties and it does not involve the means by which the goals are achieved, such as sensing and feedback. This paper takes the point of view that any autonomous system is a control system, and that to achieve higher levels of autonomy one may need to add methods traditionally developed in areas such as operations research and AI. The work presented here is based on earlier work by the author on functional architectures for autonomous spacecrafts.

Contents

1. Introduction

Autonomous vehicles have certainly captured the imagination of everyone, in recent years. It is fascinating to have machines being able to drive us around autonomously, without a human driver. In addition, the promise of reducing or even eliminating accidents via autonomy has been very appealing indeed, and more so because of the convincing marketing strategies of the world's largest high tech and automobile companies. Furthermore, significant progress in unmanned aerial vehicles (UAV) and in autonomous underwater and surface ships is announced daily. These are exciting times, especially for researchers in control systems.

* Corresponding author.
E-mail address: antsaklis.1@nd.edu

https://doi.org/10.1016/j.arcontrol.2020.05.001
1367-5788/

Autonomy in engineered systems is not a new concept. Automatic pilots for aircrafts and ships that increase the degree of autonomy of the system, have been operating very successfully for many years — the first autopilot for aircraft was introduced in 1912. Furthermore, autonomy is not a new concept in society, in politics, in companies and organizations, in biology, to mention but a few.

These engineered systems are examples of accomplishments of the *Quest for Autonomy*, a pervasive theme in engineered systems through the centuries starting even earlier than Ktesibios' water-clock with its feedback mechanism in the 3rd century BC and continuing strong today. It seems that we always wanted to build things that did more things by themselves, that served us. In fact, as it was mentioned in the works of the ancient poets Hesiod and Homer around 700BC, Hephaestus the Greek god of invention and blacksmithing had created several creatures that were accomplishing different tasks by themselves. One of them was Talos, a giant bronze man commissioned by Zeus to protect the island of Crete. As the story goes, Talos marched around the island three times every day autonomously and hurled boulders at approaching enemy ships! (Mayor, 2018), see also (Antsaklis, Vachtsevanos, & Polycarpou, 2005; Valavanis, Vachtsevanos, & Antsaklis, 2007; 2014).

When people refer to autonomous systems they often mean different things. It is important to be more precise and agree upon a common definition.

The purpose of the present paper is to define autonomy, describe concrete ways to talk about autonomy and levels or degrees of autonomy and provide quantitative relations.

The terms autonomy and autonomous imply that the system has the capability to accomplish certain goals. A system is autonomous always with respect to certain goals; it is not autonomous for the sake of being autonomous. Although this appears to be an obvious point, if it is not recognized, it can lead to misunderstandings when attempting to compare autonomous systems.

A second major point is that a system which is autonomous with respect to certain goals is always subject to uncertainties in the system and its environment; uncertainties that affect the abilities of the system to accomplish the goals. That is, uncertainties are always present, because if they were not and we had perfect knowledge of the system and its environment, we would have been able to preprogram the system to accomplish any goals — possible by its structure and its environment.

Therefore, when we state that a system is autonomous we are really saying —or we should be saying— that the system is autonomous with respect to a set of goals subject to a set of uncertainties.

Such use of terminology is analogous to the use of the term optimal which really means optimal with respect to certain optimization criteria, say minimum cost (goals), subject to constraints (restrictions such as uncertainties). Just stating that a solution is optimal is too vague. In fact, anything can be optimal with respect to something! The constraints restrict the set of possible solutions. The constraints may be severe enough for a feasible solution not to exist (set of constraints being infeasible) in which case no optimal solution exists (in fact no solution exists at all). Similarly, the uncertainties restrict the set of possible control policies that achieve the goals. The uncertainties may be large enough for no control policies to exist that achieve the goals (the set of uncertainties render the problem infeasible) in which case no control policy exists that make the system autonomous with respect to the given set of goals.

In view of the above we can characterize autonomy as follows:

> If a system has the capacity to achieve a set of goals under a set of uncertainties in the system and its environment, by itself, without external intervention, then it will be called autonomous with respect to the set of goals under the set of uncertainties.

For the same set of goals, the larger the set of uncertainties the system can handle, the higher is its degree of autonomy. The lower the needed external intervention by humans or other systems to achieve the goals under the uncertainties, the higher the degree of autonomy. So, the level of autonomy depends on both, a measure of the set of the goals that are being accomplished and a measure of the set of uncertainties present. Specifically, {Measure of the Set of Goals} $\times$ {Measure of the Set of Uncertainties}$=L$, the level of autonomy. This definition allows the comparison of the autonomy levels of different systems.

An autonomous system has goals to be accomplished and mechanisms to accomplish those goals under uncertainties. This is exactly what a control system does. Control policies are added to satisfy certain specifications under uncertainties. Therefore, *every autonomous system is a control system*. Adding to traditional control systems advanced sensing and incorporating decision making from areas such as AI is a way to increase substantially the level or degree of autonomy of a system. Control systems should be seen as the cornerstone of autonomous dynamic systems.

The issues outlined above are discussed in detail in this paper.

The present paper focuses on measures of autonomy with emphasis on comparing levels or degrees of autonomy. The definition of autonomy used here was first presented in Antsaklis (2017) and further discussed in Antsaklis and Rahnama (2018) where the main ideas behind defining levels of autonomy were elaborated upon.

It should be noted that the concepts in defining autonomy using sets of goals and uncertainties have appeared in the writings of the author, published in the open literature, much earlier; see for example (Antsaklis, Passino, & Wang, 1988; Antsaklis & Passino, 1993; Antsaklis, Passino, & Wang, 1989; 1991). Autonomy in engineering systems and its relation to intelligent behavior was discussed in the task force report (Antsaklis, 1994). Details of defining levels of autonomy were discussed in a paper draft (Antsaklis, 2018) which was circulated and commented upon by colleagues. These ideas were also presented in a keynote address at the Mathworks Research Summit in early June 2019.

To appreciate what is needed to achieve high levels of autonomy a conceptual description of a functional architecture for an autonomous spacecraft is given. This description first appeared in Antsaklis et al. (1988) and in journal paper form in Antsaklis et al. (1989). The version included in the present paper may be found in Antsaklis (2011) and it is given in Appendix B. It is recommended that the reader reads Appendix B, even before proceeding to the main body of this paper.

In the following, we start the discussion with our definition of autonomy. The interested reader may want to read materials from Antsaklis et al. (1988); Antsaklis (1998, 2011); Antsaklis et al. (1989, 1991); Antsaklis and Rahnama (2018); Antsaklis (1994) and consult the references therein which describe early research (in the late 1980s to mid 1990s) in combining control systems with intelligent methods from artificial intelligence and machine learning to design highly autonomous intelligent control systems. For additional discussion of autonomy and its levels see, for example, Beer, Fisk, and Rogers (2014) where the definitions introduced correspond to the definitions in this paper in that they use task-specific goals to be achieved by the system and refer to needed outside intervention instead of uncertainties present; that definitions also involve the means by which autonomy is achieved which is not part of the definition in our approach. See also (Barber & Martin, 1999; Durst & Gray, 2014; Hrabia, Masuch, & Albayrak, 2015; Huang et al., 2007; NIST, 2000). Note that a definition involving necessary outside interventions to achieve the goals is discussed later in this paper. It should be noted that, contrary to other definitions, our definition of autonomy does not involve descriptions of the means by which a specific level of autonomy is achieved, whether smart sensors or intelligent decision making are involved. We find it more useful to characterize autonomy using only the possible to achieve goals under given uncertainties and letting the specific means by which the level of autonomy is achieved to be used in characterizations of the system as smart, intelligent etc. In fact, as it was stated many times in our publications, "autonomy is the goal and intelligent means is one way to achieve it." Higher autonomy typically involves higher intelligence.

Our definition of autonomous behavior provides a natural way to define levels or degrees of autonomy via simple quantitative relations, specifically, as it was mentioned above, {Measure of the Set of Goals} $\times$ {Measure of the Set of Uncertainties}$= L$, the level of autonomy. This easily leads to an intriguing and interesting relation, namely {Performance} $\times$ {Robustness}$= L$, the level of autonomy. Here Performance is a measure of the set of goals that can be achieved (and it may include stability) and Robustness (Resilience) is a measure of the set of uncertainties under which the goals are reached. Systems with higher performance and/or higher Robustness/Resilience have higher degree of autonomy. These issues are discussed in detail later in the present paper.

Entropy can also be used as a general measure of the set of uncertainties. Entropy in autonomy is also discussed here.

An additional interesting measure is the degree of external intervention needed to achieve the set of goals. The higher the needed external intervention the lower the level of uncertainties under which the goals can be achieved; that is there exists an inversely proportional relation between the level of needed external intervention and the level of uncertainties or robustness under which the system operates when achieving the set of goals.

Examples are used throughout this paper to illustrate the concepts including a glimpse of how these definitions and new relations may be applied to the 5 autonomous vehicle levels used in the self-driving car literature and industry.

2. Autonomous systems

We start with the etymology of the word autonomy:

The term autonomy originated in Ancient Greek: αὐτονομία (*autonomia*), from αὐτόνομος (*autonomos*), which comes from αὐτο (*auto*) "self" and νόμος (*nomos*) "law", hence when combined it is understood to mean one who gives oneself his/her own law. *Autonomous means having the capability and authority for self-government.*

Autonomy goals: A system exhibits autonomous behavior of interest only when is achieving a goal or a set of goals. That is, autonomy without clearly identified goals, autonomy for the sake of autonomy is not interesting, if we want to build useful engineering systems. Autonomy without goals is as vague a concept as claiming that something is optimal without specifying a measure, such as a cost to be minimized. For example, a goal of an autonomous train could be to move passengers safely from station to station following a time schedule with some probability; the goal of a speed cruise control of an automobile is to control the car so to maintain approximately constant speed.

Every autonomous system is a control system: An autonomous system always has a set of goals to be achieved and a control mechanism to achieve them. This implies that *every autonomous system is a control system*. Here the term "control system" is used in a most general sense, in which control (a decision mechanism typically using sensor measurements and feedback together with ways to implement these decisions via actuators) is used to make the system (a very general collection of processes) attain desirable goals.

As it was mentioned above, the word control in autonomous control has a more general meaning than in conventional control; in fact, it is closer to the way the term control is used in every-day language; see Antsaklis et al. (1988). To illustrate, in a rolling steel mill, while conventional controllers may include the speed (rpm) regulators of the steel rollers, in the autonomous control framework one may include in addition, fault diagnosis and alarm systems; and perhaps the problem of deciding on the set points of the regulators, that are based on the sequence of orders processed, selected based on economic decisions, maintenance schedules, availability of machines etc. All these factors have to be considered as they play a role in controlling the whole production process, which is really the overall goal. Note that in order to increase autonomy it is typical to implement several layers/levels of automation. Local controllers are often referred to as level 1 automation, set points assignment as level 2, and so on.

System and its environment: As it is typically done in the field of control systems, it is useful to think of a system to be controlled as being surrounded by a boundary separating it from its environment. The system acts upon its environment through its outputs and receives inputs in the form of disturbances or additional information. What the system includes within its boundary, expressed via the particular system model used, depends of course on the goals and the characteristics/properties used to achieve its goals.

Goals and Uncertainties: In addition to the set of goals to be attained the other central component of autonomy is the set of uncertainties. For example, in the above cruise control example, the speed needs to be maintained (goal) under varying external conditions such as road incline, condition of road surface, wind gusts, as well as internal varying vehicle conditions such as hot or cold engine and age of the car (uncertainties). Clearly the uncertainties of interest in an autonomous system are the ones that affect the accomplishments of the goals.

So, autonomy is the ability of a system to achieve a set of goals under uncertainties in the system and its environment. Autonomy exists only with respect to a set of goals and it is of value when there are uncertainties. If there were no uncertainties, we could program the system ahead of time, in which case a macro-command would be adequate. In control system theory if we had complete knowledge of the system to be controlled and of the external disturbances then we could only use open loop control and the control problem would have been rather straightforward. Uncertainties however are always present in different degrees. For example, in the above case of the train moving on fixed rails from station to station, as in an airport terminal, there are reasonable guarantees that no passenger will cross the rails and there will be an unobstructed path for the train and so the uncertainties are rather limited and are primarily caused by variations in the flow of passengers in and out the train at each station. This problem is manageable and currently such automated trains are operating successfully in many airports around the world. Compare this with a car moving from point A to point B. Even if we assume that the car stays in the same lane, the problem is much harder, compared to the train example above, because there are uncertainties such as traffic lights, other cars changing lanes without warning, pedestrians crossing unexpectedly, the weather that affects sensors and braking distance and so on. Because of the increased uncertainties designing autonomous cars to operate in a city is much harder than designing autonomous trains to operate in an airport terminal. It should be noted that significant successes have been achieved in airplane automatic pilot systems that are being used thousands of times daily which maintain direction, speed and altitude under unexpected disturbances such as gusts of wind and air pockets.

In view of the above discussion we introduce the following definition which captures the fact that autonomy should always be considered in terms of goals attained under uncertainty.

Given a system *S*, let *G* be a set of goals to be achieved under a set of uncertainties *U*.

Definition 1. A system *S* is autonomous with respect to the set of goals *G* under the set of uncertainties *U*, if the system *S* is capable of achieving all goals in *G* in the presence of all uncertainties in *U*, by itself, without external intervention.

The set of uncertainties *U* is associated with the set of goals *G*. It is implied that the uncertainties considered in the above definition are the ones that are relevant to the goals considered. For example, for the goal of stability, certain uncertainty in the parameters may be relevant, but different set of uncertainties may be relevant when the goal is tracking. Other uncertainties which are irrelevant to the goals of interest do of course exist; for example, in designing the autopilot of an aircraft we do not consider the interior design of the passenger cabin unless it has implications on the weight of the aircraft.

Long term autonomy is of interest. It is assumed that the system *S* is able to perform these functions autonomously over a significant time horizon; that is, this is a repeatable function the system is capable of, over extended periods of time.

It is possible to have as a goal to control the system so that some property is attained with certain probability. For example, the goal could be to attain asymptotic stability with a probability of 95%. So, the above definition captures the realistic scenarios of achieving goals with certain likelihood.

The above definition of autonomy should and does apply to organizations and natural systems as well. For example, in an organization, a team led by a manager accomplishes a set of tasks under uncertainties such as personnel absences and equipment breakdowns, independently, without intervention by a general manager. A bacterium may be able to reach a light under normal circumstances, but needs external help to remove unexpected obstacles in its path.

Autonomous systems should be able to collaborate with humans to accomplish enhanced goals which are not attainable by just the human or just the autonomous system. Adaptive autonomy is of interest here. Imagine the scenario where the driver in an automobile carries out variable tasks depending on how these tasks are shared with the autonomous vehicle. For instance, the driver may want to take on the task of maintaining certain distance from another vehicle, that is taking on the advanced cruise control functions. The vehicle may take full control if the driver is not capable of driving safely due to tiredness or intoxication.

The impact of autonomous systems on society is of significant interest. It is most important to adopt autonomy in stages, providing education to prepare the population for the changes. Such changes accompanied by temporary loss of jobs have occurred several times in the past due to the industrial revolution and there are many lessons to be learned from that, on how to ease the impact. The hope is of course that autonomy and automation in the long run will create more jobs that the ones that were lost. The difficulty is that the new jobs will probably require new sets of skills. The social impact of autonomous systems needs to be taken very seriously.

3. Levels of autonomy

It is of interest to compare the levels of autonomy in systems. Assume that a given system is autonomous with respect to a set of goals under a set of uncertainties. If another system can achieve the same goals under higher uncertainties (under a larger set of uncertainties) then clearly the second system has higher autonomy. Similarly, if more goals can be achieved under the same set of uncertainties then the system has higher autonomy.

The autonomy level of a system can be manipulated and increased by adding feedback control, adaptation, learning, planning, failure detection and reconfiguration capabilities, which in effect increase the level of uncertainties the system can cope with autonomously.

A fixed feedback control system has low degree of autonomy, because it can achieve the stability goals under rather restricted parameter variations and external disturbances. When there are more substantial parameter changes then one could use methods from adaptive control to achieve stability. Such adaptive control system has higher degree of autonomy due to greater uncertainty in the parameters it can handle.

The degree of autonomy can be interpreted as the size of an operating region (operating sphere) defined by a set of parameters within which the system acts on its own in a safe manner towards the goal. In the example of the car speed control, a typical cruise control system can keep the car speed at acceptable levels only when the road is not too steep. And such control system has certain degree of autonomy as it acts appropriately within its operating region, which is specified by the initial design of the system. We could build cruise control systems with larger operating regions satisfying the goal of keeping the speed at a preset desired level. One way to achieve this is to anticipate, via perhaps a vision system an upcoming steep grade and prepare for it by shifting gears or accelerating slightly, which is exactly what human drivers typically do. We could also have car speed control systems that may attain additional goals thus increasing even more their operating regions. For example, we could add in a car a control system that maintains the same speed as the car in the front (these are called ACC —advanced cruise control systems), and in addition it adjusts the distance between the cars depending on the speed, for safety reasons. It is clear that these two control systems, taken together can satisfy a set of goals under quite diverse conditions. Clearly such system has higher degree of autonomy.

For given set of goals, the degree of autonomy may be quantified by characterizing the safe operating region within which the system acts appropriately. This region in control systems is sometimes referred to as region (ball, sphere) of uncertainty and it is characterized by certain norm measures, when of course normed spaces are appropriate. Control systems that act appropriately in these uncertainty regions are called robust with respect to these uncertainty regions and with respect to goals such as stability (typically Lyapunov asymptotic stability) or performance.

Note that the same system *may be autonomous or not* depending on the stated goals and the uncertainties present. Furthermore, a non-autonomous system may have several autonomous functions. For example, in cars, the cruise control, the ABS, ACC, lane preserving, etc., offer autonomous functionalities and for each one of these subsystems the set of goals and the uncertainties could be identified.

Autonomous systems deal with uncertainties primarily using sensors, but also, for example, using prior knowledge and machine learning, to improve their knowledge of the processes to be controlled and also of the outside environmental influences, so to be able to achieve the goals by applying effective decision-making methods. Intervention (human or via a controller) reduces uncertainties the system has to deal with autonomously. Successful control actions, by engineered systems or human intervention, reduce the set of uncertainties that impact the goals and must be dealt with autonomously. Human intervention or adaptive/learning controllers may provide information via, for example, cognitive abilities, data bases, prior experience that reduce the uncertainties, and lead to a smaller set of uncertainties that need to be dealt with autonomously.

Measuring the degree of autonomy is non-trivial. It is perhaps straightforward to compare systems that have the same sets of goals but different uncertainties. It was pointed out above that an adaptive control system has higher degree of autonomy than a fixed feedback controller because it can handle greater parameter uncertainty in achieving stabilization (the common goal). When the goals are different as well, then the problem of measuring degrees of autonomy and comparing autonomous systems becomes more complex.

The automotive industry currently uses a useful, descriptive classification to distinguish levels of autonomy. There is a SAE scale of 5 levels (plus a zero level) with level 5 used for full autonomy. Similarly, the AFRL Autonomy Framework is used in the UAV area, where a scale of 10 levels (plus a zero level) is being used with level 10 used for full autonomy.

4. Metrics

The above discussed relationships that help us characterize different degrees of autonomy may be captured by the following very simple relations:

Level or degree of Autonomy = {Measure of the Set of Goals G} × {Measure of the Set of Uncertainties U under which the goals in G are attained}

Let M_G be a measure of the set of goals G and M_U be a measure of the set of uncertainties U and L be a measure of the level of autonomy of the system.

Then $L = M_G \times M_U$

L, the level or degree of autonomy, depends on both, the measure of the set of uncertainties and the measure of the set of the goals that can be accomplished.

The measure of the set of goals should reflect the importance, complexity and number of goals. Importance may depend on existing priorities — tracking quickly within a few seconds may be a higher priority than tracking asymptotically and in this case the level of autonomy with respect to the finite tracking is smaller if only asymptotic tracking may be achieved. Similarly, the measure of the set of uncertainties should reflect the size, frequency and number of uncertainties.

For a given level of autonomy L, when M_U decreases, M_G increases, that is under reduced uncertainty more goals can be achieved by the system. When M_U goes up, M_G goes down that is, under increased uncertainty fewer goals can be achieved.

When the goal is just stability and the uncertainties are small, that is M_G and M_U are small then the level of autonomy L is low. This is the case for example when stabilization can be achieved via a fixed feedback controller. When stabilization can be achieved under higher uncertainties, which is the case for example when adaptive control is used to stabilize a system, the level of autonomy L is higher. To increase L, when there is a fixed set of goals, one needs to increase uncertainties under which the system is capable of achieving the goals.

Appropriate controllers in effect increase the size of the set of uncertainties relevant to the goals that can be accomplished autonomously and increase the system's level of autonomy. Note that these controllers are modifying the system. Uncertainties that can be dealt with autonomously may be increased, for example using adaptation and learning, or human intervention, where extra sensors, cognitive abilities, past experience effectively increases the set of uncertainties the system can cope with autonomously. For example, consider the case when a driver intervenes and assumes certain functions to help the vehicle cope with uncertain situations.

Clearly, by introducing restrictions on the uncertainties in autonomous vehicles (e.g. adding structure — staying in the same lane, using rails, assuming good weather etc.) more goals can be achieved.

More goals can be achieved by adding additional controllers. For example, assume that a given system is stabilized via a feedback controller, which operates successfully over a set of uncertainties. If a tracking controller is added the goals that can be achieved increase; however, the set of uncertainties that can be dealt with autonomously while tracking may be reduced compared to the stabilization case.

Given a system, if there are no uncertainties at all, a much-enhanced set of goals may be achieved with appropriate controllers. For example, we could use open loop control to cancel all existing dynamics and introduce any new desired dynamics. However, note that when a system is run open loop, uncertainties in plant parameters and disturbances could deny the ability to achieve control goals, such as stability.

Given a system is there a maximum L? The answer is affirmative. For a given system there is a maximum set of goals that can be achieved. For example, the attainable goals for a self-driving vehicle do not include the ability to fly — at least not yet. Considering this maximum set of goals, consider the set of uncertainties that affect those goals and then consider the largest set of uncertainties under which this set of goals can be attained. To find the maximum autonomy level of a system, consider the measures for the set of goals and the set of uncertainties under which these goals are achieved and then maximize their product by varying the sets of goals and for each set of goals selecting the corresponding set of uncertainties that have the maximum measure.

5. Humans in the loop and adaptive autonomy

When one considers humans collaborating with engineered systems, then the overall system that includes humans in the loop may be considered autonomous with respect to a large set of goals and under a large class of uncertainties, that is having a high level of autonomy. Depending on the role of the humans in the loop and the level of control authority humans exert, the remaining system will have different degrees or levels of autonomy. So, in an automobile, if for example the goal is to keep the vehicle inside a lane while travelling with constant speed, the system may consist of the vehicle and the driver where the system attains its goals in the presence of uncertainties/disturbances, such as gusts of wind and road inclines. The driver together with the automobile's control systems provide the correct steering and gas pedal commands so the vehicle maintains its course within a lane and at certain (approximately) constant speed in the presence of uncertainties/disturbances, such as gusts of wind and road inclines. If one considers the controller to consist of just the control systems of the car without the driver, then the system, the car, has a lower degree of autonomy, meaning that it may need extra help from humans or other systems to attain the required level of autonomy.

Humans or other systems may insert themselves at different levels of a functional hierarchy (that correspond to different levels of autonomy) used to describe the operation of autonomous intelligent systems (Antsaklis et al., 1988; Antsaklis, 1998; Antsaklis et al., 1989; 1991; Antsaklis & Rahnama, 2018), and take over control functions. For example, humans may insert themselves to take over planning, failure detection and identification, reconfiguration or learning functions. Or they may insert themselves to take over lower control functions e.g. a driver may want to take over the functions of the ABS system to perform the braking pumping action on his own. Such *adaptive autonomy,* where the authority the human operator exercises may vary, appears to be a very promising direction in autonomous systems research. The level of authority of the human operator may vary and the changes may be initiated not only by the human operator, but also by the vehicle if it detects driver errors or lowering of the driver's alertness.

6. External intervention

Autonomy may also be defined in terms of needed, outside intervention necessary to achieve the goals instead of in terms of a set of disturbances. Note that an equivalent definition of autonomy is:

Definition 2. A system is autonomous with respect to a set of goals G under a set of outside interventions I (by humans or engineered systems), when the system can achieve all the goals in G, assisted by just the interventions I.

If the goals can be achieved under a smaller set of outside (human and otherwise) interventions, then the system may cope with higher uncertainties and has higher autonomy; if more goals can be achieved under the same set of interventions or the same set of uncertainties then the system has higher autonomy.

The lower the needed intervention to accomplish the goals, the higher the level of autonomy. The uncertainty the system can cope with while achieving its goals, is inversely proportional to intervention necessary to achieve the same goals.

{Measure of the Set of Interventions I under which the goals in G are attained} × {Measure of the Set of Uncertainties U under which the goals in G are attained}= a constant which is taken to be 1. That is $M_I \times M_U = 1$.

These relationships may be captured via a simple relation:

Level or degree of Autonomy = {Measure of the Set of Goals G}/{Measure of the Set of Interventions I under which the goals in G are attained}

Let L be a constant that corresponds to the level of autonomy. Let M_G be a measure of the set of goals G and let M_I be a measure of the set of needed interventions I. Then

$$L = M_G/M_I$$

Note that here it was assumed that $M_U \times M_I = 1$. That is, M_I the measure of the set of needed interventions may be taken to be inversely proportional to M_U the measure of the set of uncertainties U.

As an example, consider a home thermostat. A simple thermostat can achieve the goal of thermal comfort with minimum energy use, with user interventions to change the set-point when residents leave or return to home, sleep, etc. A smart thermostat could achieve the goal without this level of human intervention, relying on occupancy sensors, models of thermal comfort at night versus daytime, etc. The smart thermostat has higher level of autonomy as it can achieve the goal with lower user intervention.

In certain cases, human intervention is needed to take care of a subset of the existing disturbances thus eliminating them from the set of uncertainties the system needs to cope with. Such intervention allows the system to attain the goals autonomously, under the now reduced set of disturbances. For example, the cruise control in a car that maintains the car's speed constant may be not be able to perform if the road incline is too steep. The driver may intervene using say look ahead control policies to cope with these large size uncertainties of the road incline and so reducing the set of incline uncertainties the system needs to deal with autonomously.

7. Performance and robustness

Performance may be taken to be a measure of the set of goals G achieved by the system. A performance level is assigned that captures the number of goals, their difficulty and importance. It should be noted that the term Performance here has a more general meaning than in the Controls literature, where typically it does not include stability. A level of Performance is accomplished under a level of *Robustness* which corresponds to the level of uncertainty under which the goals are achieved. For fixed performance level, higher level of robustness implies higher autonomy. Also, for fixed robustness level, higher level of performance implies higher autonomy level. For fixed autonomy level, higher performance leads to lower robustness and higher robustness leads to lower performance.

Level of autonomy $L =$ {Performance} × {Robustness}

Performance P is a particular measure of the set of goals G, M_G. Robustness R is a particular measure of the set of uncertainties U, M_U. For $P = M_G$ and $R = M_U$.

$$L = P \times R$$

For fixed level of autonomy L when Performance increases Robustness must be reduced. This brings up interesting issues regarding fundamental limitations.

Robustness R which is a measure of the uncertainties the system can cope with is inversely proportional to M_I the level of needed outside intervention.

$$R \times M_I = 1$$

8. Summary of measures and relations

Let M_G be a measure of the set of goals G. Let M_U be a measure of the set of uncertainties U.

If L is the level of autonomy of the system then

$$L = M_G \times M_U$$

Performance P can be seen as an M_G. Robustness R can be seen as an M_U. Then for $P = M_G$ and $R = M_U$ the above relation becomes

$$L = P \times R$$

Let M_I be a measure of the set of needed interventions I. Then for $M_I \times M_U = 1$ the above relation becomes

$$L = M_G/M_I$$

In view of the relation between measures of Uncertainty and Intervention, namely $M_I \times xM_U = 1$ and the fact that measure of Robustness $R = M_U$ we have that $M_U = R = 1/M_I$, that is the smaller the needed intervention the higher the robustness of the system

9. Entropy

We may use Entropy to compare autonomous systems that achieve the same set of goals. Entropy is a measure of uncertainty. If two systems accomplish the same goals, the system with higher Entropy has a higher level of autonomy since the goals are achieved under greater uncertainties. For the same goals, higher Entropy implies higher levels of autonomy.

We may use Entropy to compare autonomous systems with varying sets of goals. Entropy measures uncertainty. Reduced entropy means reduced uncertainty that implies an increase of the set of goals possible, that is a higher level of autonomy.

As Entropy decreases the set of goals that may be achieved increases. When Entropy is epsilon or zero, a very large set of goals may be accomplished — restricted only by the system's characteristics, its dynamics and structure.

As Entropy increases the set of goals that maybe achieved decreases. When Entropy is very large the set of goals that can be achieved becomes very small — epsilon size or zero.

Let M_G be a measure of the set of goals G and M_U be a measure of the set of uncertainties U.

We have seen that $L = M_G \times M_U$.
Entropy can be taken to be the measure for the uncertainties.
Let Entropy $H = M_U$. Then $L = M_G \times H$.

10. Concluding remarks

General metrics to measure autonomy levels and compare levels of autonomy among different systems were introduced. They are based on a very general definition of autonomy that involves only the set of goals to be achieved under a set of uncertainties. This point of view was developed by the author, simplified and fine-tuned over many years. It is based on research performed during a summer spent at JPL in Pasadena, California on envisioning the capabilities of a spacecraft necessary to act autonomously — in a way bringing Houston Control on board of the spacecraft. Appendix B below describes in detail how higher levels of autonomy may be accomplished. It is quite interesting that fault detection and identification and control reconfigurations needed for autonomy are not being addressed in current efforts towards autonomous cars. Autonomy is a very exciting topic with many challenges. I am glad to see that the big high-tech companies working on autonomous vehicles have started moving beyond the original hype where full autonomy was to be implemented in our cars within days, or weeks at most! There are many things to improve and more things to invent in this area. The challenges are serious, but the potential payoff makes it all worth it. The Quest for Autonomy continues at an ever increasing pace.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Autonomy in vehicles

There is a SAE scale of 5 levels (plus a zero level) with level 5 used for full autonomy. Consider the 0 to 5 levels of autonomy in vehicles. We have:

$$M_G \times M_U = L$$

and

$$M_G / M_I = L$$

Let the scale for M_G be 0–10 and the scale for M_U be 0–10. Then the range of level of autonomy will be 0–100.

Assume that the goals are the same - to drive at the level of a human driver under any normal road conditions. We shall take M_G to be equal to 10 across all levels.

At level 5 we calibrate M_U to be 10, which implies that all goals are achieved under maximum uncertainties. At level 0 we calibrate M_U to be 0.1, which implies that all goals are achieved under minimum or no uncertainties.

In summary

- At level 0, $M_U = 0.1$
- At level 1, $M_U = 2$
- At level 2, $M_U = 4$
- At level 3, $M_U = 6$
- At level 4, $M_U = 8$
- At level 5, $M_U = 10$

The levels of autonomy then will be

- At level 0, $\{M_G = 10\} \times \{M_U = 0.1\} = 1 = L$ the level of autonomy.
- At level 1, $\{M_G = 10\} \times \{M_U = 2\} = 20 = L$
- At level 2, $\{M_G = 10\} \times \{M_U = 4\} = 40 = L$
- At level 3, $\{M_G = 10\} \times \{M_U = 6\} = 60 = L$
- At level 4, $\{M_G = 10\} \times \{M_U = 8\} = 80 = L$
- At level 5, $\{M_G = 10\} \times \{M_U = 10\} = 100 = L$

Instead of Uncertainty consider now a measure of required Intervention for the goals to be achieved.

Let the scale for M_G be 0–10 and the scale for M_I be 0–10. Then the range of level of autonomy will be 0–100.

At level 5 we calibrate M_I to be $1/10 = 0.1$, which implies that all goals are achieved under minimum or no intervention. At level 0 we calibrate M_I to be $1/0.1 = 10$, which implies that all goals are achieved only under maximum intervention.

In summary

- At level 0, $M_I = 1/0.1 = 10$
- At level 1, $M_I = 1/2$
- At level 2, $M_I = 1/4$
- At level 3, $M_I = 1/6$
- At level 4, $M_I = 1/8$
- At level 5, $M_I = 1/10 = 0.1$

The levels of autonomy then will be

- At level 0, $\{M_G = 10\}/\{M_I = 1/0.1 = 10\} = 1 = L$ the level of autonomy.
- At level 1, $\{M_G = 10\}/\{M_I = 1/2\} = 20 = L$
- At level 2, $\{M_G = 10\}/\{M_I = 1/4\} = 40 = L$
- At level 3, $\{M_G = 10\}/\{M_I = 1/6\} = 60 = L$
- At level 4, $\{M_G = 10\}/\{M_I = 1/8\} = 80 = L$
- At level 5, $\{M_G = 10\}/\{M_I = 1/10 = 0.1\} = 100 = L$

We could have taken

- At level 0, $M_I = 1/0.1 = 10$
- At level 1, $M_I = 8$
- At level 2, $M_I = 6$
- At level 3, $M_I = 4$
- At level 4, $M_I = 2$
- At level 5, $M_I = 1/10 = 0.1$

In that case the levels of autonomy then will be

- At level 0, $\{M_G = 10\}/\{M_I = 1/0.1 = 10\} = 1 = L$ the level of autonomy.
- At level 1, $\{M_G = 10\}/\{M_I = 8\} = 10/8 = L$
- At level 2, $\{M_G = 10\}/\{M_I = 6\} = 10/6 = L$
- At level 3, $\{M_G = 10\}/\{M_I = 4\} = 10/4 = L$
- At level 4, $\{M_G = 10\}/\{M_I = 2\} = 10/2 = L$
- At level 5, $\{M_G = 10\}/\{M_I = 1/10 = 0.1\} = 100 = L$

The constant then takes on different values from the case when Uncertainties are considered.

Appendix B. Autonomous spacecraft

In this Appendix a conceptual case study is described that identifies the capabilities of a spacecraft necessary to exhibit autonomous behavior. It is pointed out that control theory is a cornerstone of autonomy in systems. The level of autonomy is increased by adding functions such as learning, planning, failure diagnosis and reconfiguration. The description below follow Antsaklis (2011). Additional details maybe found in Antsaklis (2011) and Antsaklis et al. (1989, 1991) and the references therein.

We begin by describing a conceptual functional architecture of the autonomous controller necessary for the operation of future advanced space vehicles that was developed in

Antsaklis et al. (1988), Antsaklis et al. (1989), and Antsaklis and Passino (1993); Antsaklis, Passino, and Wang (1991). This hierarchical architecture is certainly one of many possible control architectures. The choice is dependent on the particular problem addressed. We refer to it as a hierarchical functional architecture - hierarchies make it possible for us to handle complexity better - but the architecture in fact is a *heterarchy,* as it also allows direct communication among elements on the same level.

The concepts and methods needed to design successfully such an autonomous controller are introduced and discussed. A hierarchical functional autonomous controller architecture for a future spacecraft is described; it is designed to ensure the autonomous operation of the control system and it allows interaction with the pilot/ground station and the systems on board the autonomous vehicle. A command by the pilot or the ground station is executed by dividing it into appropriate subtasks, which are then performed by the controller. The controller can deal with unexpected situations, new control tasks, and failures within limits. To achieve this, high-level decision-making techniques for reasoning under uncertainty and taking actions must be utilized. These techniques, if used by humans, are attributed to *intelligent* behavior. Hence, one way to achieve autonomy, in some applications, is to utilize high-level decision-making techniques, "intelligent" methods, in the autonomous controller. Remember that *autonomy is the objective, and "intelligent"* or "smart" *controllers are one way to achieve it.*

B1. Autonomous controller functions

Autonomous control systems must perform well under significant uncertainties in the plant and the environment for extended periods of time and they must be able to compensate for system failures without external intervention. Such autonomous behavior is a very desirable characteristic of advanced systems. An autonomous controller provides high level *adaptation* to changes in the plant and environment. To achieve autonomy the methods used for control system design should utilize both

(a) algorithmic-numeric methods, based on the state-of-the-art conventional control, identification, estimation, and communication theory, together with advanced sensors and actuators and
(b) decision making-symbolic methods, such as the ones developed in computer science (e.g., automata theory), and specifically in the field of AI.

In addition to supervising and tuning the control algorithms, the autonomous controller must also provide a high degree of tolerance to failures. To ensure system reliability, failures must first be detected, isolated, and identified (and if possible contained), and subsequently a new control law must be designed if it is deemed necessary.

The autonomous controller must be capable of planning the necessary sequence of control actions to be taken to accomplish a complicated task.

It must be able to interface to other systems as well as with the operator, and it may need learning capabilities to enhance its performance while in operation. It is for these reasons that advanced planning and learning, among others, must work together with conventional control systems in order to achieve autonomy.

The need for quantitative methods to model and analyze the dynamical behavior of such autonomous systems presents significant challenges. The development of autonomous controllers requires significant interdisciplinary research effort as it integrates concepts and methods from areas such as control, identification, estimation, and communication theory, computer science, artificial intelligence, and operations research.

Autonomous controllers evolve from existing controllers in a natural way fueled by actual needs, as is now discussed.

B2. Design methodology — History

Conventional control systems are designed using mathematical models of physical systems. A mathematical model, which captures the dynamical behavior of interest is chosen and then control design techniques are applied, aided by software packages, to design the mathematical model of an appropriate controller. The controller is then realized via hardware or software and it is used to control the physical system. The procedure may take several iterations. The mathematical model of the system must be "simple enough" so that it can be analyzed with available mathematical techniques, and "accurate enough" to describe the important aspects of the relevant dynamical behavior. It approximates the behavior of a plant in the neighborhood of an operating point or a region. The first mathematical model to describe plant behavior for control purposes is attributed to J.C. Maxwell, who in 1868 used differential equations to explain instability problems encountered with James Watt's flyball governor; the governor was introduced in 1769 to regulate the speed of steam engine vehicles (the first feedback control mechanism in the historical record is the water clock of Ktesibios, 3rd century BC).

Control theory made significant strides in the past 140 years, with the use of frequency domain methods and Laplace transforms in the 1930s and 1940s and the introduction of the state space analysis in the 1960s. Optimal control in the 1950s and 1960s, stochastic, robust and adaptive control methods in the 1960s to today, have made it possible to control more accurately, significantly more complex dynamical systems than the original flyball governor. The control methods and the underlying mathematical theory were developed to meet the ever-increasing control needs of our technology. The evolution in the control area is fueled by three major needs:

(a) The need to deal with increasingly complex dynamical systems.
(b) The need to accomplish increasingly demanding design requirements.
(c) The need to attain these design requirements with less precise advanced knowledge of the plant and its environment, that is, the need to control under increased uncertainty.

The need to achieve the demanding control specifications for increasingly complex dynamical systems has been addressed by using more complex mathematical models such as nonlinear and stochastic ones, and by developing more sophisticated design algorithms for, say, optimal control. The use of highly complex mathematical models however, can seriously inhibit our ability to develop control algorithms. Fortunately, simpler plant models, for example linear models, can be used in the control design; this is possible because of the feedback used in control, which can tolerate significant model uncertainties. Controllers can then be designed to meet the specifications around an operating point, where the linear model is valid and then via a scheduler a controller emerges which can accomplish the control objectives over the whole operating range. This is, for example, the method typically used for aircraft flight control. *In autonomous control systems we need to significantly increase the operating range;* we must be able to deal effectively with significant uncertainties in models of increasingly complex dynamical systems in addition to increasing the validity range of our control methods. This will involve the use of intelligent decision-making processes to generate control actions so that a performance level is maintained even though there are drastic changes in the operating conditions.

There are needs today that cannot be successfully addressed with the existing conventional control theory. They mainly pertain to the area of uncertainty. Heuristic methods may be needed to tune the parameters of an adaptive control law. New control laws to perform novel control functions should be designed while the system is in operation. Learning from past experience and planning control actions may be necessary. Failure detection and identification is needed. Many of these functions have been performed in the past by human operators. *To increase the speed of response, to relieve the pilot from mundane tasks, to protect operators from hazards, autonomy is desired.* It should be pointed out that several functions proposed in later sections, to be part of the autonomous controller, have been performed in the past by separate systems; examples include fault trees in chemical process control for failure diagnosis and hazard analysis, and control reconfiguration systems in aircrafts, planning the sequence of order execution in steel mills and setting control set-points.

In the next section the functions, characteristics, and benefits of autonomous control are outlined. Next it is explained that plant complexity and design requirements dictate how sophisticated a controller must be. From this it can be seen that often it is appropriate to use methods from operations research and computer science to achieve autonomy. An autonomous control functional architecture for future space vehicles is then presented, which incorporates the concepts and characteristics described earlier. *The controller is hierarchical, with three levels, the execution level (lowest level), the coordination level (middle level), and the management and organization level (highest level).* The general characteristics of the overall architecture, including those of the three levels are explained, and an example to illustrate their functions is given. In the following section the fundamental issues and attributes of intelligent autonomous systems are described. Then we discuss mathematical models for autonomous systems including "logical" discrete event system models. A "hybrid' approach that includes both conventional analysis techniques based on difference and differential equations, together with new techniques for the analysis of systems described with a symbolic formalism such as finite automata appears to offer advantages.

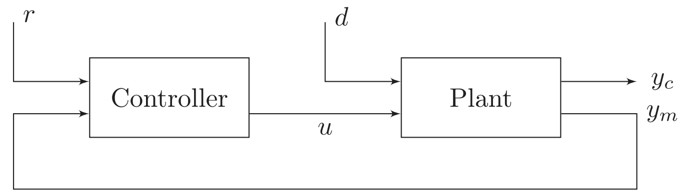

Fig. B.1. Conventional Fixed Controller for Robust Control.

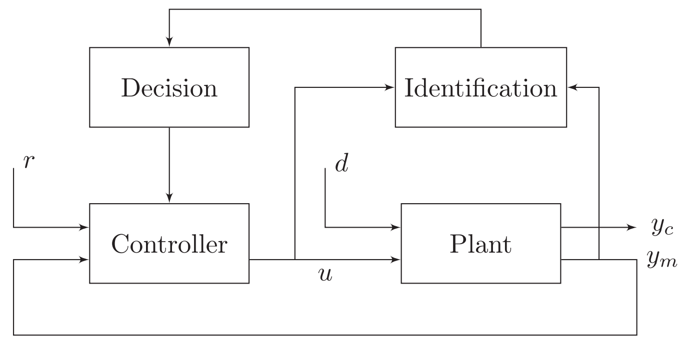

Fig. B.2. Conventional Indirect Adaptive Controller.

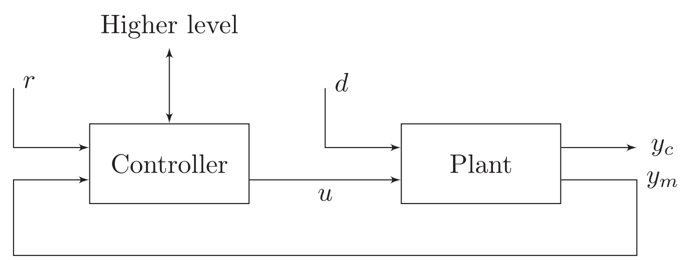

Fig. B.3. Highly Adaptive Controller for Autonomous Control.

B3. Functional architecture of an autonomous controller

Intelligent Autonomous Control Motivation: Sophistication and Complexity in Control. The complexity of a dynamical system model and the increasingly demanding closed loop system performance requirements, necessitate the use of more complex and sophisticated controllers. For example, highly nonlinear systems normally require the use of more complex controllers than low order linear ones when goals beyond stability are to be met. The increase in uncertainty, which corresponds to the decrease in how well the problem is structured or how well the control problem is formulated, and the necessity to allow human intervention in control, also necessitate the use of increasingly sophisticated controllers. *Controller complexity and sophistication is then directly proportional to both the complexities of the plant model and of the control design requirements.*

Based on these ideas, Saridis in Saridis (1979, 1989) suggested a hierarchical ranking of increasing controller sophistication on the path to *intelligent* controls. At the lowest level, deterministic feedback control based on conventional control theory is utilized for simple linear plants. As plant complexity increases, such controllers will need for instance, state estimators. When process noise is significant, Kalman or other filters may be needed. Also, if it is required to complete a control task in minimum time or with minimum energy, optimal control techniques are utilized. When there are many quantifiable, stochastic characteristics in the plant, stochastic control theory is used. If there are significant variations of plant parameters, to the extent that linear robust control theory is inappropriate, adaptive control techniques are employed. For still more complex plants, self-organizing or learning control may be necessary. At the highest level in their hierarchical ranking, plant complexity is so high, and performance specifications so demanding, that intelligent control techniques are used. In the hierarchical ranking of increasingly sophisticated controllers described above, the decision to choose more sophisticated control techniques is made by studying the control problem using a controller of a certain complexity belonging to a certain class. When it is determined that the class of controllers being studied (e.g., adaptive controllers) is inadequate to meet the required objectives, a more sophisticated class of controllers (e.g., intelligent controllers) is chosen. That is, if it is found that certain higher-level decision-making processes are needed for the adaptive controller to meet the performance requirements, then these processes can be incorporated. These intelligent autonomous controllers are the next level up in sophistication. They are *enhanced adaptive controllers,* in the sense that they can adapt to more significant global changes in the plant and its environment than conventional adaptive controllers, while meeting more stringent performance requirements. *One turns to more sophisticated controllers only if simpler ones cannot meet the required objectives. The need to use intelligent autonomous control stems from the need for an increased level of autonomous decision-making abilities in achieving complex control tasks.*

A brief literature overview of the early literature on autonomous intelligent control may be found in Antsaklis (1999); Antsaklis et al. (1989, 1991). The architecture in Fig. B.4 has three levels. At the lowest level, the execution level, there is the interface

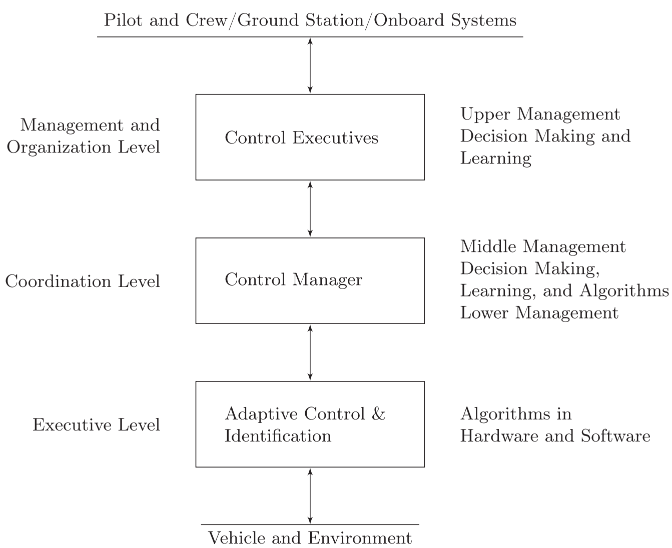

Fig. B.4. Autonomous Controller Functional Architecture — Spacecraft JPL.

to the vehicle and its environment via the sensors and actuators. At the highest level, the management and organization level, there is the interface to the pilot and crew, ground station, or onboard systems. The middle level, called the coordination level, provides the link between the execution level and the management level. Note that we follow the somewhat standard viewpoint that there are three major levels in the hierarchy. *It must be stressed that the system may have more or fewer than three levels. Some characteristics of the system, which dictate the number of levels, are the extent to which the operator can intervene in the system's operations, the degree of autonomy or level of intelligence in the various subsystems, the dexterity of the subsystems, the hierarchical Characteristics of the plant. Note however that the three levels shown here in Fig. B.4 are applicable to most architectures of autonomous controllers, by grouping together sublevels of the architecture if necessary; the levels are the lower execution level, the higher management level with everything else in between being included in the mid coordination level.* Notice that as it is indicated in the figure, the lowest, execution level involves conventional control algorithms, while the highest, management and organization level involves only higher-level decision-making methods. The middle, coordination level is the level, which provides the interface between the actions of the other two levels and it uses a combination of conventional and intelligent decision-making methods. The sensors and actuators are implemented mainly with hardware. They are the connection between the physical system and the controller. Software and perhaps hardware are used to implement the execution level. Mainly software is used for both the coordination and management levels. There are multiple copies of the control functions at each level, more at the lower and fewer at the higher levels. For example, there may be one control manager, which directs a number of different adaptive control algorithms to control the flexible modes of the vehicle via appropriate sensors and actuators. Another control manager is responsible for the control functions of a robot arm for satellite repair. The control executive issues commands to the managers and coordinates their actions. Note that the autonomous controller is only one of the autonomous systems on the vehicle. It is responsible for all the functions related to the control of the physical system and allows for continuous online development of the autonomous controller and to provide for various phases of mission operations. The tier structure of the architecture allows us to build on existing advanced control theory. Development progresses, creating each time, higher level adaptation and a new system, which can be operated and tested independently. The autonomous controller performs many of the functions currently performed by the pilot, crew, or ground station. The pilot and crew are thus relieved from mundane tasks and some of the ground station functions are brought aboard the vehicle. In this way the degree of autonomy of the vehicle is increased.

Functional Operation: Commands are issued by higher levels to lower levels and response data flows from lower levels upwards. Parameters of subsystems can be altered by systems one level above them in the hierarchy. There is a delegation and distribution of tasks from higher to lower levels and a layered distribution of decision-making authority. At each level, some preprocessing occurs before information is sent to higher levels. If requested, data can be passed from the lowest subsystem to the highest, e.g., for display. All subsystems provide status and health information to higher levels. Human intervention is allowed even at the control implementation supervisor level, with the commands however passed down from the upper levels of the hierarchy.

The specific functions at each level are described in detail in Antsaklis et al. (1989) and Antsaklis and Passino (1993). Here we present a simple illustrative example to clarify the overall operation of the autonomous controller. Suppose that the pilot desires to repair a satellite. After dialogue with the control executive, the task is refined to "repair satellite using robot A'. This is arrived at using the capability assessing, performance monitoring, and planning functions of the control executive. The control executive decides if the repair is possible under the current performance level of the system, and in view of near term planned functions. The control executive, using its planning capabilities, sends a sequence of subtasks, sufficient to achieve the repair, to the control manager. This sequence could be to order robot A to: "go to satellite at coordinates *xyz*", "open repair hatch", "repair". The control manager, using its planner, divides say the first subtask, "go to satellite at coordinates *xyz*", into smaller subtasks: "go from start to $x_1y_1z_1$, " then "maneuver around obstacle," "move to $x_2y_2z_2$."..."arrive at the repair site and wait." The other subtasks are divided in a similar manner. This information is passed to the control implementation supervisor, which recognizes the task, and uses stored control laws to accomplish the objective. The subtask "go from start to $x_1y_1z_1$" can for example, be implemented using stored control algorithms to first, proceed forward 10m, to the right15//, etc. These control algorithms are executed in the controller at the execution level utilizing sensor information; the control actions are implemented via the actuators.

Some Design Guidelines for Autonomous Controllers: There are certain functions, characteristics, and behaviors that autonomous systems should possess. These are outlined below. Some of the important characteristics of autonomous controllers are that they relieve humans from time consuming mundane tasks thus increasing efficiency, enhance reliability since they monitor health of the system, enhance performance, protect the system from internally induced faults, and they have consistent performance in accomplishing complex tasks. There are autonomy guidelines and goals that should be followed and sought after in the development of an autonomous system. Autonomy should reduce the work-load requirements of the operator or, in the space vehicle case discussed here, of the pilot/crew/ground station, for the performance of routine functions, since the gains due to autonomy would be superficial if the maintenance and operation of the autonomous controller taxed the operators. Autonomy should enhance the functional capability of the system. Since the autonomous controller will be performing the simpler routine tasks, persons will be able to dedicate themselves to even more complex tasks. There are certain autonomous system architectural characteristics that should be sought after in the design process. The autonomous control architecture should be amenable to evolving future needs and updates in the state of the art. The autonomous control architecture should be functionally hierarchical; for lower level subsystems to take some actions, they have to clear it with a higher-level authority. The system must, however, be able to have lower level subsystems, that are monitoring and reconfiguring for failures, and act autonomously to certain extent to enhance system safety. There are also certain operational characteristics of autonomous controllers. Human operators should have ultimate supervisory override control of autonomy functions. Autonomous activities should be highly visible, “transparent”, to the operator at the maximum extent possible. Finally, there must be certain features inherent in the autonomous system design. Autonomous design features should prevent failures that would jeopardize the overall system mission goals or safety. These features should enhance safety, and avoid false alarms and unnecessary hardware reconfiguration. This implies that the controller should have self-test capability. Autonomous design features should also be tolerant to transient errors, they should not degrade the reliability or operational lifetime of functional elements, they should include adjustable fault detection thresholds, avoid irreversible state changes, and provide protection from erroneous or invalid external commands.

B4. Characteristics of autonomous control systems

Based on the architecture described above we identify the important fundamental concepts and characteristics that are needed for an autonomous control theory. Note that several of these have been discussed in the literature as outlined above. Here, these characteristics are brought together for completeness. Furthermore, the fundamental issues which must be addressed for a quantitative theory of intelligent autonomous control are introduced and discussed. There is a *successive delegation* of *duties* from the higher to lower levels; consequently, the *number* of *distinct tasks* increases as we go down the hierarchy. Higher levels are concerned with slower aspects of the system’s behavior and with its larger portions, or broader aspects. There is then a *smaller contextual horizon at lower levels,* i.e. the control decisions are made by considering less information. Also notice that higher levels are concerned with *longer time horizons* than lower levels. Due to the fact that there is the need for high-level decision-making abilities at the higher levels in the hierarchy, there is *increasing intelligence* as one moves from the lower to the higher levels. This is reflected in the use of fewer conventional numeric-algorithmic methods at higher levels as well as the use of more symbolic-decision making methods. This is the “principle of increasing intelligence with decreasing precision” described in Saridis (1979, 1989). The decreasing precision is reflected by a decrease in *time scale density,* decrease in *bandwidth* or *system rate,* and a decrease in the *decision (control action) rate.* All these characteristics lead to a decrease in *granularity of models* used, or equivalently, to an *increase in model abstractness.* Model granularity also depends on the *dexterity* of the autonomous controller. The execution level of a highly dexterous controller is very sophisticated and it can accomplish complex control tasks. The control implementation supervisor can issue high level commands to a dexterous controller, or it can completely dictate each command in a less dexterous one. The simplicity, and level of abstractness of macro commands in an autonomous controller depends on its dexterity, which really corresponds to its level of autonomy. The more able the execution level is, the simpler are the commands that the control implementation supervisor needs to issue. Notice that a very dexterous robot arm may itself have a number of autonomous functions. If two such dexterous arms were used to complete a task, which required the coordination of their actions then the arms would be considered to be two dexterous actuators and a new supervisory autonomous controller would be placed on top for the supervision and coordination task. In general, this can happen recursively, adding more intelligent autonomous controllers as the lower level tasks, accomplished by autonomous systems, need to be supervised.

There is an ongoing *evolution* of the intelligent functions of an autonomous controller and this is now discussed. It was pointed out above that complex control problems required a controller sophistication that involved the use of AI methodologies. It is interesting to observe the following: Although there are characteristics, which separate intelligent from non-intelligent systems, as intelligent systems evolve, the distinction becomes less clear. Systems, which were originally considered intelligent evolve to gain more character of what are considered to be non-intelligent, numeric algorithmic systems. An example is a route planner. Although there are AI route planning systems, as problems like route planning become better understood, more conventional numeric-algorithmic solutions are developed. The AI methods which are used in intelligent systems, help us to understand complex problems so we can organize and synthesize new approaches to problem solving, in addition to being problem solving techniques themselves. AI techniques can be viewed as research vehicles for solving very complex problems. As the problem solution develops, purely algorithmic approaches, which have desirable implementation characteristics, substitute AI techniques and play a greater role in the solution of the problem. It is for this reason that we concentrate on achieving autonomy and not on whether the underlying system can be considered “intelligent”.

B5. Mathematical models for autonomous systems

For autonomous control problems, normally the plant is so complex that it is either impossible or inappropriate to describe it with conventional system models such as differential or difference equations. Even though it might be possible to accurately describe some system with highly complex nonlinear differential equations, it may be inappropriate if this description makes subsequent analysis too difficult to be useful. *The complexity of the plant model needed in design depends on both the complexity of the physical system and on how demanding the design specifications are. There is a tradeoff between model complexity and our ability to perform analysis on the system via the model. However, if the control performance specifications are not too demanding, a more abstract, higher level, model can be utilized, which will make subsequent analysis simpler. This model intentionally ignores some of the system characteristics, specifically those that need not be considered in attempting to meet the particular performance specifications.* For example, a simple temperature controller could ignore almost all heat related dynamics of the house or the office and consider only a temperature threshold model of the system to switch the furnace off or on. Logical discrete event system (DES) models and Petri nets are quite useful for modeling the higher-level decision-making processes in the autonomous controller together with logics, semantic networks, rule-based descriptions etc. Queuing network models, Markov chains, etc. will be useful in the study. The choice of whether to use such models will, of course, depend on what properties of the autonomous system need to be studied.

The quantitative, systematic techniques for modeling, analysis, and design of control systems are of central and utmost practical importance in conventional control theory. Similar techniques for autonomous controllers do not exist to a similar degree. This is of course because of their novelty, but for the most part, it is due to the *“hybrid”* structure (nonuniform, nonhomogeneous nature) of the dynamical systems under consideration. The systems are hybrid since in order to examine autonomy issues, a more global, macroscopic view of a dynamical system must be taken than in conventional control theory. Modeling techniques for autonomous systems must be able to support this macroscopic view of the dynamical system, hence it is necessary to represent both numeric and symbolic information. We need modeling methods that can gather all information necessary for analysis and design. For example, we need to model the dynamical system to be controlled (e.g., a space platform), we need models of the failures that might occur in the system, of the conventional adaptive controller, and of the high-level decision-making processes at the management and organization level of the intelligent autonomous controller (e.g., an AI planning system performing actions that were once the responsibility of the ground station). The heterogeneous components of the autonomous controller all take part in the generation of the low-level control inputs to the dynamical system, therefore they all must be considered in a complete analysis. It is our viewpoint that research should begin by using different models for different components of the autonomous controller. Full hybrid models that can represent large portions or even the whole autonomous system should be examined but much can be attained by using the best available models for the various components of the architecture and joining them via some appropriate interconnecting structure. For instance, research in the area of systems that are modeled with a logical DES model at the higher levels and a difference equation at the lower level, that is hybrid dynamical systems, should be used. In any case, our modeling philosophy requires the examination of *hierarchical* models. Much work needs to be done on hierarchical DES modeling, analysis, and design, let alone the full study of hybrid hierarchical dynamical systems. Abstractions are of course at the center of any such study.

References

Antsaklis, P. (2017). Control systems and the quest for autonomy. *Editorial of Editor-in-Chief, IEEE Transactions on Automatic Control, 62*(3), 1013–1016. doi:10.1109/TAC.2017.2658298.

Antsaklis, P., Passino, K., & Wang, S. (1988). Autonomous Control Systems: Architecture and Fundamental Issues. In *1988 American Control Conference (Invited)* (pp. 602–607). Atlanta, GA, USA: IEEE. doi:10.23919/ACC.1988.4789789. See also, P.J. Antsaklis and K.M. Passino, "Autonomous control systems: Architecture and concepts for future space vehicles," Final Rep., Jet Propulsion Laboratory Contract 957856, Oct. 1987.

Antsaklis, P., Vachtsevanos, G. J., & Polycarpou, M. (2005). Is feedback control a mediterranean invention? *Panel Discussion on the Early History of Autonomous Mechanisms* Limassol, Cyprus: in Proceedings 13th Mediterranean conference on Control and Automation and IEEE International Symposium on Intelligent Control.

Antsaklis, P. J. (1998). Setting the Stage: Some Autonomous Thoughts on Autonomy. Introduction to the Panel Discussion: "Autonomy in Engineering Systems: What is it and Why is it Important?". In *Proceedings of the ISIC/CIRA/ISAS'98 A Joint Conference on the Science and Technology of Intelligent Systems*. National Institute of Standards and Technology (NIST), Gaithersburg, MD

Antsaklis, P. J. (1999). Intelligent control. In J. G. Webster (Ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering: 10* (pp. 493–503). Hoboken, NJ, USA: John Wiley & Sons, Inc.. doi:10.1002/047134608X.W1019.

Antsaklis, P. J. (2011). The quest for autonomy revisited. *ISIS Technical Report ISIS-2011-004*. University of Notre Dame.

Antsaklis, P. J. (2018). Autonomy, Levels of Autonomy, Goals, Uncertainties, External Interventions, Entropy, Performance and Robustness. Paper Draft.

Antsaklis, P. J., & Passino, K. M. (1993). Introduction to intelligent control systems with high degrees of autonomy. In P. J. Antsaklis, & K. M. Passino (Eds.), *An Introduction to Intelligent and Autonomous Control* (pp. 1–26). USA: Kluwer Academic Publishers.

Antsaklis, P. J., Passino, K. M., & Wang, S. J. (1989). Towards intelligent autonomous control systems: Architecture and fundamental issues. *Journal of Intelligent and Robotic Systems, 1*(4), 315–342. doi:10.1007/BF00126465.

Antsaklis, P. J., Passino, K. M., & Wang, S. J. (1991). An introduction to autonomous control systems. *IEEE Control Systems Magazine, 11*(4), 5–13. doi:10.1109/37.88585. Reprinted in Neuro-Control Systems: Theory and Applications, M.M. Gupta and D.H. Rao Eds., Chapter 4, Part 1, pp 81-89, IEEE Press 1994.

Antsaklis, P. J., & Rahnama, A. (2018). Control and machine intelligence for system autonomy. *Journal of Intelligent & Robotic Systems, 30th Year Anniversary Special Issue,*, 1–12. doi:10.1007/s10846-018-0832-6.

Antsaklis, P. (Chair). (1994). Defining intelligent control. *Report of the Task Force on Intelligent Control, IEEE Control Systems Magazine*, 4–5. 58–66

Barber, K., & Martin, C. (1999). Agent autonomy: Specification, measurement, and dynamic adjustment. In *Proceedings of the autonomy control software workshop at autonomous agents* (pp. 8–15).

Beer, J. M., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction, 3*(2), 74–99. doi:10.5898/JHRI.3.2.Beer.

Durst, P. J., & Gray, W. (2014). Levels of autonomy and autonomous system performance assessment for intelligent unmanned systems. *Report ERDC/GSL SR-14-1*. Fort Belvoir, VA: Defense Technical Information Center. doi:10.21236/ADA601656.

Hrabia, C.-E., Masuch, N., & Albayrak, S. (2015). A metrics framework for quantifying Autonomy in complex systems. In J. P. Mõller, W. Ketter, G. Kaminka, G. Wagner, & N. Bulling (Eds.), *Multiagent System Technologies*. In *Lecture Notes in Computer Science: 9433* (pp. 22–41). Springer International Publishing. doi:10.1007/978-3-319-27343-3_2.

Huang, H.-M., Pavek, K., Ragon, M., Jones, J., Messina, E., & Albus, J. (2007). Characterizing unmanned system autonomy: Contextual autonomous capability and level of autonomy analyses. In *Proc. SPIE 6561, Unmanned Systems Technology IX, 65611N* (p. 65611N). doi:10.1117/12.719894. Orlando, Florida, USA

Mayor, A. (2018). *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*. Princeton University Press.

NIST (2000). Measuring Performance of Systems with Autonomy: Metrics for Intelligence of Constructed Systems: A White Paper for Explaining Goals of the Workshop (PERMIS).

Saridis, G. (1979). Toward the realization of intelligent controls. *Proceedings of the IEEE, 67*(8), 1115–1133. doi:10.1109/PROC.1979.11407.

Saridis, G. N. (1989). Analytic formulation of the principle of increasing precision with decreasing intelligence for intelligent machines. *Automatica, 25*(3), 461–467. doi:10.1016/0005-1098(89)90016-2.

Valavanis, K. P., Vachtsevanos, G. J., & Antsaklis, P. J. (2007). Technology and autonomous mechanisms in the mediterranean: From ancient Greece to Byzantium. In *2007 European Control Conference (ECC'07)* (pp. 263–270). doi:10.23919/ECC.2007.7068808. Kos, Greece

Valavanis, K. P., Vachtsevanos, G. J., & Antsaklis, P. J. (2014). Technology and autonomous mechanisms in the mediterranean: From ancient Greece to byzantium. *IEEE Control Systems Magazine, 34*(6), 110–119. doi:10.1109/MCS.2014.2359589.

Panos Antsaklis is the H.C. & E.A. Brosey Professor of Electrical Engineering at the University of Notre Dame. He is graduate of the National Technical University of Athens, Greece, and holds MS and PhD degrees from Brown University. His research addresses problems of control and automation and examines ways to design control systems that will exhibit high degree of autonomy. His current research focuses on Cyber-Physical Systems and the interdisciplinary research area of control, computing and communication networks, and on hybrid and discrete event dynamical systems. He is IEEE, IFAC and AAAS Fellow, President of the Mediterranean Control Association, the 2006 recipient of the Engineering Alumni Medal of Brown University and holds an Honorary Doctorate from the University of Lorraine in France. He served as the President of the IEEE Control Systems Society in 1997 and was the Editor-in-Chief of the IEEE Transactions on Automatic Control for 8 years, 2010–2017.