Artif Intell Rev (2019) 51:149–186
https://doi.org/10.1007/s10462-017-9560-8

# Adjustable autonomy: a systematic literature review

**Salama A. Mostafa[1] · Mohd Sharifuddin Ahmad[2] · Aida Mustapha[1]**

Published online: 2 June 2017


**Abstract** Developing autonomous systems that operate successfully in dynamic environments entails many challenges. Researchers introduce the concept of adjustable autonomy to mitigate some of these challenges. Adjustable autonomy enables a system to operate in different autonomic conditions and transfers control between the system's operators. To gauge the extent to which such autonomy has been studied, this paper presents a systematic literature review of adjustable autonomy. It reviews 171 research papers and examines, in detail, 78 research papers. The review provides a fundamental understanding of adjustable autonomy and its application in multi-agent systems. The paper contributes to (1) identifying adjustable autonomy approaches and evaluating their utility, (2) specifying the requirements of formulating adjustable autonomy, (3) presenting adjustable autonomy assessment techniques, and (4) exploring the adjustable autonomy research and identify the research gaps.

**Keywords** Adjustable autonomy · Flexible autonomy · Autonomy assessment · Software agent · Multi-agent system · Systematic literature review (SLR)

## 1 Introduction

Autonomous systems are gradually incorporating flexibility in operational sectors (Ball and Callaghan 2012). Their role is expanding to cover many public and personal applications. Nowadays, the services provided by autonomous systems have become substantial parts of

✉ Salama A. Mostafa
salama@uthm.edu.my

Mohd Sharifuddin Ahmad
sharif@uniten.edu.my

Aida Mustapha
aida@uthm.edu.my

1 Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

2 College of Information Technology, Universiti Tenaga Nasional, Selangor, Malaysia

humans' lives. The next leap of technological advancement in this field, it seems, is the ability of machines to interact with complex environments and autonomously perform tasks with minimum humans' intervention (Truszkowski et al. 2009). Examples of autonomous systems are different types of unmanned system, robots, smart home et cetera.

Developing an autonomous system to do some tasks on behalf of humans are fraught with software and hardware challenges (Pratihar and Jain 2010; Zieba et al. 2011). The challenges result from dealing with the environment's dynamism complexity, high workload and risk assessment measures (Mostafa et al. 2013a; Valero-Gomez et al. 2011). These challenges mandate the need for maintaining human's global control over autonomous systems in order to avoid autonomy surprises (Alan et al. 2014; Mostafa et al. 2013b). Subsequently, researchers propose adjustable autonomy to manage systems' autonomy and maintain human's global control over the systems' autonomous behaviors (De Visser et al. 2008). Another critical challenge is to produce a deliberative artifact that works in a dynamic environment and performs in a viable manner (Bibel 2010; Ceballos et al. 2011).

Subsequently, autonomy and autonomous systems have aroused the interest of many researchers in software agents and multi-agent systems due to agents' different capabilities, e.g., autonomy, intelligence, situatedness and embodiment (Py et al. 2010; Hexmoor et al. 2012). The multi-agent systems paradigm reinforces the concept of autonomy to be much more significant as autonomy is a core characteristic of software agents (Hexmoor et al. 2012; Codognet 2011; Pătraşcu and Drăgoicea 2014). Adjustable autonomy supports the agents with a graded autonomy properties. These properties manifest flexibility and reliability to the agents' performance (Mostafa et al. 2013b, c).

The scope of this paper is a Systematic Literature Review (SLR) that covers adjustable autonomy and its application in agent-based systems. The study targets understanding the autonomy distribution among agents and humans in systems. The motivation behind the paper is to facilitate the understanding of adjustable autonomy and promote its application technology. The objective of the paper is to review existing adjustable autonomy models' formalization, development, and, assessment. Subsequently, the paper discusses a number of adjustable autonomy research issues and extracts the underlying research questions. Consequently, the paper attempts to answer the following questions using the Brereton et al. (2007) guidelines for a SLR:

1. What are the main requirements of building adjustable autonomy for multi-agent systems?
2. How to assess the ability of adjustable autonomy that satisfies a domain's autonomy requirements of multi-agent systems?
3. What are the shortcoming and possible points of improvement in adjustable autonomy research?

This section introduces the scope, motivation, objectives, and research questions of the paper. Section 2 presents the research motivation of adjustable autonomy. Section 3 presents the research methodology of a systematic literature review to complete this work. Section 4 illustrates the concept of adjustable autonomy. Sections 5, 6 and 7 review the literature on adjustable autonomy approaches formalization and areas of application in agent-based systems. Section 8 presents the various adjustable autonomy assessments methodologies, attributes and techniques. Section 9 presents the milestones of the adjustable autonomy research and possible gaps. Finally, Sect. 10 summarizes the paper's contents and contributions.

## 2 Motivation

Users are intimidated by the ability of autonomous systems to make decisions on their own due to different reasons including the presence of risk and the absence of trust, and confidence (Valero-Gomez et al. 2011; Sadaghiyani 2011). There is a fear that autonomous behavior could wreak havoc and cause harm, fatalities, or catastrophes (Alan et al. 2014; Schurr 2007; Scerri et al. 2009). Subsequently, autonomous systems are still human-dependent. This dependency has a positive impact on systems' performance, especially, in dynamic, complex and interactive environments (Ball and Callaghan 2011; Mostafa et al. 2014b). Humans' contributions in such environments increase the initiative and awareness of the systems (Zieba et al. 2010; Johnson et al. 2012).

In agent-based autonomous systems, giving agents absolute control over their actions is a risky practice (Mostafa et al. 2013b; Sadaghiyani 2011). Obviously, it is impossible for the agents to have a global knowledge about their environments (Fleming and Cohen 2004). The agents decide based on some local states and theoretically, the agents cannot always make optimal decisions, which justify the need for adjustable autonomy. Adjustable autonomy facilitates sharing control among different operators (Bradshaw et al. 2004; Durand et al. 2009). It is widely adopted to manage the autonomy of humans and agents in systems.

While adjustable autonomy is considered as a successful approach, it shows some crucial deficiencies. Human-dependency has a negative impact on systems where many operators are involved in the control (Fleming and Cohen 2004; Mostafa et al. 2015a). The interactions and autonomy adjustment result in continuous interruptions. The interruptions cause disturbances that make the system destabilized, dependent, and slow (Johnson et al. 2012, 2014). The problem is exacerbating when there are communication delays or progressive decisions (Schurr et al. 2009; Rogers et al. 2012). Consequently, in adjustable autonomy, there are still many aspects that need to be further studied and improved (Bradshaw et al. 2003; Zilberstein 2015). Schurr et al. (2008) among others, stated that "adjustable autonomy in teams is an inherently distributed and complex problem that cannot be solved optimally and completely online."

## 3 Research methodology

The research methodology of this work encompasses a Systematic Literature Review (SLR) (Brereton et al. 2007). The SLR entails formulating a number of research questions to be investigated (Jatoth et al. 2015). It explicitly describes the applied methods for reviewing a topic of interest including identifying the relevant materials, exploring the current research trends, presenting the achievements, evaluating the research gaps, and reporting the findings (Calero et al. 2013).

In this paper, we identify, record, classify, synthesize and report adjustable autonomy research by a SLR. Our aim is to gather and transfer the research state to the interested communities. To the best of our knowledge, no SLR has been made on adjustable autonomy. Subsequently, a research methodology that adopts the SLR is not available. The methodology of the SLR is adopted from Brereton et al. (2007) guidelines. The guidelines consist of three phases and their corresponding steps (Brereton et al. 2007). It provides protocols and methods that help on setting the frame and flow of the SLR. It is applied in many SLR studies within the software engineering domain e.g., Calero et al. (2013) and Jatoth et al. (2015). Figure 1 shows the SLR methodology based on Brereton et al. (2007).

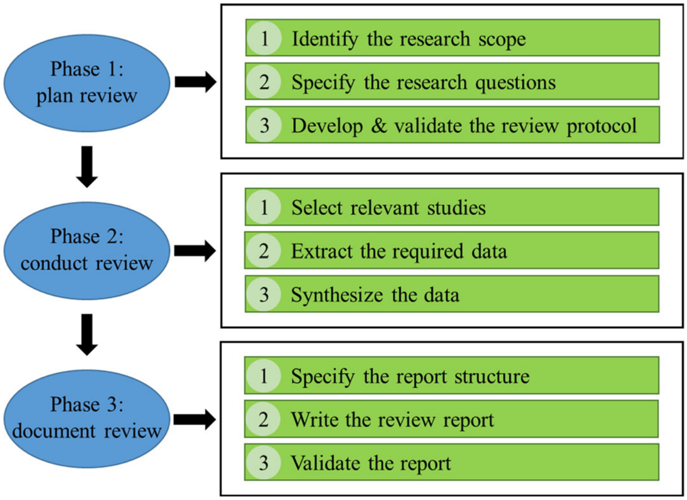


**Fig. 1** The SLR methodology

### 3.1 Plan review

The plan review phase provides a detailed proposal for conducting the review. It consists of three steps that are performed to identify the research scope, specify the research questions and develop the review protocol.

Identifying the research scope is an important step of the SLR. It determines the extent of the relevant research area that the review covers. This step also ensures that the selected scope is not covered in other similar SLRs. Consequently, we search the well-known research scholar digital libraries to verify that there is no SLR that has been made on adjustable autonomy topic. The scope covers the review of the literature in adjustable autonomy and its applications in which humans and agents operate autonomous systems.

Three research questions are specified to identify the research points of interest and extract the critical outcomes. The questions are empirical methods that shape the scope and outcomes of the SLR. Table 1 presents the research questions and their corresponding objectives.

The review protocol defines the review plan including the review process, conditions and quality measures of a conducted SLR study. It determines search strategies that refine the

**Table 1** The research questions and objectives

| Question (Q) | Objective (O) |
|---|---|
| Q1: What are the main requirements of building adjustable autonomy for multi-agent systems? | O1: To formulate a general framework for building adjustable autonomy of multi-agent systems |
| Q2: How to assess the ability of adjustable autonomy that satisfies a domain's autonomy requirements of multi-agent systems? | O2: To provide a comprehensive view of what have been achieved in adjustable autonomy research, applications and assessment |
| Q3: What are the shortcomings and possible points of improvement in adjustable autonomy research? | O3: To identify the gaps and future directions of adjustable autonomy research |

**Table 2** The elements of the systematic research plan

| Element | Specification |
|---|---|
| Context | A systematic review on adjustable autonomy that provides a global understanding of the topic, presents what have been achieved, and addresses a number of research issues |
| Objectives | The objective of the review is to answer the proposed three research questions that concern with adjustable autonomy research development, assessment and gaps |
| Methods | The methods used in the systematic review process are data source selection, filtering, classification, analysis, synthesis, and reporting |
| Results | The results of the review are general requirements for building adjustable autonomy, examples of adjustable autonomy modeling, application and assessment, and a number of research opportunities for adjustable autonomy |
| Conclusion | A brief illustration of the objectives successful achievement and the review critical outcomes |

review scope, and the data source selection criteria (Calero et al. 2013). Ultimately, the review protocol ensures the consistency and minimize the bias of the study. We implement the review protocol according to the guidelines of Brereton et al. (2007) and the SLR example of Jatoth et al. (2015). Table 2 shows the elements of the proposed systematic review plan. They are further detailed within this section and applied in the coming sections.

## 3.2 Conduct review

The conduct review phase implements the review process. It is represented by three steps that are performed to select the relevant studies, extract the required data and synthesize the data.

### *3.2.1 Initial search*

The initial search is performed to build the review database. The source of the data is a number of digital libraries that are accessed online including Association for *Computing Machinery (ACM), Google Scholar, IEEE Xplore, Science direct, and Springer link*. These libraries provide an access to high-quality technical articles of adjustable autonomy topic. The search query used is *adjustable autonomy*. The overall extracted and viewed papers in the initial search are 360.

### *3.2.2 Intermediate search*

The intermediate search is performed to determine the relevant and irrelevant papers to the conducted review. The conditions for the papers' inclusion of the intermediate search are: (1) the paper must have a key name adjustable autonomy, (2) the publication date of the paper is between 2003 and 2015, (3) the paper has four pages' length or above, (4) the paper has a clear contribution to adjustable autonomy theory or application. The objectives of this search are to extract papers, keywords, and search conditions in order to find relevant papers and filter out the irrelevant papers. The following table summarizes the intermediate search steps.

**Intermediate Search**
*Step1*: Retrieve an article;
*Step2*: Perform a quick study to the article;
*Step3*: Gather relevant data of the article;
Step5: Decide on the article's inclusion or exclusion;
*Step4*: Extract the included article's keywords;
Step5: Filter out the unrelated keywords;

The selected papers in the intermediate search in total are 171 after filtering out the irrelevant papers that do not fulfill the four conditions. The selected keywords in total are 10 and some of the keywords have a number of synonyms. The keywords selection is based on their relevance to the research scope and frequency of occurrence. We configure the keywords and their synonyms in the intermediate search as follows:

*adjustable autonomy, flexible autonomy, adaptive autonomy, mixed initiative, autonomous agent, multi-agent system, teamwork, interaction, autonomous system, assessment*

### *3.2.3 Final search*

The final search is performed to determine the most relevant papers for the review. The search condition is to exceed the keywords matching threshold. The following formula is used to set the bias for the papers' inclusion or exclusion threshold.

$$R^y = \frac{\sum_i^n \frac{k_i}{n}}{p^y} \tag{1}$$

where $R$ is the papers' relevance ratio of a particular year, $y$, $k$ is a number of the matched keywords, $n$ is the total number of the proposed keywords, $p$ is the number of the initial papers in a particular year.

Then, formula 2 determines the included and excluded papers.

$$f(k_i) = \begin{cases} included, & R^y < \frac{k_i}{p^y} \\ excluded, & otherwise \end{cases} \tag{2}$$

The final number of selected papers is 87 and the number of excluded papers in total is 84. The total number of used references is 99. It includes the final number of selected papers for review, 87; the papers of the key contributors, 7; and 5 additional necessary references. Table 3 shows the relevance of the papers according to the year of publication.

Figure 2 shows the adjustable autonomy research relevance against the publication years which covers the years between 2003 and 2015. The highest relevance scores for the review topic are in 2008 and 2014 and the lowest is in 2005.

The contents of the required data are extracted from the selected papers and classified into a number of sections that are associated with the research objectives. A detailed examination is performed on the data which results in a number of outcomes. The synthesis of the outcomes represents draft solutions to Q1 and Q2 which include adjustable autonomy approaches, requirements applications, and assessment. After several rounds of analysis on a number of proposed adjustable autonomy models, we are able to identify five

**Table 3** The relevance measure of the papers

| Year | Initial | *R* | Excluded | Included |
|---|---|---|---|---|
| 2003 | 7 | 0.54 | 4 | 3 |
| 2004 | 9 | 0.46 | 4 | 5 |
| 2005 | 5 | 0.58 | 3 | 2 |
| 2006 | 14 | 0.52 | 7 | 7 |
| 2007 | 8 | 0.42 | 5 | 3 |
| 2008 | 16 | 0.61 | 6 | 10 |
| 2009 | 11 | 0.55 | 4 | 7 |
| 2010 | 15 | 0.47 | 7 | 8 |
| 2011 | 17 | 0.39 | 11 | 6 |
| 2012 | 19 | 0.71 | 10 | 9 |
| 2013 | 16 | 0.60 | 7 | 9 |
| 2014 | 23 | 0.57 | 10 | 13 |
| 2015 | 11 | 0.52 | 6 | 5 |

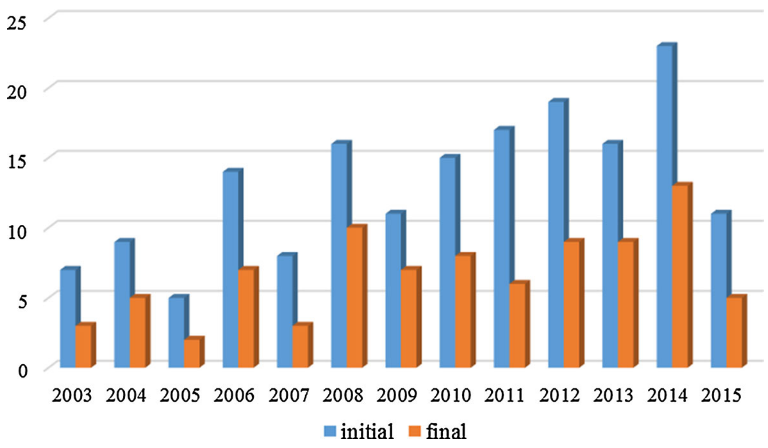


**Fig. 2** Papers' distribution according to the publication years

research gaps. The gaps are illustrated in Sect. 9 and hence, providing potential solutions to Q3.

### 3.3 Document review

The review is documented in four sections. The first section introduces a general description of the study. The second section summarizes the adjustable autonomy's reported data. The report reflects the literature solutions to the review questions. The validation of the report is considered only if the research works have common and clear directions, otherwise, they are excluded. The validation improves the report structure and shapes its flow.

## 4 The concept of adjustable autonomy

The abstract view of a system's autonomy is the ability of the system to make unaided decisions. The literature proposes two main concepts of autonomy, which are absolute autonomy and adjustable autonomy. Absolute autonomy is a general implicit concept of autonomous agents, whereby autonomy is considered as an emergent property of an agent (Dallaire et al. 2014). Therefore, an agent's internal state has an absolute control over the agent's behavior (Wilson et al. 2014). The agent autonomously generates its goals and decides on its actions. Particularly, the agent's proactivity and reactivity conditions to perform particular actions determine its autonomy. The agent's designers setup its autonomy during the design stage and the agent implicitly comply with its desires (Pătraşcu and Drăgoicea 2014; Magill and Erden 2012). The dependence or external controls that the agent encounters is considered as complementary to its interactions (Myers and Morley 2003; Ermon et al. 2012). Hence, there is no explicit representation, formulation, measurement, distribution, and adjustment to autonomy (Pătraşcu and Drăgoicea 2014; Luck et al. 2003).

The main reasons behind the adaptation of the absolute autonomy approach are: (1) there is no essential need to formulate an autonomy model, and (2) to avoid the complications of using an explicit model to autonomy. A key aspect of advanced autonomous systems is their ability to concurrently communicate and cooperate with each other in order to fulfill different situations' constraints (Jennings et al. 2014; Vecht et al. 2008). However, autonomous systems' operators of humans and agents operate at different levels of ability, intelligence, and authority which mandates adjustable autonomy (Birk and Pfingsthorn 2006; Cote et al. 2013; Mostafa et al. 2014a).

The concept of adjustable autonomy is introduced in the late Nineties. In this review, we could not specify who originates the concept as the research focuses on the increasing literature from 2003 to 2015. Nevertheless, the historical character of the paper on this new and strategic domain mandates the acknowledgment of the key contributors to this research domain including Ferguson et al. (1996), Dorais et al. (1999), Musliner and Pell (1999), Falcone and Castelfranchi (1999), Scerri and Reed (2001), Bradshaw et al. (2001), and Goodrich et al. (2001).

Adjustable autonomy or also known as adaptive autonomy is a means to formulate a flexible autonomy to a system's operators (Durand et al. 2009; Bradshaw et al. 2003; Luck et al. 2003). It provides an autonomous system with a variable autonomy in which its operators have the options to work in different autonomy states (Zieba et al. 2010; Scerri and Reed 2001). Adjustable autonomy enables operators to share, oversight, or intervene control in order to avoid impasses (Moffitt et al. 2006; Alzahrani et al. 2013). Some of the commonly accepted definitions of adjustable autonomy or adjustable autonomous system in the literature are:

- Scerri and Reed (2001) define an adjustable autonomous system as "an intelligent system where the distribution of autonomy is changed dynamically to optimize overall system performance."
- Moffitt et al. (2006) define adjustable autonomy as "a mechanism through which an operator delegates authority to the system that can be taken back or shared dynamically throughout mission execution."
- van der Vecht (2009) defines adjustable autonomy as "dynamically dealing with external influences on the decision-making process based on internal motivations."
- Zieba et al. (2010) define adjustable autonomy as "the property of an autonomous system to change its level of autonomy while the system operates. The human operator, another system or the autonomous system itself can adjust the autonomy level."

## 5 The approaches of adjustable autonomy

Different approaches are proposed for adjustable autonomy to resolve some of the autonomy problems. Based on our review, the adjustable autonomy's main approaches are teamwork-centered, supervised, degree-based, level-based, sliding, policy-based and predictive approaches. These approaches mainly focus on plan generation (Mostafa et al. 2013b), agent behavior (Bradshaw et al. 2004), autonomy measurement and adjustment (Rajabzadeh 2014) and human–agent interaction (Petersen and Von Stryk 2011) aspects. The aim of the approaches is to enhance the autonomy distribution dynamism, reduce human workload and decrease system disturbance (Truszkowski et al. 2009; Zieba et al. 2011; Mostafa et al. 2015a; Moffitt et al. 2006). The following illustrates the main approaches to adjustable autonomy.

### 5.1 The teamwork-centred approach

This approach is also known as a mixed-initiatives adjustable autonomy. It structures and draws operators' (agents and humans) autonomy margins in a system to work in a cohesive unit as teamwork (Bradshaw et al. 2003). Basically, it gives the operating control for specific activities of the system to agents and other specific activities to humans so as to work in a collaborative manner (Freedy et al. 2008).

### 5.2 The supervised approach

In this approach, agents' control most of the system's activities but humans monitor and maintain (or 'adjust') some of the system's activities whenever necessary to avoid its failure (Zieba et al. 2011; Valero-Gomez et al. 2011; Bradshaw et al. 2002). The adjustment transfers the decision of critical, uncertain, or unseen situations from the agents to the humans (Johnson et al. 2012; Petersen and Von Stryk 2011). Subsequently, the agents revert to operation after the humans readjust the autonomy of the agents.

### 5.3 The degree-based approach

It is also known as the dialed approach which is one of the simplest approaches to adjustable autonomy (Moffitt et al. 2006). It considers a linear scale autonomy adjustment in which the autonomy transfers between two ends, non-autonomous and fully-autonomous and mixed-autonomy in between (Mostafa et al. 2014a). The autonomy degrees correspond to some autonomy properties that direct agents' behavior as shown in Fig. 3.

The autonomy adjustment can be internally performed by the agents or externally by humans or other systems. It is performed by changing the autonomy degree value based on some autonomy measurement formulas (Roehr et al. 2010; Bush et al. 2012). This approach

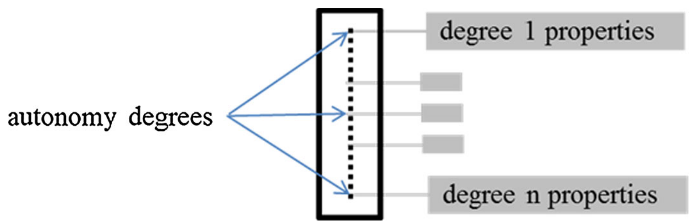


**Fig. 3** A representation of the degree-based adjustable autonomy

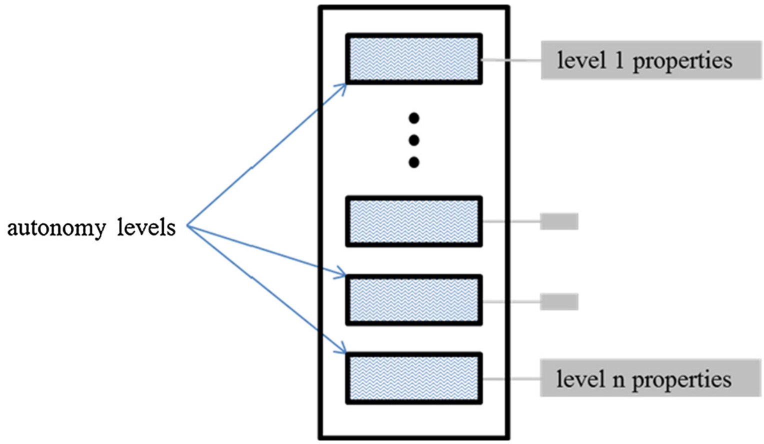


**Fig. 4** A representation of the level-based adjustable autonomy

is considered as an efficient adjustable autonomy as the adjustment reaches the behavior level of the agent. However, it is efficient in systems with a limited number of actions (Moffitt et al. 2006).

### 5.4 The level-based approach

This approach hierarchically manages the autonomy of the agents (Reed 2005). The autonomy levels are distributed to agents based on certain conditions. An autonomy level has specific autonomy properties that direct an agent's authority of operations as shown in Fig. 4. An agent might strictly correspond to an autonomy level (Pătraşcu and Drăgoicea 2014), or can switch between different autonomy levels (Huber 2007). The adjustment implements the operations of adding, suspending, blocking and/or terminating agents or some of their behaviors or goals (Moffitt et al. 2006). Agents that practice unconstraint behaviors and perform different types of actions are said to have high autonomy level and vice versa.

However, this approach lacks the flexibility of autonomy distribution and adjustment. It distributes the autonomy to agents according to a delegated operation or plan (Roehr et al. 2010). The operational autonomy distribution is mainly fixed during the design stage and the autonomy adjustment is mainly performed by humans.

### 5.5 The sliding approach

It is an adjustable autonomy approach that provides a slider as an autonomy distribution and adjustment mechanism (Roehr et al. 2010; Sellner et al. 2006). The slider gives the flexibility to adjust (increase or decrease) the autonomy dimensions or its options by a user (Lin and Goodrich 2015). The adjustment constrains the agents' choice of operations. For instance, Brookshire et al. (2004) proposed a sliding approach to distribute autonomy among agents. The slider is manually adjusted by a human operator to determine the responsible agents for a task achievement. Sellner et al. (2006) went further with sliding autonomy by proposing two authority modes: Mixed-Initiative Sliding Autonomy (MISA) which gives the authority to a human operator to intervene agents and System-Initiative Sliding Autonomy (SISA) which allows agents to decide on seeking assistance from a human. While this approach allows sharing of tasks between system operators, it is insufficient for critical time and dynamic systems (Schurr et al. 2009). The sliding autonomy has the limitations of the level-based approach. In addition, the sliding autonomy application is extensively opened to application domains' constraints (Mercier et al. 2008b).

### 5.6 The policy-based approach

This approach entails a set of policies that the agents of a system submit to Cote et al. (2013). The policies represent principles of action that define agents' behaviors in a system (Beal et al. 2010). The policies allow an agent to perform without explicit human interference. Autonomy adjustment by a human is not primary during runtime, hence, the agent by itself performs autonomy assessment and adjustment which reduce system disturbances and human workload (Alzahrani et al. 2013; Wallace and Henry 2008). The policy-based approach limits agents' actions choice to some permitted actions that are scoped by a system's policies. However, the policies cannot cover a wide range of aspects which make the approach useful for a limited domain of applications. Nevertheless, agents can seek human assistance via interfaces (Myers and Morley 2003). The interface design for such purpose and in dynamic systems is found to be challenging (Moffitt et al. 2006).

### 5.7 The predictive approach

This approach analyzes the performance of a system by studying the behaviors of agents in order to distribute autonomy among them (Bush et al. 2012). Some adjustable autonomy models use predictive displays in a real-time simulator to decide on the autonomy distribution as in the recent National Aeronautics and Space Administration (NASA) Mars projects (Truszkowski et al. 2009). The objective is to adjust a system's autonomy based on its simulation estimated results. The significance of this approach is minimizing system disturbances that result from the autonomy distribution via the prediction mechanisms (e.g., reducing re-planning cost). The disadvantage of the approach is its high cost. It is only applicable and useful for highly uncertain and ambiguous environments.

## 6 Formulating adjustable autonomy

There are many opinions and diverse understanding of what adjustable autonomy is and how it can be efficiently formulated. Formulating adjustable autonomy entails a number of requirements. The most visible requirements for building adjustable autonomy for an individual agent or a multi-agent system are autonomy representation, measurement, distribution, and adjustment. It also includes formulating Human–Agent Interaction (HAI) mechanisms and entails autonomy assessment techniques. We comprehensively discuss each of these in the following subsections. Figure 5 represents an abstract of adjustable autonomy formulation. It summarizes the requirements' structure and sequence of processing.

### 6.1 Autonomy representation

Autonomy representation provides an explicit means for autonomy measurement and manipulation (Luck et al. 2003). It is a primary stage of adjustable autonomy. It entails explicit attributes that enforce adjustable autonomy (Mostafa et al. 2014a). Without such enforcement, an assignment could be subverted by an agent who simply fails or even refuses to carry out the agreed upon tasks (Mercier et al. 2008a). Huber (2007) emphasizes that an agent's autonomy is imposed by external influences, which make it difficult to be defined only in

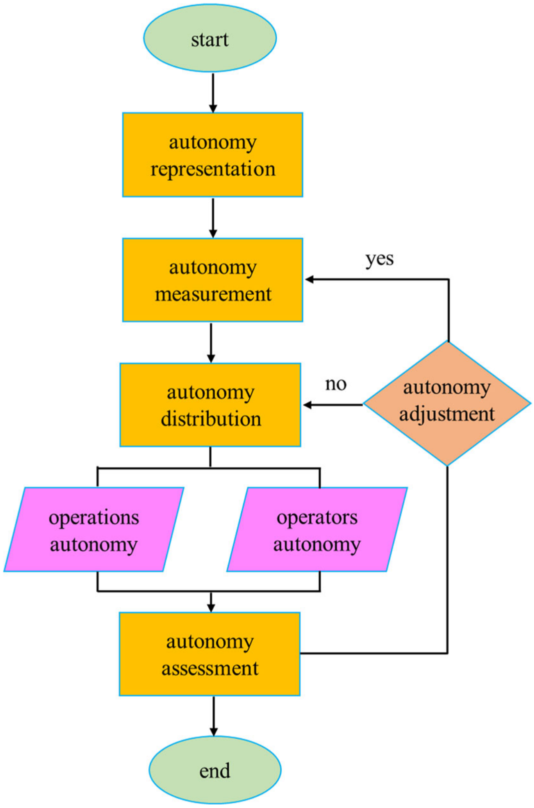


**Fig. 5** Adjustable autonomy formulation flow

terms of a single attribute. Table 4 presents the trade-off attributes/parameters that are used for autonomy representation.

Some of these attributes are treated as indicators to other attributes. For example, *performance* can be derived from *complexity-level*, *time* and *goal-achievement* or authority can be derived from *task-orientation*, *trust* and *confidence* as shown in Fig. 6. The task-orientation attribute might be derived from other sub-attributes (Mostafa et al. 2014c).

Selecting the best collection of autonomy attributes as criteria for autonomy formulation is always a domain specific issue (Mostafa et al. 2013a; Huber 2007). It requires considerable analysis to the application domain requirements for manifesting optimal combination (Mostafa et al. 2014a). The configuration of the attributes can take three forms:

- Derived from the specifications of actions, tasks or goals of a system, which we call an operations autonomy as in Mostafa et al. (2013a) and Brookshire et al. (2004) works.
- Derived from the specifications of a system's operators, which we call an operators autonomy as in Mostafa et al. (2014a) and Huber (2007) work.
- Derived from both parties as in Mostafa et al. (2013b) and Sellner et al. (2006).

The values of the autonomy attributes can be retrospectively taken by analyzing past situations or experiences (Sellner et al. 2006) or can be concurrently configured based on analyzing present situations (Ball and Callaghan 2011). In some advanced models, past and/or present situations determine the values of the autonomy attributes (Mostafa et al. 2014a; Mercier et al. 2008b).

**Table 4** Examples of autonomy attributes

| Attribute | Context | Reference |
|---|---|---|
| *Authority* | The right to make decisions | Bradshaw et al. (2004), Mostafa et al. (2014a), Lin and Goodrich (2015), Mercier et al. (2008b), Mercier et al. (2008a) |
| *Complexity* | The difficulty of handling a problem or completing a task | Brookshire et al. (2004), Mercier et al. (2008a), Mostafa et al. (2014c) |
| *Confidence* | The assurance of the ability to handle a problem or complete a task | Ball and Callaghan (2012), Alzahrani et al. (2013), Scerri and Reed (2001), Sellner et al. (2006) |
| *Consistency* | The ability of continually deliver the same performance | Durand et al. (2009) and Schurr (2007) |
| *Dependency* | The reliance on others when handling a problem or completing a task | Bradshaw et al. (2004), Luck et al. (2003) and Huber (2007) |
| *Goal- achievement* | The independency in a goal generation and achievement | Luck et al. (2003) |
| *Influence* | The effect of others on a decision made or the resistance to the effect (*integrity*) | Huber (2007) |
| *Knowledge* | The mental and physical abilities to handle a problem or complete a task | Bradshaw et al. (2004), Mostafa et al. (2014a) |
| *Motivation* | The act based on inner needs | Luck et al. (2003) and Witkowski and Stathis (2004) |
| *Performance* | The quality of actions in handling a problem or completing a task | Sellner et al. (2006) and Mercier et al. (2008b) |
| *Risk* | The potential failure of handling a problem or completing a task | Hexmoor et al. (2012) and Roehr et al. (2010), Bush et al. (2012) |
| *Task-orientation* | The association of a task through position, location or situation | Mostafa et al. (2014c) |
| *Time* | The process of handling a problem or completing a task is constrained by a time factor | Johnson et al. (2012), Sellner et al. (2006) and Cheng and Cohen (2005) |
| *Trust* | The belief of the ability to handle a problem or complete a task | Hexmoor et al. (2012), Luck et al. (2003) and Roehr et al. (2010) |

### 6.2 Autonomy measurement

Measuring autonomy of an agent or a multi-agent system is a complex process (Hexmoor et al. 2012). It mainly relies on approximations (Rajabzadeh 2014; Roehr et al. 2010). The measurement entails quantifying the autonomy of agents for its future behavior based on the retrospective effect of its past behavior or present situations (Luck et al. 2003; Mostafa et al. 2014a). Nevertheless, quantifying autonomy enhances the performance of the autonomous systems (Hexmoor et al. 2012). It helps in tracking, analyzing and predicting the performance of the agents (Truszkowski et al. 2009). Consequently, autonomy measurement is a primary operation for autonomy distribution and adjustment (Mostafa et al. 2014a).

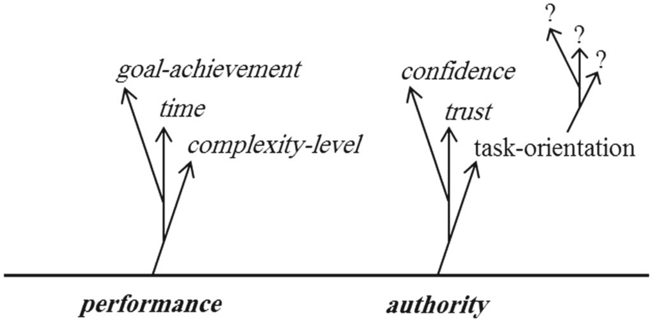


**Fig. 6** A possible derivative of autonomy attributes

Hexmoor et al. (2012) introduce a quantitative measurement approach to different cases of agents' autonomy in a multi-agent environment. The cases they cover are decision and action autonomy; user agent autonomy; and distributing autonomy among a group and for different groups of agents along with the required measurements. They address the difficulty of modeling an algorithm that dynamically matches and deploys an agent or a group of agents with a task to be optimally achieved. Many autonomy measurements are proposed in the literature and we provide insights about some of these in the following.

Let $A$ be a set of agents, $S \subseteq A$ a nonempty subset of agents and $a \in S$ is an agent. Let $P_a^a$ represents an agent $a$ performance according to a performance scale ratio measure when only $a$ is present. Let $P_S^a$ represents the agent $a$ performance when $S$ agents are present; then, according to Hexmoor et al. (2012), the autonomy of the agent $a$ regarding the agent set $S$:

$$A_S^a = \frac{P_S^a}{P_a^a} \tag{3}$$

The autonomy of the agent set $S$ is:

$$A^S = \sum_{a \in S} A_{S\{a\}}^a \tag{4}$$

The autonomy of the agent set $S$ regarding an agent $b$ is:

$$A_b^S = \frac{\sum_{a \in S} A_{(S \cup \{b\}) \setminus \{a\}}^a}{A^S} \tag{5}$$

In the above formulas, agents acquire autonomy based on their performance. An agent that performs a wide range of activities is said to have higher autonomy and vice versa.

Roehr et al. (2010) propose a mixed-initiative sliding autonomy for the humans-robot team. The autonomy distribution and adjustment are performed based on trust and self-confidence which are measured based on human observation of the robots' performance. The computation of trust is confined to prior successes likelihood of tasks in which a task corresponds to a specific autonomy as follows:

$$\text{ɟ}(a) = wP\left(\frac{S}{\mathcal{E}}, a\right) + (1 - w)\, P\,(S) \tag{6}$$

where ɟ is a trust function, $a$ is an autonomy level, $w$ is a weight as min $\left(1, \frac{number\ of\ samples}{minimum\ required\ samples}\right)$, $P\,(S)$ is a prior probability of success and $\mathcal{E}$ is a command and environment parameters.

Consequently, the computation of self-confidence is confined to the probability of successes $P(S)$ given a task as follows:

$$P(S) = \frac{\textit{number of successful task execution}}{\textit{number of task execution}} \tag{7}$$

They assume that the optimum autonomy configuration is reached when the system operates with 10% failure rate.

### 6.3 Autonomy distribution

In adjustable autonomy, an explicit mechanism is needed to demonstrate different kinds of autonomy (Witkowski and Stathis 2004). This mechanism manages the operators' roles of an autonomous system. It concurrently updates the autonomy states of the operators in the system. The autonomy distribution in the literature is performed based on different aspects which include tasks type, teamwork setting, situations type and time factors. Considerable measures for autonomy distribution might include agent authority, task complexity and environment dynamism (Roehr et al. 2010).

Three main types are proposed for autonomy distribution in the literature: (1) the client and the contractor (bilateral) which is a goal-oriented autonomy distribution type; (2) the degree/level of autonomy is either increased or decreased bi-directional which is an operations autonomy distribution and; (3) the multi-dimensional adjustable autonomy which is an action-oriented autonomy distribution type.

- The bilateral adjustable autonomy: It is also called as a dichotomous adjustable autonomy. The bilateral type provides options of two autonomy states: autonomous and non-autonomous (Inyiama et al. 2012). The autonomous state means the system is operated by agents while the non-autonomous state means that the system is operated by a human.
- The bi-directional adjustable autonomy: It is a very common model in formulating adjustable autonomy. The autonomy is scaled between fully- and non-autonomous states (Roehr et al. 2010). There is a mechanism that switches between the autonomy states and distributes the autonomy. A good example of the bi-directional adjustable autonomy is the sliding autonomy approach (Lin and Goodrich 2015).
- The multi-dimension adjustable autonomy: Huber (2007) stated that "autonomy is a complex and multi-faceted concept that cannot be determined by looking solely at a single agent and that is best captured using a multi-dimensional model." Subsequently, researchers such as Mostafa et al. (2013a), Valero-Gomez et al. (2011), Johnson et al. (2012) and Myers and Morley (2003) propose autonomy with multi-dimension. In a multi-dimension adjustable autonomy, agents of a system perform different types of tasks. A task completion entails performing actions. The actions of a task might have different characteristics such as primitive (low-level), deducible (intermediate-level) or critical (high-level) actions (Schermerhorn and Scheutz 2009). These actions are distributed between dimensions of an autonomy spectrum. The agents that execute the actions of a task operate at multi-dimensional autonomy (Mostafa et al. 2014c). Figure 7 shows an example of autonomy distribution according to actions categorization proposed by Bradshaw et al. (2002) and adapted by Moffitt et al. (2006).

The literature offers few autonomy distribution mechanisms, e.g., Mostafa et al. (2013a), Johnson et al. (2012) and Huber (2007). The mechanisms enable agents' and humans' operators to perform actions according to adjustable autonomy. They manage the autonomy states

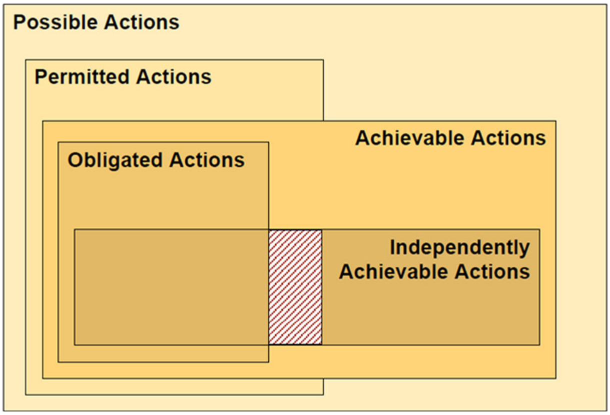


**Fig. 7** Autonomy distribution of actions (Bradshaw et al. 2002)

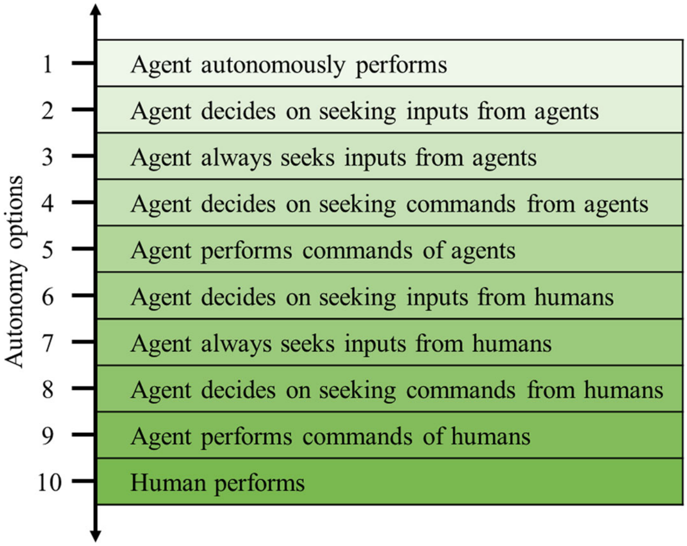


**Fig. 8** Autonomy distribution options of a system

of humans and the agents of systems. At some specific period of time, the operators have specific autonomy states and each autonomy state is vulnerable to change according to the system needs (Bradshaw et al. 2002). The following figure shows ten autonomy state options of a system. The autonomy states represent explicit autonomy degrees or levels. The autonomy state options consider an agent's compliance to other agents and humans.

Figure 8 can be viewed as a spectrum of adjustable autonomy in which the first option represents an absolute- or fully-autonomous state, the tenth option represents a non-autonomous state and in between the two, there is a semi-autonomous region. Figure 9 depicts the spectrum of adjustable autonomy.

Witkowski and Stathis (2004) propose a mechanism that distributes the autonomy to agents based on their motivations. They divide the agents' main operational activities into five categories and each category has a specific autonomy state. An agent motivational state

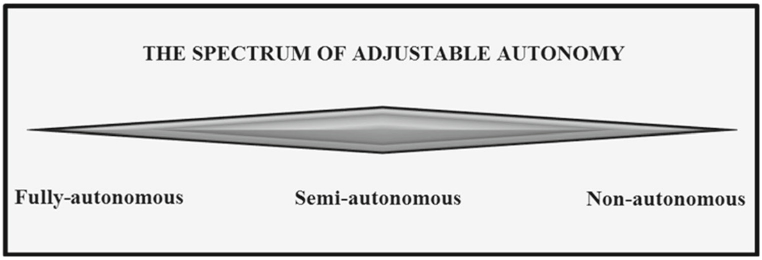


**Fig. 9** The adjustable autonomy spectrum

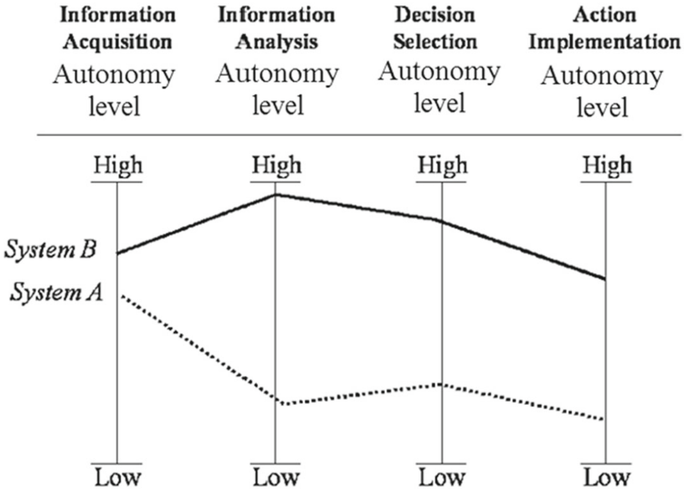


**Fig. 10** Autonomy distribution of systems A and B (Miller and Parasuraman 2007)

determines its autonomy state and accordingly its activities. An agent's motivation is adjusted based on an argumentation process which is directed to achieve a given goal.

Reed (2005) proves, in an aircraft simulation, that autonomy distribution to agents creates fixable autonomy and maintain better results. Figure 10 shows autonomy distribution through the cycle runs of two systems A and B with varying levels of autonomy.

A group of researchers from NASA, Truszkowski et al. (2009), Bradshaw et al. (2004), and Bradshaw et al. (2003) use adjustable autonomy for mixed-initiative adjustable autonomy. They propose a human-centered adjustable autonomy model that integrates the Brahms and KaoS agent frameworks. The model has mechanisms of interaction between human and machine that enables the system to work in a collaborative manner. The model includes scheduled and unscheduled activities, work practices emergency and resource management modules. The aim of their work is to simulate realistic work situations of robots in space.

Cheng and Cohen (2005) propose a utility mechanism to distribute autonomy based on time and cost measurements. The autonomy is divided into three states: fully-autonomous, in which an agent decides on its own; partially-autonomous, in which an agent partially transfers control to humans or other agents to seek feedback to improve its decision; and non-autonomous, in which an agent transfers control to human to make a decision.

Durand et al. (2009) propose a hybrid control architecture of a mobile robot that consists of decision and execution levels. They propose an Inconsistency Detection Mechanism (IDM)

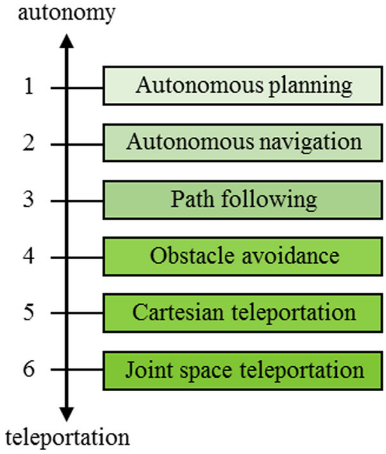


**Fig. 11** The autonomy distribution of a mobile robot (Durand et al. 2009)

to notify a human supervisor to switch between autonomy states. The autonomy is divided into six states and each state corresponds to specific tasks to be autonomously performed by a set of executable functions. The states are autonomous planning, autonomous navigation, path following, obstacle avoidance, cartesian teleportation and joint space teleportation. The human based on the IDM feedback switches between the autonomy states to assist the mobile robot in encountering complex situations. Figure 11 shows the six autonomy states of the proposed mobile robot autonomy distribution architecture.

Tipaldi and Glielmo (2015) propose a layered architecture for autonomy distribution. The architecture divides the autonomy into a number of operational autonomy levels. Markovian decision mechanism distributes the autonomy between the operational levels. The work is intended to be applied to autonomous spacecraft.

### 6.4 Autonomy adjustment

The autonomy adjustment is a process of redistributing the autonomy among system's operators according to some specific causes or reasons (Schurr et al. 2009). The adjustment is directed to change one or more of the following in a system:

- Adjusting authorities: changing the rights of acting permissions.
- Adjusting responsibilities: changing tasks' delegations.
- Adjusting possibilities: changing decisions' options.
- Adjusting capabilities: changing operators' behaviors, skills or resources.

The autonomy adjustment is triggered due to different reasons such as seeking assistance, situation awareness or coordination and cooperation (Johnson et al. 2012; Roehr et al. 2010). The adjustment objective is to prevent the occurrence of error or failure. A primary challenge in adjustable autonomy design is the determination of the need for the autonomy adjustment (when and how). A solution is to formulate a measurement mechanism that determines the autonomy adjustment as shown in Fig. 12.

Amalberti (2006) proposes a mechanism of autonomy adjustment that is efficient in dynamic environments. The mechanism identifies the need for an adjustment when a system encounters unplanned event. It is built based on a transmission time metric in which the adjustment causes the system to transit between two autonomy modes as follows:

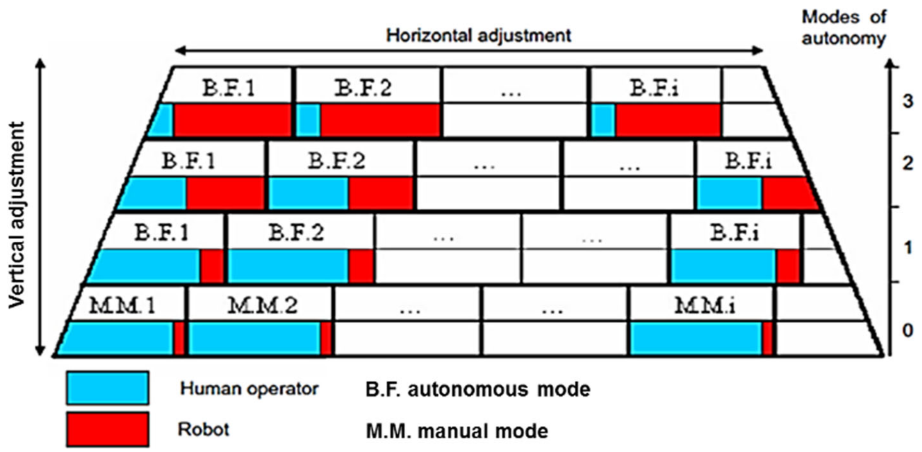


**Fig. 12** The autonomy adjustments in the TAROT project (Zieba et al. 2011)

$$T_{TRANSITION} = T_{DETECTION} + T_{DECISION} + T_{SWITCH} \tag{8}$$

where $T_{TRANSITION}$ is the total transition time, $T_{DETECTION}$ is the adjustment need identification time, $T_{DECISION}$ is the adjustment level or value decision time and $T_{SWITCH}$ is the switching time to the decided upon autonomy mode.

Autonomy adjustment of agents can be internally made based on the agents' desires as an internal adjustment or given by a third party as an external adjustment (Zieba et al. 2011; Luck et al. 2003; Sellner et al. 2006). In the internal or self-adjustment, an agent's internal state is responsible for autonomy adjustment (Huber 2007; Lewis et al. 2013). External adjustment can be performed by a human, e.g., through a Graphical User Interface (GUI), an algorithm or other agents via directly adjusting an agent's autonomy parameters, or indirectly adjusting some activities of a system (Myers and Morley 2003; Alzahrani et al. 2013; Reed 2005).

### *6.4.1 Internal adjustment*

Many researchers argue that autonomy adjustment is an inner process that an agent itself reasons and assigns based on its local observations and not a separate mechanism that is performed by a third party (Vecht et al. 2008; van der Vecht 2009; Witkowski and Stathis 2004). An internal adjustment entails that the agent by itself manages its autonomy (Cheng and Cohen 2006). It implies fully-autonomous agent, since, the agent by itself is controlling its autonomy state. This approach promotes a flexible decision-making capability to agents and reduces human workload (Myers and Morley 2003; Cheng and Cohen 2005). An example of an internal adjustment usefulness is when an agent seeks a human assistance but the human does not respond to the agent due to some reasons, then, the agent should be able to take the initiative and act on his own. The internal autonomy adjustment mechanism can be a separated module or embedded within the agent architecture. It is clearly useful in the case of a learning agent and when the agent adopts new goals (Lewis et al. 2013).

### *6.4.2 External adjustment*

In ambiguous and uncertain circumstances, an agent's behavior needs to be guided to make successful decisions (Zieba et al. 2011; Cote et al. 2013). External adjustment is meant to prevent undesirable decisions that might cause serious unintended consequences, hence, manifesting reliable systems (Schurr et al. 2009; Petersen and Von Stryk 2011). Consequently, the external adjustment implies interventions in agents' decisions and mainly by humans (Ball and Callaghan 2012; Bradshaw et al. 2004). Interventions can be made on the agents directly by adjusting their autonomy states via a GUI or indirectly by adjusting some activities of the system (Mostafa et al. 2013c; Sellner et al. 2006). Some external autonomy adjustment methods are as follows:

1. The external adjuster selects an appropriate autonomy state for an agent. The autonomy state selection implies some tasks to be completed by the agent.
2. The external adjuster modifies the solution that is proposed by some agents.
3. The external adjuster interacts with the agents to cooperatively achieve some tasks.
4. The external adjuster takes full control in achieving a particular task.

### *6.4.3 Flexible adjustment*

The existence of both the internal and external adjustment capabilities in a system manifests flexible autonomy (Alan et al. 2014). Jennings et al. (2014) state that flexible autonomy "allows agents to sometimes take actions in a completely autonomous way without reference to humans, while at other times being guided by much closer human involvement." Flexible adjustment implies interventions in agents' autonomy management. As mentioned before, with a flexible adjustment, the system benefits from both the agents' internal adjustment that provides flexible decisions and reduces human workload and the external adjustment that provides reliable actions via enabling a human global control. However, the main drawback of flexible autonomy is the disturbance that results from the possible conflict between the internal adjuster and the external adjuster which destabilizes the system.

## **6.5 Human–agent interaction**

The principle idea of Human–Agent Interaction (HAI) is that machines' and humans' abilities are complementary and they are likely to provide better performance when joined efficiently than when acting separately (Birk and Pfingsthorn 2006; Petersen and Von Stryk 2011). The decision of the interaction can be made by agents or humans depending on the situation that the system encounters (Sycara and Sukthankar 2006). The interactions have a positive impact by increasing the initiatives and negative impact by exposing the system to disturbance. The literature suggests several scenarios for HAI when agents share control with humans. Figure 13 shows possible configurations of HAI.

The HAI consists of heterogeneous entities for which autonomy can be viewed as the ability of agents to perform independently or with lesser intervention from humans (Vecht et al. 2008; Freedy et al. 2008; Cheng and Cohen 2005). Some researchers suggest that interactions affect agents' autonomy (Hexmoor et al. 2012; Reed 2005). In contrast, others suggest that when an agent interacts with a human, it still autonomously decides on the interaction and the human intervention or assistance does not affect the autonomy of the agent (Fleming and Cohen 2004).

Fleming and Cohen (2004) and Cheng and Cohen (2005) highlight the challenges of formulating initiatives for agents' decisions of interactions. Fleming and Cohen (2004) propose a computational model for autonomous agents to reason about the needs for interaction with humans. The model has a utility function that enables the agents to decide on the interaction based on costs and benefits measures. Cheng and Cohen (2005) also propose a utility function

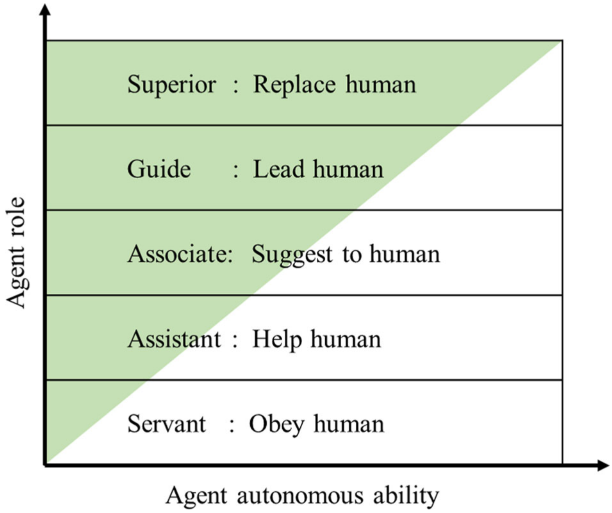


**Fig. 13** A spectrum of an agent roles in HAI (Bradshaw et al. 2002)

that enables the agents to reason about the needs for interaction but based on time and cost measures. However, both models do not consider the possibility of humans' interactions with agents.

In some HAI systems like personal assistant agents, the agents comply with a human (Mercier et al. 2008b; Scerri et al. 2012). For this type of interaction, Myers and Morley (2003) propose consultative and permissive actions for HAI. The consultative actions enable the agents to seek feedback from a human in order to pursue them. The permissive actions enable the agents to elicit authority from a human in order to pursue them. This categorization of actions improves HAI in which agents interact for a particular reason. Alzahrani et al. (2013) set policies to manage autonomy adjustment. Some of which are certified to do the HAI.

However, there are several perspectives of HAI including the problem that triggers the interaction, the teamwork that intends to handle the problem and autonomy distribution between the team members. Each of these perspectives has several issues that need to be considered in HAI. Some of them are summarized in Table 5.

## 7 Adjustable autonomy applications

Autonomous system applications are meant to facilitate humans' life. People have different characteristics, as a result, their preferences, living styles, trust and confidence regarding autonomous systems differ (Ball and Callaghan 2011). Alzahrani et al. (2013) state that people prefer lower autonomy for systems that directly interfere with their personal activities and higher autonomy for multiple-user and public systems such as lighting and cooling systems. Many surveys discuss autonomy influence on people and validate the usefulness of adjustable autonomy applications (Ball and Callaghan 2012). Adjustable autonomy covers a wide range of application domains including robotics, unmanned systems, smart homes, electronic commerce, organizational coordination, and space missions.

**Table 5** Issues of HAI modified from Bradshaw et al. (2003)

| Perspective | Issue |
|---|---|
| Problem perspective | Is the problem achievable? |
| | Is the problem routine or new? |
| | Is the problem high or low priority? |
| | What types of solution strategies are matching with the problem? |
| | What dependencies are considered? |
| | How is the problem state evolving? |
| | Is the problem solution satisfactory? |
| Teamwork perspective | How did we get into this state? |
| | What are they doing now? |
| | Why are they doing it? |
| | Are they having difficulties? Why? |
| | What are they doing to cope with difficulties? |
| | Are they likely to fail? |
| | How long will they be busy? |
| | What will they do next? |
| Autonomy perspective | Who should make the decision to do the interaction? |
| | When is the interaction required? |
| | How the decision of the interaction can be implemented? |
| | What type of actions is fully autonomous? |
| | What type of actions is not autonomous? |
| | What type of actions is semi-autonomous? |
| | How can the system handle each of them? |

### 7.1 Robotics

In this section, we review the application of adjustable autonomy in robotics. Adjustable autonomy is mainly used as a means to model human–robot cooperation (Zieba et al. 2010). It basically considers the options of human–robot teaming and its related issues like interactivity, initiative, intuitiveness, and independence (Schwarz et al. 2014).

Birk and Pfingsthorn (2006) use the concept of agents and adjustable autonomy to model a Human-Machine Interface (HMI) for rescue robots. Their research considers the part of improving direct control, visualization and the portability of the hardware. The second part emphasizes on the robot dynamically adjusts its control based on the situation need and the human user preferences. The autonomy is distributed and adjusted according to the environment inputs and needs of the implemented missions.

Similarly, Schermerhorn and Scheutz (2009) propose adjustable autonomy for Human–Robot Interaction (HRI). They argue that the robot should reason about its compliance with the human commands to avoid mistakes. This issue entails implementing internal autonomy distribution and adjustment mechanisms. The autonomy distribution and adjustment consider the parameters of goal priority that indicate its benefit (i.e., importance and urgency) and time that indicates the cost. The study shows that adjustable autonomy enhances human–robot team performance. It further shows that adjustable autonomy reduces the human cognitive load.

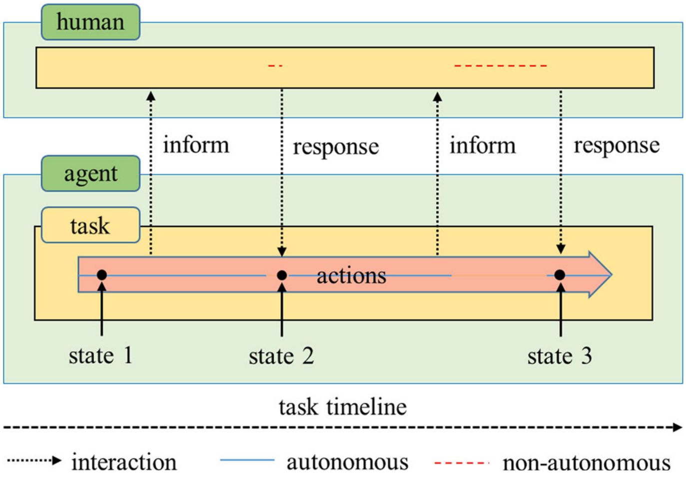


**Fig. 14** An adjustable autonomy for a mixed team (modified from (Roehr et al. 2010))

Roehr et al. (2010) propose a mixed-initiative sliding autonomy for a humans-robot team. The sliding mechanism segregates the autonomy to fully-autonomous, mixed-initiative and manual states. The autonomy distribution and adjustment are performed based on trust and self-confidence measures which are performed by a human. Figure 14 illustrates the system's performance according to the proposed adjustable autonomy.

Valero-Gomez et al. (2011) study the impact of adjustable autonomy on HRI scalability. They evaluate the performance of two adjustable autonomy models to identify their interactions scalability quality. In the first model, a human takes full control over robots when the control is switched to him/her. In the second model, the human controls specific activities of the robots and the robots dynamically maintain other activities. The results show that the robots that use the second model perform better and the second model is found more usable.

Carlin et al. (2010) propose adjustable autonomy for human–agent coordination and cooperation. They develop the CHAMP model (Coordinating with Humans by Adjustable-autonomy for Multi-robot Pursuit) that is simulated for military purposes. The challenge of the CHAMP model is to operate in an environment that incorporates small-unit tactical team and semi-autonomous robots team effectively. Both teams are dynamically controlled via distributed optimization strategies to perform the coordination and operate via adjustable autonomy. The aim of the CHAMP model is to dynamically adjust the coordination of the robots' team towards goal achievement based on the dynamics of the mission. Figure 15 shows the CHAMP robot and its related platforms.

Rau et al. (2013) study the effects of a robot's autonomy on humans' decisions. The robot's autonomy is scaled into ten levels from non-autonomous to fully-autonomous. Each level represents different autonomous capabilities. For example, in one level, the robot only suggests actions, while on another level, it acts autonomously. The study measures the reliability (trust and credibility) and workload as assessment factors (Cheng and Cohen 2006). The results show that when the robot operates in higher autonomy level, it contributes greater impacts on the humans' decisions than when it operates in the lower autonomy level. The higher autonomy level significantly extenuates the humans' workload and gradually acquires their trust.

Petersen and Von Stryk (2011) proposes a supervised agent-based adjustable autonomy model for human–robot teams. The model has a situation overview technique that

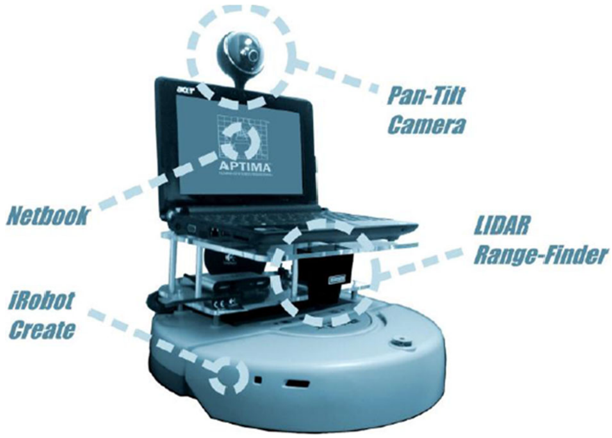


**Fig. 15** CHAMP robot (Carlin et al. 2010)

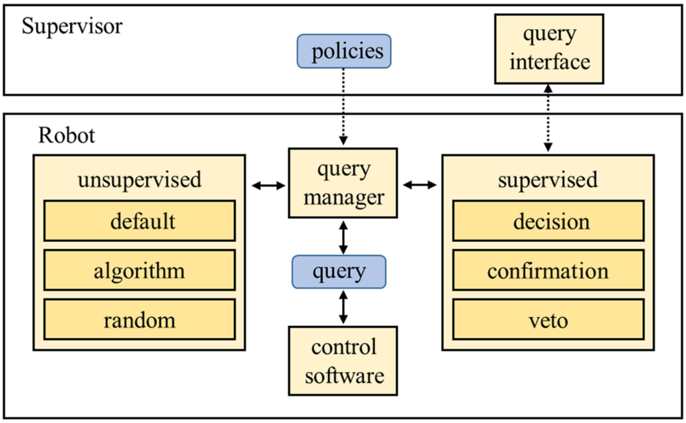


**Fig. 16** A supervised adjustable autonomy model of a robot (Petersen and Von Stryk 2011)

insights a human supervisor about the robot's conditions and a query mechanism to perform the interaction. The situation overview technique shows the human supervisor information about the system's current state, the environment, and the mission through a GUI. The query mechanism distributes the autonomy between the robot and the human. The model considers the contribution of both the agent and human in the autonomy distribution and adjustment. Figure 16 shows the proposed supervised adjustable autonomous robot model.

The actions of the system are categorized based on an operational hierarchy from high-level commands to low-level actions. Human interactions are limited to the high-level commands to minimize system disturbance via excluding human interference in the operational loop. The model is tested via experimenting and simulating urban search and rescue. It is able to maintain an autonomy level that allows the robot to cope with difficult situations. The research finding confirms the need for human supervisors for autonomous systems in the real world and dynamic environment.

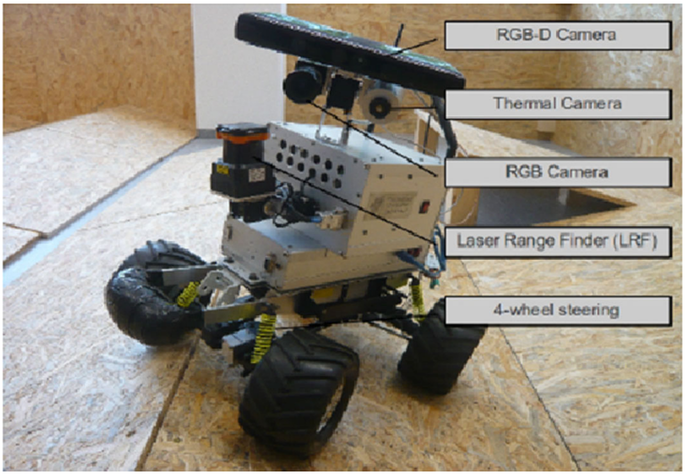


**Fig. 17** Hector UGV (Petersen and Von Stryk 2011)

### 7.2 Unmanned systems

Unmanned systems are considered as mission-specific solutions. They are deployed to operate on the ground, Unmanned Ground Vehicles (UGV); on the air, Unmanned Aerial Vehicle (UAV); on the water, Unmanned Surface Vehicles (USV); underwater, Unmanned Underwater Vehicles (UUV) or both surface and underwater, Unmanned Maritime Vehicle (UMV). Figure 17 shows Hector, a UGV autonomous system which is equipped with several sensors and is capable of exploration missions.

Adjustable autonomy is an essential part of unmanned systems development (Scerri et al. 2009; Schwarz et al. 2014). Recently, the development of adjustable autonomous unmanned systems represents the cutting edge for many technology companies. The technology is used in many applications including reconnaissance, surveillance, inspection, terrain mapping and photography tasks. Figure 18 shows an experiment of applying adjustable autonomy in a USV for solving a real-world challenge of flood disaster mitigation.

The most common adjustable autonomy in unmanned systems considers applying the behaviors of sensing, processing and actuating (Moffitt et al. 2006; Parasuraman et al. 2008; Insaurralde and Lane 2014). de Brun et al. (2008) define adjustable autonomy of unmanned systems as “a mechanism that enables a change in the autonomous level of an unmanned system during mission execution.” In general, there are four levels of autonomy that determine humans’ roles and interactions with unmanned systems (Lin and Goodrich 2015; Lewis et al. 2013; Fagnant and Kockelman 2014):

- Operated autonomy: The system operates based on humans’ instructions.
- Delegated autonomy: Humans delegate authority of operations to the system.
- Supervised autonomy: The system autonomously operates but susceptible to humans’ intervention.
- Full-autonomy: The system operates without interventions.

Reed (2005) proposes an adjustable autonomy model for UAV. He defines adjustable autonomy as a spectrum that ranges from fully-autonomous to non-autonomous command driven autonomy. The agents are distributed in a hierarchy of different autonomy levels as contractors with higher autonomy levels, as engineers with lower autonomy levels. The

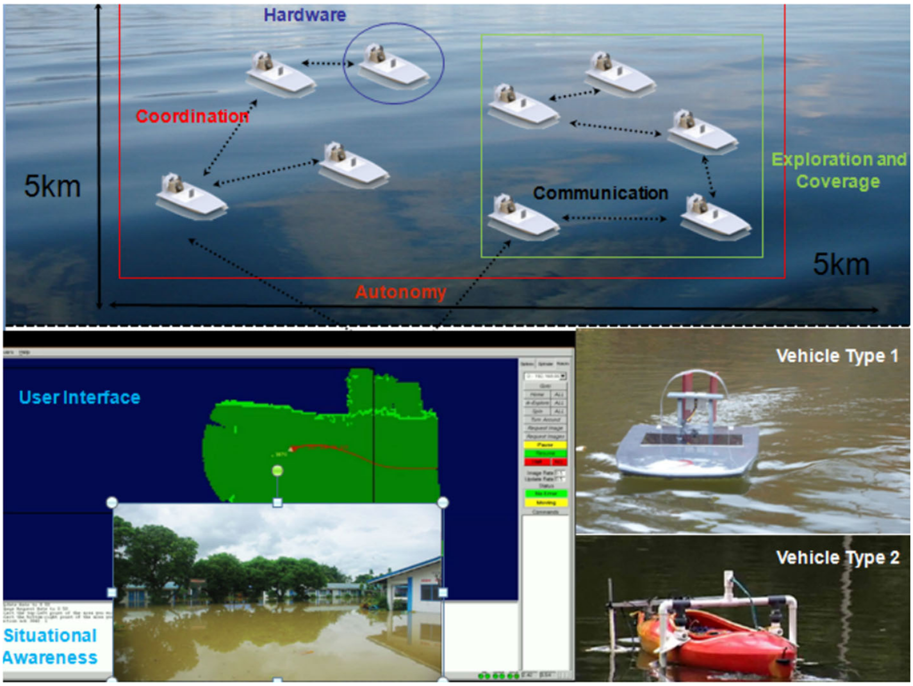


**Fig. 18** Adjustable autonomy of a USV application (Scerri et al. 2012)

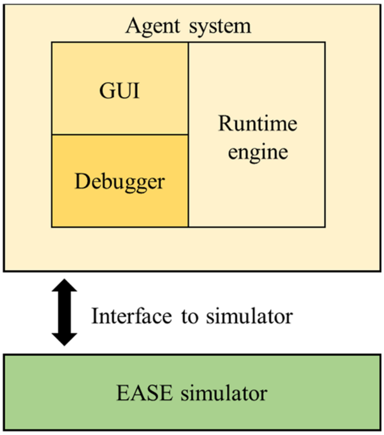


**Fig. 19** The simulated system architecture (Reed 2005)

contractor and engineer agents negotiate and cooperate to achieve piloting tasks. The autonomy is adjusted based on the capability of the agent to complete the given task. When an agent's states indicate no progress, it is switched with another agent. The model is validated via simulated pilot application using End-user Actor Specification Environment (EASE). A human can add, remove and suspend agents and manipulate their parameters through a GUI. Figure 19 shows the architecture of the system.

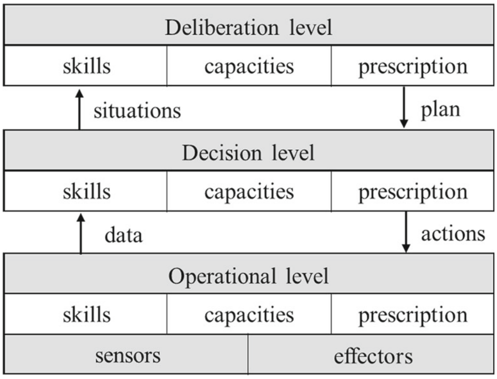


**Fig. 20** The adjustable autonomous agent of Zieba et al. (2011)

Zieba et al. (2011, 2010) propose an adjustable autonomy model with human–machine cooperative ability. They address the need for dynamic autonomy distribution and adjustment mechanisms in highly dynamic environments or high workloads. The model manages the interaction and the cooperation between humans and agents. Humans assist agents to handle difficult situations that the system encounters in three cases: interference, prevention, and recovery. Both humans and agents are capable of adjusting the autonomy. The agents operate in three functional levels of *deliberate-level*, *decisional-level* and *operational-level* as shown in Fig. 20. The autonomy of agents is formulated according to two dimensions: (1) semantic dimension, which represents agents' skills, capacities and prescriptions to perform a particular task, and (2) activity dimension, which contains the three functional levels. Autonomy is measured based on risk assessment. It considers a human error in the assessment loop. The model is simulated in a three-dimensional environment of a UGV system to accomplish object detection and tracking tasks.

Bush et al. (2012) propose an Autonomous Robot Control via Autonomy Levels (ARCAL) model. The ARCAL aims to dynamically distribute the autonomy of a system according to a confidence level of a mission's success. The model has a risk and trust assessment mechanisms that support the confidence level of the mission. The risk is measured via considering two factors: the probability of mission failure and uncertainty of the planning decisions. Agents request human intervention based on the risk, which determines the type of the intervention and hence the autonomy level. The model is simulated in a natural disaster recovery of UGV and UAV systems. The UGV performs the main mission's tasks while the UAV performs the scouting mission's tasks. The UAV assists the UGV and calculates the expected risk. Figure 21 shows the ARCAL model.

Cote et al. (2013) propose a Human Help Provider in a Markov Decision Process-based (HHP-MDP) adjustable autonomy model. The HHP-MDP model works as a middle ground between autonomous agents and a system. It integrates a human supervisor feedback in the agents' decision process. The feedback represents the human's observation and it intends to help the agents in critical situations when the agents (1) lack information or (2) reach some looping states. The human feedback is translated into a policy and the policy is processed according to transition and reward functions and performed via a Markov Decision Process (MDP) algorithm. The MDP algorithm outcomes are associated with states and actions and integrated within the agents' decision processes. The autonomy adjustment is valid only when the human supervisor conducts an intervention. The proposed adjustable autonomy model is designed to be implemented in a UAV system.

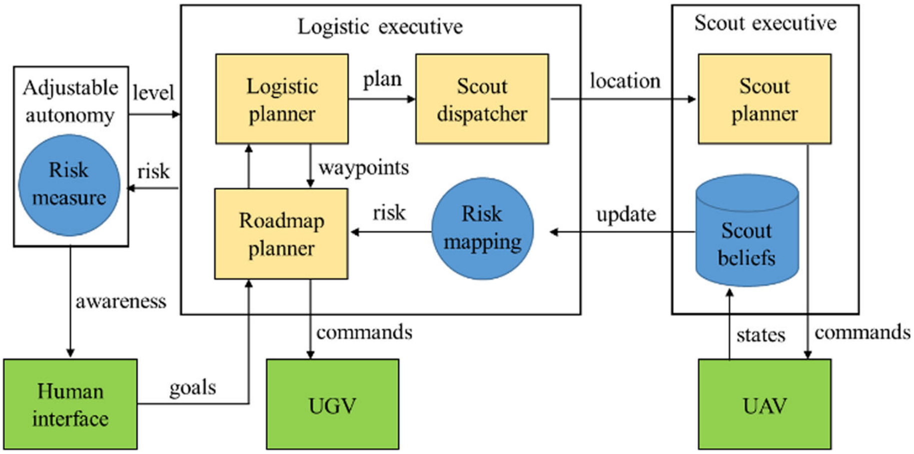


**Fig. 21** The ARCAL model of (Bush et al. 2012)

### 7.3 Other applications

Myers and Morley (2003) propose a Taskable Reactive Agent Communities (TRAC) model for HAI using adjustable autonomy. The TRAC model considers the agents' capability of deciding at different degrees of autonomy and studies their tasks accomplishment. The model aims to achieve reliable autonomous systems through a human directability of agents. Human supervisor defines policies to influence agents' actions. Human intervention is customized in TRAC to overcome unexpected situations and ensure better preference. The TRAC is applied in an Intelligence Gathering and Emergency Response (TIGER) simulation.

A survey conducted by Ball and Callaghan (2011) reveal people's impression about autonomy in smart homes. The autonomy of a smart house is organized into four levels: fully-autonomous, semi-autonomous with low and high autonomy, and non-autonomous. The results of the survey show that different people prefer different levels of autonomy in different situations and for different sub-systems. In time, peoples' views about the autonomy level configuration may shift as a consequence of their experience with the technology (Ball and Callaghan 2012). The survey concludes that peoples' attitudes are directed towards adjustable autonomy. They perceive a preferred autonomy configuration to complete a particular task based on the task performances.

Consequently, Waytz et al. (2014) study people's opinions about personified agent-based autonomous systems. They acquire people's preferences to a normal car, autonomous car and autonomous car with humanoid or anthropomorphic features. The preference measurement attributes are anthropomorphism, liking, and trust. The study shows that people prefer human like attitudes of reasoning and interaction.

Alzahrani et al. (2013) propose an adjustable autonomy model for Adaptive Course Sequencing Application (ACSS) of intelligent tutoring system. The model is meant to improve students' learning. Particularly, the adjustable autonomy makes the ACSS flexible via adjusting its tutoring agents to an autonomy level that fits students' needs. The tutoring agents perform learning guidance actions via setting the learning objects sequence according to the students' profiles analysis results. The tutoring agents reasoning engine contains a set of rules that are conditioned based on the students' profiles. The rules have confidence level values that indicate each rule usage number. Active rules have high confidence levels and potential rules have lower confidence levels. The adjustable autonomy model enables the students to adjust the autonomy of the tutoring agents to intervene some of their rules in order to change the presentation of the learning objects. The autonomy adjustment options are: (1) full-autonomy is the default setting in which the tutoring agents take full control; (2) partial-autonomy, in which the tutoring agents recommend and the students choose, and (3) none-autonomy in which the students manually choose the presentation of the learning objects.

Rajabzadeh, [52] propose an adjustable autonomy model to minimize smart grid power network system disturbance during system's operation. Agents' autonomy over the environment is denoted as authority. The authority is evaluated according to the agents' dependence on others when achieving their goals (i.e., Hexmoor et al. 2012 approach). Autonomy distribution and adjustment are formulated via access control mechanism according to the agents' authority. The system disturbance reduction results from permitting authority to less dependent agents. The proposed model improves system's performance and reduces the effects of some constraints.

## 8 Adjustable autonomy assessment

Autonomy assessment of a system is the process of measuring, evaluating and validating the behavior of the autonomous system in its environment (Mostafa et al. 2015b). The assessment covers different aspects of the system which might include human standpoint assessment such as reliability, usability, dependency and technical standpoint assessment such as flexibility, viability, and decision choice quality (Mostafa et al. 2015a; Parasuraman et al. 2008). Assessing the autonomous capabilities of systems is a challenging process due to many reasons. The United State army considers the capabilities of observation, perception, situation awareness and decision-making in the autonomy assessment of systems (Insaurralde and Lane 2014).

Adjustable autonomy opens the door to autonomy assessment because of its qualitative and quantitative features. Principally, adjustable autonomy assessment concerns with the disturbances that result from autonomy distribution and adjustment between humans and agents (Valero-Gomez et al. 2011; Zieba et al. 2010). For example, Durand et al. (2009) assess the autonomy of a UAV by measuring its performance consistency. They identify the inconsistencies sources in the UAV via its physical parts, algorithms and a number of approximations.

### 8.1 Assessment methodologies

Autonomy assessment in the literature considers quantitative and qualitative measures. Quantitative assessment contributes a number of measurement metrics. The matrices help to trace a system's process flow and identify deficiencies that need to be fine-tuned (Rajabzadeh 2014). However, the meaning of the results is hard to understand (Mercier et al. 2008a). On the other hand, a qualitative assessment is an important methodology, especially, for complex autonomous systems or reactive automation that are difficult to be quantitatively measured. Subsequently, there are two main methodologies to assess autonomy as follows:

- Expert assessment: It is mainly a qualitative assessment that focuses on a system's development aspects. Experts assess the autonomy requirements of a system and their specifications according to some guidelines or criteria (Insaurralde and Lane 2014). The requirements might include system's goal, interaction, usability, autonomy distribution organization, environment, decision-making and situation awareness.

- Computational assessment: It is a widely-used autonomy assessment methodology, especially, in robotics and unmanned systems. The computational assessment implies a quantitative measure for a workable or a simulated autonomous system (Roehr et al. 2010; Huang et al. 2009). It is achieved via implementing measurement metrics to assess the autonomy of a system based on its performance. The measurement metrics consider two dimensions: (1) performance, independence from human, successfulness and consistency, (2) complexity, of the environment and the goal achievement.

### 8.2 Assessment attributes and criteria

There is no general term or specific matrix that can be used in the autonomy assessment process (Insaurralde and Lane 2014). This is due to the autonomy models that are tightly coupled with a wide range of application domains (Mostafa et al. 2013a). Subsequently, the notion of autonomy assessment is a loosely defined term. The commonly used notions in the autonomy assessment literature are the attributes of usability, viability, effectiveness, efficiency and resilience (Mostafa et al. 2015a). These attributes are defined by a number of criteria. The criteria are associated with the assessed autonomous system domain.

Conceivably, an individual attribute is not sufficient to assess the autonomy. Parasuraman et al. (2008) show that reliability consideration only as an autonomy attribute is not sufficient. There are other critical demands of autonomous systems that need to be fulfilled. Hence, we see a combination of different attributes is used in the literature in autonomy assessment. The following includes a number of autonomy assessment examples.

Insaurralde and Lane (2014) define autonomy usability as the ability of an autonomous system to successfully perform actions without human intervention. Hence, the usability entails measurement to human workload (Jennings et al. 2014; Schwarz et al. 2014; Naderpour and Lu 2013). Ball and Callaghan (2012, 2011) define autonomy usability according to human users' preferences.

Roehr et al. (2010) consider efficiency and reliability as attributes for autonomy assessment. These are aimed to maximize overall system autonomy and ensure better performance. Autonomy efficiency is defined by the criteria of interaction constraints, e.g., cost and number of successes in performance. Autonomy reliability is measured according to the criteria of the cost of losing, damaging or replacing the autonomous system.

Miller and Parasuraman (2007) define a viable autonomy for a UAV system via autonomy usability and flexibility. They measure usability according to the criteria of human mental workload, user satisfaction, and overall system performance. They measure autonomy flexibility according to HAI and situation awareness capabilities. Consequently, Parasuraman et al. (2008) define a viable autonomy by its ability to reduce human mental workload, situation awareness, and reliable performance.

Mostafa et al. (2015a) define a viable autonomy by three autonomy attributes: usability, flexibility, and reliability. The autonomy usability is defined by applicability, complexity, domain independence and human user satisfaction. The autonomy flexibility is defined by the reasoning initiative of HAI in which both humans and agents are involved in the autonomy distribution and adjustment process; the autonomy has a representation down to the action level and there are situation awareness capabilities. The autonomy reliability is defined by human's authority of control and system resistance to disturbances.

Zieba et al. (2010, 2011) consider resilience as the main attribute in assessing adjustable autonomy. The resilience measurement considers other attributes of efficiency, adaptability, stability and interaction capabilities (i.e., autonomy flexibility as in Cheng and Cohen (2005)).

Finally, De Visser et al. (2008, 2010) consider system's ability of situation awareness as key criteria of autonomy assessment.

### 8.3 Assessment techniques

Applying adjustable autonomy to a system causes a number of drawbacks. The drawbacks include increased human operator workload, system's disturbances, and autonomy constraints problems (e.g., the flexibility of decisions). These drawbacks are crucial when the system is operating in a dynamic environment. The literature proposes different assessment techniques to investigate the adjustable autonomy drawbacks and the following are some examples.

Sellner et al. (2006) focus on assessing human operators' ability in controlling sliding autonomy of robots. They address that humans' responses are the bottleneck of system's responses' speed compared with other autonomous entities. Nevertheless, autonomous agents with uncertainties need humans' assistance to perform successful actions.

Insaurralde and Lane (2014) propose a measurement metric for UMV autonomous capability assessment. The assessment objective is to configure satisfactory autonomy degrees or levels for the UMV system. They assess the autonomy by two measurement metrics: Degree of Autonomy which is a real number, the value of which indicates the amount of autonomy that the system has and Level of Autonomy which is a natural number that indicates the grade of operations that are autonomously performed. They define the autonomy measurement via the usability of an autonomous system with or without human intervention. Figure 22 shows the proposed autonomy assessment technique for UMV.

Situation awareness assessment is another technique that measures, evaluates and validates a process, procedure or mental ability of a system. It investigates the flow points in a system that might cause or caused risk or failure (Wardziński 2006) Situation awareness assessment techniques have been used to assess the operational autonomy of many systems, e.g., Naderpour and Lu's (2013) technological disaster avoidance system.

Scholtz et al. (2004) propose a situation awareness assessment technique for the human–robot interface. It is applied in a vehicle-based robotic autonomous system to evaluate the supervisory interface of HRI. The human supervisor should be aware when the vehicle is in risk or trouble and provides assistance. The interface informs the supervisor about the system, the environment, and the route. The supervisor is timely notified about the system's attitude

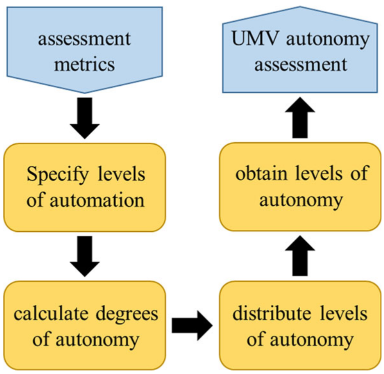


**Fig. 22** The assessment process of a system's autonomy (Insaurralde and Lane 2014)

(i.e., normal, cautionary or hazardous) to do the required assistance. The results show that the autonomy assessment is correlated with the given workload and time.

## 9 Discussion

The required degree or level of autonomy to which operators of humans and agents adhere is a highly-debated topic (Hexmoor et al. 2012). Researchers argue that strict minimal autonomy to agents is sufficient in producing reliable systems, while others deliberate that agents with fully autonomous capabilities represent an essential aspect of advanced intelligent and flexible systems. Adjustable autonomy is proposed to provide a variable range of autonomy to agents. The autonomy adjustment within this range determines the autonomous capabilities of the agents including the operational roles and intervention rules.

Autonomy adjustment represents an interventional power that redistributes the operational control between humans and agents of a system. The adjustment can be internally made by an agent or externally set by human or other systems. The agent must be equipped with mechanisms that enable it to trigger or respond to the adjustment. The adjustment affects the agent's decisions and actions. It either changes the ability of the agent to make some particular decisions or influences the decisions and the final outcome.

Adjustable autonomy incorporates flexibility to autonomous systems' development in a generic way. The literature proposes different adjustable autonomy approaches to overcome some of the autonomous systems' challenges. Figure 23 summarizes the proposed approaches and their associated requirements. The approaches are categorized according to humans' role, adjustable autonomy structure, and adjustable autonomy application mechanisms in an autonomous system.

Each of the adjustable autonomy approaches attempts to achieve different objectives. For instance, the teamwork and the supervised adjustable autonomy are built upon setting human roles in the autonomous system operation. The teamwork-centered approach improves

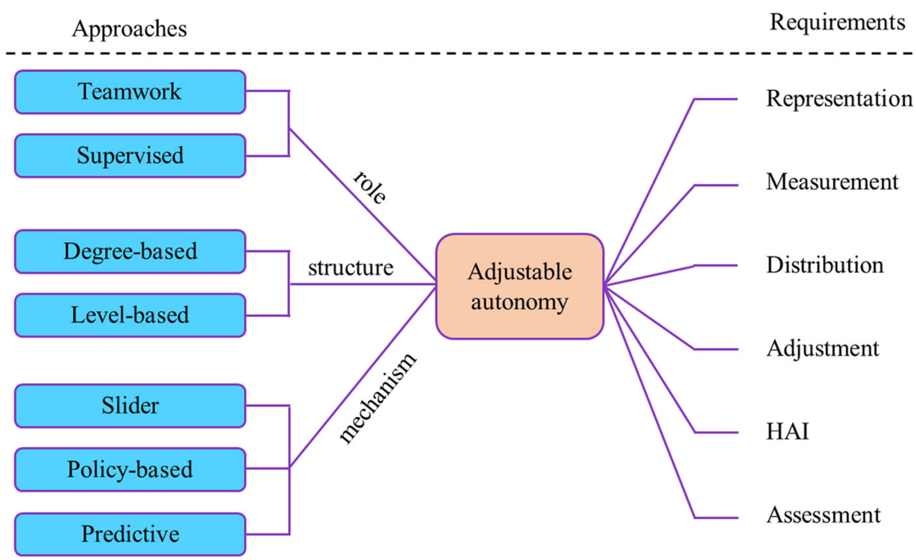


Fig. 23 Adjustable autonomy approaches and requirements

the team performance for the systems. The supervised adjustable autonomy approach improves the autonomous capabilities of systems. The degree and level-based approaches improve the autonomy distribution and adjustment. Subsequently, different combinations or hybrids of approaches can be adopted to satisfy different objectives. It is clear that the teamwork-centered or supervised adjustable autonomy approaches can be combined with other approaches. For instance, Moffitt et al. (2006) propose a policy-based mixed-initiative adjustable autonomy and Cote et al. (2013) propose a policy-based supervised adjustable autonomy.

However, until today, there is no specific model or algorithm that yields viable adjustable autonomy. The following issues pinpoint the shortcomings of adjustable autonomy research according to this study.

- Autonomy distribution is a flexible structure that manages dynamic autonomy configuration in a system's operation. It is considered as the backbone of adjustable autonomy. However, autonomy distribution in some systems is found to be scattered within the autonomous system design. It is mostly human-based as in the case of the mixed-initiative and the sliding autonomy, e.g., Cote et al. (2013) and Lin and Goodrich (2015) approaches which constrain agents' autonomy flexibility and increase human's workload.
- There are a few attempts in the literature that include users' preferences in adjustable autonomy (Ball and Callaghan 2012; Durand et al. 2009). The autonomy distribution and adjustment operations can be employed to satisfy a system user's preference (Waytz et al. 2014). For example, the Internet of Things research can be a suitable coverage to users' preferences in adjustable autonomy.
- Software agents and adjustable autonomy are integrated into many autonomous systems. However, there are a few attempts to develop a software agent model that can efficiently perform in an environment with adjustable autonomy (Mostafa et al. 2013b). Consequently , developing adjustable autonomous agents demands explicit mechanisms to determine agents' variable ranges of autonomy and its related measures (Johnson et al. 2012; Alzahrani et al. 2013).
- Situation awareness is an important aspect of autonomous systems' development (Roehr et al. 2010). We found that there are a few attempts to apply situation awareness in adjustable autonomy. It can be used to determine when and how a human should intervene (Petersen and Von Stryk 2011).
- Most of the adjustable autonomy research results are obtained based on simulation programs, e.g., Bush et al. (2012) and Pantelis et al. (2014). Many important considerations about the constraints of the real-world and dynamic environments might be neglected in the simulations' settings (De Visser et al. 2008; Insaurralde and Lane 2014), hence, the results might lack valid testing.

## 10 Conclusion

This paper presents a Systematic Literature Review (SLR) on adjustable autonomy. The applied SLR methodology has three phases of Plan Review, Conduct Review, and Document Review. It structures the review process and provides search strategies that reduce the bias of the study. The study reviews more than 171 references that include papers, articles, and books. It summarizes and analyzes more than 87 sources that focus on adjustable autonomy formalization of multi-agent systems.

The paper provides a comprehensive review of adjustable autonomy and highlights its research challenges. The paper contributions manifest solutions to three research question pertaining to adjustable autonomy formulation of multi-agent systems. Answering these questions adds value to adjustable autonomy research progress. We summarize the paper's contributions in the following points:

- It identifies seven adjustable autonomy approaches and assesses their utilities. This contribution could help researchers to select and apply a convenient adjustable autonomy approach for their work.
- It specifies six general requirements of formulating adjustable autonomy which are: representation, measurement, distribution, adjustment, HAI and assessment. Some of these requirements are not clearly specified or found to be combined with others. This contribution offers a complete framework for adjustable autonomy researchers in developing adjustable autonomy models. It supports a dynamic spectrum of autonomy distribution and works beyond discrete autonomy configuration.
- It presents the various adjustable autonomy assessment techniques and their criteria. The assessment techniques could be deployed to evaluate the abilities or configure the autonomy distribution of autonomous systems.
- It summarizes the adjustable autonomy research contributions and identifies five research gaps. The research gaps include problems which have not been appropriately or completely solved. The current adjustable autonomy research lacks advanced autonomy distribution mechanisms, application of user preferences, agent models that are designed to be compatible with adjustable autonomy, deployment of situation awareness techniques to improve adjustable autonomy and valid testing platforms.

## References

Alan A, Costanza E, Fischer J, Ramchurn SD, Rodden T, Jennings NR (2014) A field study of human-agent interaction for electricity tariff switching. In: Proceedings of the 2014 international conference on autonomous agents and multi-agent systems. International Foundation for Autonomous Agents and Multi-agent Systems, pp 965–972

Alzahrani A, Callaghan V, Gardner M (2013) Towards adjustable autonomy in adaptive course sequencing. In: Intelligent environments (workshops), pp 466–477

Amalberti R (2006) Optimum system safety and optimum system resilience: agonistic or antagonistic concepts. In: Hollnagel E et al (eds) Resilience engineering: concepts and precepts. Ashgate, Farnham, pp 253–271

Ball M, Callaghan V (2011) Introducing intelligent environments, agents and autonomy to users. In: 7th international conference on intelligent environments (IE)IEEE, pp 382–385

Ball M, Callaghan V (2012) Explorations of autonomy: an investigation of adjustable autonomy in intelligent environments. In: 2012 8th international conference on intelligent environments (IE)IEEE, pp 114–121

Beal J, Webb J, Atighetchi M (2010) Adjustable autonomy for cross-domain entitlement decisions. In: Proceedings of the 3rd ACM workshop on artificial intelligence and security. ACM, pp 65–71

Bibel W (2010) General aspects of intelligent autonomous systems. In: Pratihar DK, Lakhmi CJ (eds) Intelligent autonomous systems. Springer, Berlin, pp 5–27

Birk A, Pfingsthorn M (2006) A HMI supporting adjustable autonomy of rescue robots. In: Robocup 2005: Robot Soccer World Cup IX. Springer, Berlin, pp 255–266

Bradshaw JM, Sierhuis M, Gawdiak Y, Jeffers R, Suri N, Greaves M (2001) Teamwork and adjustable autonomy for the personal satellite assistant. In: Workshop on autonomy, delegation and control: interacting with autonomous agents. Seventeenth international joint conference on artificial inelligence (IJCAI-2001), Seattle, WA, vol 6

Bradshaw JM, Boy G, Durfee E, Gruninger M, Hexmoor H, Suri N, Tambe M, Uschold M, Vitek J (2002) Software agents for the warfighter. ITAC consortium report. AAAI Press/The MIT Press, Cambridge, MA. (14) Bradshaw, vol 1, pp 89–95

Bradshaw JM, Sierhuis M, Acquisti A, Feltovich P, Hoffman R, Jeffers R, Prescott D, Suri N, Uszok A, Van Hoof R (2003) Adjustable autonomy and human-agent teamwork in practice: an interim report on space applications. In: Hexmoor H, Castelfranchi C, Falcone R (eds) Agent autonomy. Springer, US, pp 243–280

Bradshaw JM, Jung H, Kulkarni S, Allen J, Bunch L, Chambers N, Feltovich P, Galescu L, Jeffers R, Johnson M, Taysom W (2004) Toward trustworthy adjustable autonomy and mixed-initiative interaction in KAoS. In: Proceedings of the AAMAS 2004 trust workshop, New York City, NY

Brereton B, Kitchenham A, Budgen D, Turner M, Khalil M (2007) Lessons from applying the systematic literature review process within the software engineering domain. J Syst Softw 80(4):571–583

Brookshire J, Singh S, Simmons R (2004) Preliminary results in sliding autonomy for coordinated teams. In: Proceedings of the 2004 spring symposium series

Bush LA, Wang AJ, Williams BC (2012) Risk-based sensing in support of adjustable autonomy. In: Aerospace conference, 2012. IEEE, pp 1–18

Calero C, Bertoa M, Angeles Moraga M (2013) A systematic literature review for software sustainability measures. In: Proceedings of the 2nd international workshop on green and sustainable Software (GREENS), pp 46–53

Carlin A, Ayers J, Rousseau J, Schurr N. (2010). Agent-based coordination of human-multi-robot teams in complex environments. In: 9th international conference on autonomous agents and multi-agent systems: industry track. International Foundation for Autonomous Agents and Multi-agent Systems, pp 1747–1754

Ceballos A, Bensalem S, Cesta A, De Silva L, Fratini S, Ingrand F, Ocon J, Orlandini A, Py F, Rajan K, Rasconi R (2011) A goal-oriented autonomous controller for space exploration. ASTRA, p 11

Cheng MY, Cohen R (2005) A hybrid transfer of control model for adjustable autonomy multi-agent systems. In: Proceedings of the fourth international joint conference on autonomous agents and multi-agent systems. ACM, pp 1149–1150

Cheng MY, Cohen R (2006) Integrating information gathering interaction into transfer of control strategies in adjustable autonomy multi-agent systems. In: Lamontagne L, Marchand M (eds) Advances in artificial intelligence. Springer, Berlin, pp 1–12

Codognet P (2011) A simple language for describing autonomous agent behaviors. In: The seventh international conference on autonomic and autonomous systems (ICAS 2011)

Cote N, Bouzid M, Mouaddib AI (2013) Integrating human recommendations in the decision process of autonomous agents. In: 2013 IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT)IEEE, vol 2, pp 195–200

Dallaire P, Giguère P, Émond D, Chaib-Draa B (2014) Autonomous tactile perception: a combined improved sensing and Bayesian nonparametric approach. Robot Autonom Syst 62(4):422–435

de Brun ML, Moffitt VZ, Franke JL, Yiantsios D, Housten T, Hughes A, Fouse S, Housten D (2008) Mixed-initiative adjustable autonomy for human/unmanned system teaming. In: AUVSI unmanned systems North America conference

De Visser EJ, LeGoullon M, Freedy A, Freedy E, Weltman G, Parasuraman R (2008) Designing an adaptive automation system for human supervision of unmanned vehicles: a bridge from theory to practice. In: Proceedings of the human factors and ergonomics society annual meeting, vol 52, no. 4. SAGE Publications, pp 221–225

De Visser EJ, LeGoullon M, Horvath D, Weltman G, Freedy A, Durlach P, Parasuraman R. (2010). TECRA: C2 application of adaptive automation theory. In: Aerospace conference, 2010 IEEE, pp 1–12. IEEE

Dorais G, Bonasso, RP, Kortenkamp D, Pell B, Schreckenghost D (1999) Adjustable autonomy for human-centered autonomous systems. In: Working notes of the sixteenth international joint conference on artificial intelligence workshop on adjustable autonomy systems, pp 16–35

Durand B, Godary-Dejean K, Lapierre L, Crestani D (2009) Inconsistencies evaluation mechanisms for a hybrid control architecture with adaptive autonomy. In: CAR'09: 4th national conference on control architectures of robots

Ermon S, Gomes C, Selman B, Vladimirsky A (2012) Probabilistic planning with non-linear utility functions and worst-case guarantees. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems, vol 2, pp 965–972

Fagnant DJ, Kockelman KM (2014) The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transp Res Part C Emerg Technol 40:1–13

Falcone R, Castelfranchi C (1999) Levels of delegation and levels of help for agents with adjustable autonomy. In: 1999 AAAI spring symposium on agents with adjustable autonomy, Mar 22–24, 1999. Stanford University, pp 25–32

Ferguson G, Allen JF, Miller BW (1996) TRAINS-95: towards a mixed-initiative planning assistant. In: AIPS, pp 70–77

Fleming M, and Cohen R (2004) A decision procedure for autonomous agents to reason about interaction with humans. In: Proceedings of the AAAI 2004 spring symposium on interaction between humans and autonomous systems over extended operation, pp 81–86

Freedy A, Sert O, Freedy E, McDonough J, Weltman G, Tambe M, Gupta T, Grayson W, Cabrera P (2008) Multi-agent adjustable autonomy framework (MAAF) for multi-robot, multi-human teams. In: International symposium on collaborative technologies and systems. CTS 2008 IEEE, pp 498–505

Goodrich MA, Olsen DR, Crandall JW, Palmer TJ (2001) Experiments in adjustable autonomy. In: Proceedings of IJCAI workshop on autonomy, delegation and control: interacting with intelligent agents, pp 1624–1629

Hexmoor H, Castelfranchi C, Falcone R (eds) (2012) Agent autonomy, vol 7. Springer, Berlin

Huang HM, Messina E, Jacoff A (2009) Performance measures framework for unmanned systems (PerMFUS): initial perspective. In: Proceedings of the 9th workshop on performance metrics for intelligent systems, pp 65–72. ACM

Huber MJ (2007) Agent autonomy: social integrity and social independence. In: Fourth international conference on information technology, 2007. ITNG'07. IEEE, pp 282–290

Insaurralde CC, Lane DL (2014) Metric assessment of autonomous capabilities in unmanned maritime vehicles. Eng Appl Artif Intell 30:41–48

Inyiama HC, Ufoaroh SU, Chijindu VC. (2012). Intelligent agents autonomy issues. Afr J Comput ICT 5(4)

Jatoth C, Gangadharan GR, Buyya R (2015) Computational intelligence based QoS-aware web service composition: a systematic literature review. IEEE Trans Serv Comput, pp 1–18

Jennings NR, Moreau L, Nicholson D, Ramchurn S, Roberts S, Rodden T, Rogers A (2014) Human–agent collectives. Commun ACM 57(12):80–88

Johnson M, Bradshaw JM, Feltovich PJ, Jonker C, Van Riemsdijk B, Sierhuis M (2012) Autonomy and interdependence in human–agent-robot teams. Intell Syst IEEE 27(2):43–51

Johnson M, Bradshaw JM, Feltovich PJ, Jonker CM, Van Riemsdijk MB, Sierhuis M (2014) Coactive design: designing support for interdependence in joint activity. J Human Robot Interact 3(1):2014

Lewis B, Tastan B, Sukthankar G (2013) An adjustable autonomy paradigm for adapting to expert-novice differences. In: 2013 IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 1656–1662. IEEE

Lin L, Goodrich MA (2015) Sliding autonomy for UAV path-planning: adding new dimensions to autonomy management. In: Proceedings of the 2015 international conference on autonomous agents and multi-agent systems. International Foundation for Autonomous Agents and Multi-agent Systems, pp 1615–1624

Luck M, D'Inverno M, Munroe S (2003) Autonomy: variable and generative. In: Hexmoor H, Castelfranchi, C, Falcone R (eds) Agent autonomy. Springer US, pp 11–28

Magill K, Erden YJ (2012) Autonomy and desire in machines and cognitive agent systems. Cognit Comput 4(3):354–364

Mercier S, Dehais F, Lesire C, Tessier C (2008a) Resources as basic concepts for authority sharing. Humans Operating Unmanned Systems, HUMOUS

Mercier S, Tessier C, Dehais F (2008b) Adaptive autonomy for a human-robot architecture. In: 3rd national conference on control architectures of robots, Bourges

Miller CA, Parasuraman R (2007) Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. Hum Fact J Hum Fact Ergon Soc 49(1):57–75

Moffitt VZ, Franke JL, Lomas M (2006) Mixed-initiative adjustable autonomy in multi-vehicle operations. In: Proceedings of AUVSI, Orlando, Florida

Mostafa SA, Ahmad MS, Annamalai M, Ahmad A, Gunasekaran SS (2013a) A conceptual model of layered adjustable autonomy. Adv Intell Syst Comput 206:619–630

Mostafa SA, Ahmad MS, Annamalai M, Ahmad A, Gunasekaran SS (2013b) A dynamically adjustable autonomic agent framework. Adv Intell Syst Comput 206:631–642

Mostafa SA, Ahmad MS, Annamalai M, Ahmad A, Basheer GS (2013c) Frontiers in artificial intelligence and applications. A layered adjustable autonomy approach for dynamic autonomy distribution. IOS Publisher, New York

Mostafa SA, Ahmad MS, Ahmad A, Annamalai M, Mustapha A (2014a) A dynamic measurement of agent autonomy in the layered adjustable autonomy model. Stud Comput Intell 513(513):25–35

Mostafa SA, Ahmad MS, Tang AY, Ahmad A, Annamalai M, Mustapha A (2014b) Agent's autonomy adjustment via situation awareness. Lect Notes Comput Sci 8397:443–453

Mostafa S, Gunasekaran SS, Ahmad MS, Ahmad A, Annamalai M, Mustapha A (2014c) Defining tasks and actions complexity-levels via their deliberation intensity measures in the layered adjustable autonomy model. In: International conference on intelligent environments (IE), 2014 IEEE, pp 52–55

Mostafa SA, Ahmad MS, Ahmad A, Annamalai M, Gunasekaran SS (2015a) An autonomy viability assessment matrix for agent-based autonomous systems. In: International symposium on agents, multi-agent systems and robotics (ISAMSR), pp 53–58. IEEE

Mostafa SA, Ahmad MS, Annamalai M, Ahmad A, Gunasekaran SS (2015b) Formulating dynamic agents' operational state via situation awareness assessment. Adv Intell Syst Comput 320:545–556

Musliner D, Pell B (1999) Agents with adjustable autonomy. Papers from the AAAI spring symposium. Technical Report SS-99-06, AAAI Press, Menlo Park, CA
Myers KL, Morley DN (2003) Policy-based agent directability. In: Hexmoor H, Castelfranchi C, Falcone R (eds) Agent autonomy. Springer, US, pp 185–209
Naderpour M, Lu J (2013) A human situation awareness support system to avoid technological disasters. In: Vitoriano B, Montero J, Ruan D (eds) Decision aid models for disaster management and emergencies. Atlantis Press, pp 307–325
Pantelis PC, Baker CL, Cholewiak SA, Sanik K, Weinstein A, Wu CC, Tenenbaum JB, Feldman J (2014) Inferring the intentional states of autonomous virtual agents. Cognition 130(3):360–379
Parasuraman R, Sheridan TB, Wickens CD (2008) Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. J Cognitive Eng Decis Mak 2(2):140–160
Pătraşcu M, Drăgoicea M (2014) Integrating agents and services for control and monitoring: managing emergencies in smart buildings. In: Borangiu T, Trentesaux D, Thomas A (eds) Service orientation in holonic and multi-agent manufacturing and robotics. Springer, pp 209–224
Petersen K, Von Stryk O (2011) Towards a general communication concept for human supervision of autonomous robot teams. In: Proceedings of the fourth international conference on advances in computer-human interactions (ACHI). pp 228–235
Pratihar DK, Jain LC (2010) Towards intelligent autonomous systems. In: Intelligent autonomous systems. Springer, Berlin, pp 1–4
Py F, Rajan K, McGann C (2010) A systematic agent framework for situated autonomous systems. In: Proceedings of the 9th international conference on autonomous agents and multi-agent systems: volume 2—Volume 2. International Foundation for Autonomous Agents and Multi-agent Systems, pp 583–590
Rajabzadeh RA (2014) Measuring autonomy and solving general stabilization problems with multi-agent systems. Master diss., University of Nebraska, Lincoln
Rau PLP, Li Y, Liu J (2013) Effects of a social robot's autonomy and group orientation on human decision-making. Adv Hum Comput Interact 2013(11):1–13
Reed NE (2005) A user controlled approach to adjustable autonomy. In: Proceedings of the 38th annual Hawaii international conference on system sciences, 2005. HICSS'05. IEEE, pp 295b–295b
Roehr TM, Shi Y (2010) Using a self-confidence measure for a system-initiated switch between autonomy modes. In: Proceedings of the 10th international symposium on artificial intelligence, robotics and automation in space, Sapporo, Japan, pp 507–514
Rogers A, Ramchurn S, Jennings NR (2012) Delivering the smart grid: challenges for autonomous agents and multi-agent systems research. In: Twenty-sixth AAAI conference on artificial intelligence (AAAI-12), Toronto, CA, 22–26 July 2012, pp 2166–2172
Sadaghiyani JS (2011) Concepts of decision making under uncertain, risky and deterministic situations. J Math Comput Sci 2(3):529–545
Scerri P, Reed N (2001) Designing agents for systems with adjustable autonomy. Department of Computer and Information Science, Linköpings Universitet, Linköpings
Scerri P, Sycara K, Tambe M (2009) Safety in the context of coordination via adjustable autonomy. In: Safety and security in multi-agent systems. Springer, Berlin, pp 208–224
Scerri P, Kannan B, Velagapudi P, Macarthur K, Stone P, Taylor M, Dolan J, Farinelli A, Chapman A, Dias B, Kantor G (2012) Flood disaster mitigation: a real-world challenge problem for multi-agent unmanned surface vehicles. In: Dechesne F, Hattori H, ter Mors A, Such JM, Weyns D, Dignum F (eds) Advanced agent technology. Springer, Berlin, pp 252–269
Schermerhorn P, Scheutz M (2009) Dynamic robot autonomy: investigating the effects of robot decision-making in a human–robot team task. In: Proceedings of the 2009 international conference on Multimodal interfaces. ACM, pp 63–70
Scholtz J, Antonishek B, Young J. (2004). Evaluation of a human–robot interface: development of a situational awareness assessment methodology (ETSIB03). In: International conference on annual Hawaii system sciences, pp 130–130
Schurr N (2007) Toward human–multi-agent teams. ProQuest, Ann Arbor
Schurr N., Marecki J, Tambe M. (2008). RIAACT: a robust approach to adjustable autonomy for human-multi-agent teams. In: Proceedings of the 7th international joint conference on autonomous agents and multi-agent systems, Volume 3. International Foundation for Autonomous Agents and Multi-agent Systems, pp 1429–1432
Schurr N, Marecki J, Tambe M (2009) Improving adjustable autonomy strategies for time-critical domains. In: Proceedings of the 8th international conference on autonomous agents and multi-agent systems, Volume 1. International Foundation for Autonomous Agents and Multi-agent Systems, pp 353–360

Schwarz M, Stückler J, Behnke S (2014) Mobile teleoperation interfaces with adjustable autonomy for personal service robots. In: Proceedings of the 2014 ACM/IEEE international conference on human–robot interaction, pp 288–289. ACM

Sellner B, Heger FW, Hiatt LM, Simmons R, Singh S (2006) Coordinated multi-agent teams and sliding autonomy for large-scale assembly. Proc IEEE 94(7):1425–1444

Sycara K, Sukthankar G (2006) Literature review of teamwork models. Robotics Institute, Carnegie Mellon University, Pittsburgh

Tipaldi M, Glielmo L (2015) A Markovian based approach for autonomous space systems. In: Metrology for aerospace (metroaerospace), 2015 IEEE, pp 426–430. IEEE

Truszkowski W, Hallock H, Rouff C, Karlin J, Rash J, Hinchey M, Sterritt R (2009) Autonomous and autonomic systems: with applications to NASA intelligent spacecraft operations and exploration systems. Springer, Berlin

Valero-Gomez A, De La Puente P, Hernando M (2011) Impact of two adjustable autonomy models on the scalability of single-human/multiple-robot teams for exploration missions. Hum Factors J Hum Factors Ergon Soc 53(6):703–716

van der Vecht, B. (2009) Adjustable autonomy: controlling influences on decision making (No. 2009–20). Utrecht University

van der Vecht B, Dignum F, Meyer JJC, Neef M (2008) A dynamic coordination mechanism using adjustable autonomy. Springer, Berlin

Wallace SA, Henry M (2008) Towards a generic infrastructure to adjust the autonomy of soar agents. In: FLAIRS conference, pp 119–120

Wardziński A. (2006). The role of situation awareness in assuring safety of autonomous vehicles. In: Górski J (ed) Computer safety, reliability, and security. Springer, Berlin, pp 205–218

Waytz A, Heafner J, Epley N (2014) The mind in the machine: anthropomorphism increases trust in an autonomous vehicle. J Exp Soc Psychol 52:113–117

Wilson MA, McMahon J, Aha DW (2014) Bounded expectations for discrepancy detection in goal-driven autonomy. In: Workshops at the twenty-eighth AAAI conference on artificial intelligence

Witkowski M, Stathis K (2004) A dialectic architecture for computational autonomy. In: Nickles M, Rovatsos M, Weiss G (eds) Agents and computational autonomy. Springer, Berlin, pp 261–273

Zieba S, Polet P, Vanderhaegen F, Debernard S (2010) Principles of adjustable autonomy: a framework for resilient human-machine cooperation. Cognit Technol Work 12(3):193–203

Zieba S, Polet P, Vanderhaegen F (2011) Using adjustable autonomy and human–machine cooperation to make a human–machine system resilient-application to a ground robotic system. Inf Sci 181(3):379–397

Zilberstein S (2015) Building strong semi-autonomous systems. In: AAAI, pp 4088–4092