Chapter 2
Literature Review

"The scientist is not a person who gives the right answers, he's one who asks the right questions."

Claude Lévi–Strauss

Contents

2.1 Introduction

This chapter provides an overview of works relevant to the topic of the thesis. It is divided into three sections roughly corresponding to the chapters of the thesis.

Section 2.2 provides an overview of collaboration theories that are inspired from human–human collaboration. Section 2.3 includes an overview of the approaches of human–robot collaboration and application known from the available works. This section also provides background on cognitive architecture. Section 2.3 provides a brief summary of some of the work on cognitively enhanced control that has inspired this research, together with an argument for why idea of cognitively enhanced is compelling. This section also includes background on robot control architecture. Section 2.4 goes more in depth into human navigation, which is a domain that is of particular interest. Section 2.4 starts with background on intelligent wheelchair.

2.2 Human Robot Collaboration

Quest for natural and human like intelligent artificial systems influences the emergence of Human-machine (Robot) interaction (HRI) as a field. HRI research community readily acknowledges that NASA's *vision for space exploration*[11] has been influential in much of the research and development in HRI around the world.

An inclination in recent work on HRI has been to view human-robot interaction as collaboration. This can be marked by the fact that it was in the year 2006 when the inaugural HRI conference was on the theme:

> *"Toward human-robot collaboration, highlights the importance of creating robot capabilities and interfaces that dynamically balance human and robot competencies while addressing human concerns such as social appropriateness, safety, and quality of service"*.
>
> HRI 2006 [12]

In this thesis the human-robot system is understood as a system that act together with a human to achieve a common goal.

2.2.1 Human Robot Collaboration: The human

According to [13], collaboration is an indispensable feature of human social interaction. Human join in collaboration because it enables them to achieve a task which, either individual is unable to achieve on its own or is difficult for a single person to do. Human ability to engage in joint action with others is crucial for participating in collaborative activities; requires forming shared intentions and having shared goals [14],[13]. It was Green [15] who specifically mentioned that a way human collaborate with the other human in a task is a source of inspiration to how robot should collaborate with humans. There is an extensive body of work focused on understanding what is involved in the human-human interaction for successful collaboration. Within this section, a few key points are addressed such as the rules that govern the formation and maintenance of human-human collaboration.

Joint Intention

How do human engage in collaboration is a question which is addressed in terms of joint intention [16],[17], [18]. Tuomela and Miller [13] proposed the concept of "we-ness" as a means of describing human-human collaboration. According to Warneken [19], young children even have tendency of joint intentions in collaborative activities. Warneken [19] describes joint as well as individual intentions as follows:

> "*Roughly speaking, the notion of a person acting merely with an individual intention in the form of "I intend to do x by means of y", or recognizing another person's individual intention ("You intend to do x by means of y"), is sufficient to explain engagement in social interactions such as competition, parallel play, instrumental helping, or imitation. However, when two individuals engage in a collaborative activity, monitoring individual intentions is not sufficient; the partners also have to be aware that they are pursuing a joint goal, which both jointly intend to achieve in the manner of "We intend to do x together by means of me doing y1 and you doing y2"*.
>
> Warneken [19]

Supporting Bratman's guiding principle Levescque and Cohen propose joint intention theory [20]. A more detailed description is found in [20]. As Hoffman and Breazeal [21] stated:

> "*Their notion of joint intention is viewed not only as a persistent commitment of the team to a shared goal, but also implies a commitment on part of all its members to a mutual belief about the state of the goal*"
>
> Hoffman and Breazeal [21]

Joint Activity

The other way of exploring human–human collaboration is with the help of the concept of Joint activity [22].

> "*A joint activity is an activity carried out by an ensemble of people acting in coordination with each other*"

Clark [22]

Bratman's study of Shared Co-operative Activity (SCA) [16] identifies three essential features of joint activity. These features are:

- Mutual responsiveness: In SCA, each participating agent attempts to be responsive to the intentions and actions of the other, knowing that the other is attempting to be similarly responsive. Each seeks to guide his behavior with an eye to the behavior of the other, knowing that the other seeks to do likewise.
- Commitment to the joint activity: In SCA, each participants have an appropriate commitment (though perhaps for different reasons) to the joint activity, and their mutual responsiveness is in pursuit of this commitment.
- Commitment to mutual support: In SCA, each agent is committed to supporting the efforts of the other to play her role in the joint activity.

According to [23],[24], all participants in joint activity must:

- Enter into an agreement, that the participants intend to work together
- Be mutually predictable in their actions
- Be mutually directable
- Maintain common ground

According to Klein [24] for a joint activity, there are two primary criteria: one is the intention and commitment to take part within the joint activity; the other one is the participants' inter-dependence of action with the activity.

Common ground

Common ground is the base for inter-predictability and inter-dependence of actions with joint activity [25]. As mentioned in [21], a primary feature of a collaborative interaction is *common ground.* According to Clark [22]

> "*once we have formulated a message, we must do more than just send it off. We need to assure ourselves that it has been understood as we intended it to be.(...) For whatever we say, our goal is to reach the grounding criterion: that we and our addressees mutually believe that they have understood what we meant well enough for current purposes*"

Clark [22]

Clark [22], defined common ground as :

> “*the sum of ....mutual, common, or joint knowledge, beliefs, or suppositions*”
>
> Clark [22]

Common ground in teamwork inherits the principle of joint closure.

Shared plan

> “*To adequately model collaboration it is necessary that both to accommodate difference among the belief of individual participants and to distinguished between knowledge about the action performance and intention to act, agent differ not only in their beliefs about the ways to perform an action and the state of world, but also in their assessments of the ability and willingness of individual to perform an action.* ”
>
> Grosz and Kraus [26]

Grosz and Kraus’s theory stated that to be successful in collaboration, each participant need to have mutual beliefs about goals and actions, capabilities, intentions and commitment of the other participants. The original formulation of Shared Plans [26]was formulated for a model of collaborative planning in which it was not necessary for one agent to have intentions toward an act of other agent. This formulation was without introducing the notion of joint intentions. The original formulation of Shared Plans theory describe [26] :

- The notion of *intention-that* for agents to avoid adopting intentions that conflict with those arising from the group’s plan and engenders helpful behavior.
- The notion of Shared Plan itself and its possible partiality.

According to Grosz and Kraus [26], Bretmans’ description for intention is imperative for collaboration. Grosz and Kraususe use the terms “recipe”and “plan” to differentiate between “knowing how ” and “having a plan” to do an action.

Shared mental model

In interacting with the environment, with others and with artifacts of technology, people form internal mental models of themselves and of the things with which they are interacting [27]. According to Rouse and Morris [28] a mental model is a:

> "*mechanism whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states*"
>
> Rouse and Morris [28]

In literature, a number of works explain the functioning of a team by Shared mental model (SMM) theory. Within research for team cognition, shared mental models are the most studied hypothetical construct to explain certain human behaviors [28]. According to [29], the shared mental model theory presents an explanation of how a team can quickly and efficiently adjust their strategy "on the fly"–what are the mechanisms of adaptability for a team. A mental model provide three key roles: "help people to describe, explain, and predict events in their environment" [29]. Cannon-Bowers *et al.* [30] describe types of shared mental models in teams. In their work, Mathieu *et al.*[29] have shown that the similarity of knowledge structures between two team members can better predict the quality of team processes and performance.

2.2.2 Human Robot Collaboration: Related works

Integrating robots into human teams is challenging in many aspects. This subsection addressed how these challenges are addresses in literature. The term agent here personify any robotic or agent system.

There are lines of research on human-agent collaborative team work. Collagen [31] is a collaborative agent that has adopted the principles that underlie human collaboration from research in discourse theory and shared plans [31]. R-CAST [32] is a multi-agent system that supports human decision-making in a team context by using a collaborative RPD process (Recognition–Primed Decision model) as part of the team members' shared mental model. Miller *et al.* [33] presented intelligent team training systems called Collaborative Agent architecture for Simulating Teamwork (CAST). A life support control system based on collaborative agents and agent-human collaboration is presented in [34]. Miao *et al.* [35] presented a agent based system to train learners to handle abnormal situations while driving a car. For product design Hedfi *et al.* [36] developed a agent based negotiation architecture.

In work so far presented, the core design concept of collaborative agent were based on a comparison of tasks in which a human or agent is superior , which result in fixed allocation of tasks between human and agent. Billings [37], point out that Human and agents are not comparable, rather they are complementary. In this context, Brashaw *et al.* [38] summarize that the point is not to think so much about which tasks are best performed by human and agents but rather how tasks can be best shared and performed by both humans and agents working in concert. Hanna and Richards [39] identified two important factors for a collaborative agent with human–agent team:

- An understanding of factors that manage human social collaboration.
- Extending the abilities of agents with cognitive capabilities, social and affective characteristics.

Sycara and Lewis [40] identified three primary roles played by agents interacting with humans:

- Agents supporting individual team members in completion of their own tasks
- Agents supporting the team as a whole
- Agents assuming the role of an equal team member

Recently, shared mental models based on shared belief map [41] is proposed for realistic shared expectations among human-agent team member. It is based on the fact that the establishment of shared mental models among human and agent team members is a crucial step to advance human-centered teamwork research. Hidden Markov Models are used to capture cognitive load in a way that can be used in team contexts to make predictions about other team members' workload [41]. In their work, *Collaboration in human-robot teams*, Hoffman and Breazeal have used Bratman's SCA [16] to develop an agent that meshes human and robot sub-plans [21].

Human-agent teams have been used and explored extensively in search and rescue operations. Search and rescue exercises have shown that a team with an effective SMM are nine times more likely to find victims [42]. During the aftermath of the World Trade Center disaster small mobile robots collaborated with humans in order to locate and rescue victims[43]. The rescue operations revealed that both humans and robots contributes with unique qualities to the team.

Reasoning about "the human awareness" is an important instance of human-robot collaboration [44]. In their works, Alami and his group [44] introduce a framework to manage shared knowledge for a robotic system dedicated to interactive task achievement with a human. One of their papers [45] addresses high level robot planning in collaboration with human. Task planner called-Human aware task planner is designed for human robot collaborative task achievement. Human aware motion planning is also important for robot to collaborate with human [46]. Joint intention theory is the source of inspiration for Alami and his group [46] [44][45].

> *"Many tasks are parallelizable or involve complex interactions with objects in the environment, and can be more efficiently completed if human and robot collaborate. When robots work alongside humans for performing collaborative tasks, they need to be able to anticipate human's future actions and plan appropriate actions.."*
>
> Koppula *et al.* [[47], p.3]

The above quote is taken from the recent work of [47] where in Koppula *et al.* model the human's and robot's behavior through Markov Decision Processes. Simultaneous motion for human-robot collaboration has been studied in terms of hand-over tasks. [48] demonstrate that robots and people can effectively and intuitively work together by directly handing objects to one another. In [49], a framework was presented that allows a human and a robot to perform simultaneous manipulation tasks safely in close proximity. The proposed framework was based on early prediction of the human's motion.

Human intention can be inferred from his observed actions by modeling human behavior in the Partially Observable Markov Decision Process (POMDP) transition function [50]. In Oracular Partially Observable Markov Decision Processes (OPOMDP) [51], it is supposed that human is available permanently to provide the information. Whereas, HOP-POMDP (Humans Observation Provider POMDP) [52] and MI-MDP (Mixed Markov Decision Processes) [53] study the likelihood that the human is available and propose a decision-theoretic framework called Mixed-Initiative Markov Decision Process (MI-MDP). In this model, a robot and a human operator can each control a process using different sets of actions. Each action is characterized by a probabilistic transition model and its cost. In addition to the state of the environment, MI-MDPs also maintain a state variable that determines the current level of autonomy. The transition between fully autonomous operation and human control can be made over multiple time steps, during which the operator increases the level of attention paid to the system. [54] used "Learning the Model of Humans as Observation Providers" (LM–HOP) POMPDs to learn the accuracy of human availability. [55] present a framework for automatically learning human user models from joint-action. It is based on a mixed–observability Markov decision process (MOMDP). [56] address how robot could generate collaborative plans involving a human.

2.2.3 Human-Robot Collaboration: Existing approaches

According to [21], collaboration is often viewed as a control or a communications problem.

Collaborative control

Collaborative Control is an approach proposed by Terry Fong [4]. According to Fong collaborative control :

> "allows robots to benefit from human assistance during perception and cognition, and not just planning and command generation"
>
> Fong [4]

In his works, Fong [4] argue that working partnership between a human and a robot should be viewed as collaborative control.

Mixed initiative

Mixed-initiative approach focus on "initiative" shifts between human and agent. Mixed-initiative approach is associated to agent tasks. An agent can take a degree of "initiative"based on context and even the difficulty of the task at hand. James Allen defines the term "mixed-initiative" as

> *"a flexible interaction strategy, where each agent can contribute to the task what it does best"*
>
> Allen *et al.* [57]

In Allen's work, the system is able to reason about which party should initiate action with respect to a given task or communicative exchange. The original concept of mixed initiative interaction contributes the valuable insight that joint activity is about interaction and negotiation, and that dynamic shifts in control is useful. Mixed-initiative use a control architecture that allows agent to have different levels of autonomy. It can be in tele-operated, safe mode, shared control, collaborative task mode (CTM) and totally autonomous mode. In their work, [58] used planner based mixed initiative approach in search and rescue scenario. Its architecture is based on model based execution monitoring (activities model defined) and a reactive planner monitors task execution using that model. If the human operator changes execution order, planner responds by proposing a new execution order. Adams *et al.* [59] defines and develops a mixed-initiative human–robot collaborative architecture in which affect-based sensing plays a critical role in initiative switching. Adams *et al.* applied Riley's mixed-initiative interaction [60] to the architecture development. Acosta *et al.* [61] describes emotion based planning for mixed initiative interaction. According to [62] current trend of mixed-initiative interaction is moving away from the initial goal of a "flexible interaction strategy"; currently inclined to focus on task or authority assignment [63][64].

Human-Centered Automation

Human-centered automation is where the human is allowed the ultimate responsibility for system safety.

> *" Human-centered automation is an approach to realize work environment in which humans and machines collaborate......"*

Inagaki [65]

Human-centered automation was originally defined for aviation. In his pioneering work, Billings [66] stated the concept of human-centered automation as:
the human bears the ultimate responsibility for safety of aviation system

- The human must be in command.
- To command effectively, human must be involved .
- To be involved human must be informed.
- Functions must be automated only if there is a good reasons for doing so.
- Automated system must be predictable.
- Automated system must be able to monitor the human operator.
- Each element of the system must have the knowledge of the others' intent.
- Automation must be designed to be simple to learn and operate.

However, it was argued that human-centered automation is domain-dependent and thus vary depending on transportation modes [65].

Shared control

Shared Control is an early approach that appeared from the traditional task assignment view [67]. Shared Control was originally described as

> *''a supervisory with respect to control of some variables and direct controller with respect to other variables "*

Sheridan[68]

Recently, it has been described as blended Shared Control [67]:

> *...shared control as a control scheme that causes the output to be influenced ... by a set of two or more entities... be a human agent or operator and an autonomous electronic agent or robot.*

Enes and Book[67]

2.2.3.1 Human Robot Collaboration: Cognitive Architectures

> *Most of the works on collaborative agents have roots in the BDI logic or assume that the agents are BDI controlled.*
>
> Ciger[69]

As quoted, background of cognitive agent such as BDI is worthy.

Background on Cognitive Architectures

> *"Cognitive agents are intelligent human-like systems developed using insights from cognitive science. They represent computational models of human thought and reasoning: they perceive information from the environment, assess situations using knowledge obtained from human experts, and act to affect the external or internal environment."*
>
> Fan *et al.* [41]

Cognitive architecture as Trafton *et al.* [9] states:

> *"A cognitive architecture is a set of computational modules that, working together, strive to produce human level intelligence. The modules are typically designed to emulate different components of human cognition and can be tightly or loosely based on what is known about human cognition and how the brain functions. A cognitive architecture executes cognitive "programs," also called cognitive models."*
>
> Trafton *et al.* [[9], p.2]

Figure 2-1 is taken from the latest review of Chong *et al.* [1]. The figure shows the six cognitive architectures that they had studied and classify according to their roots and emphasis. Soar [70], [71], one of the first cognitive architectures , has its root in the classical artificial intelligence [72]. ICARUS [73] is developed with the primary aim of producing artificial intelligence mimicking human cognition . Adaptive Control of Thought-Rational (ACT-R) [74] is a cognitive architecture: it reflects psychological theories about human cognitive processes. ACT-R–the theory-was developed by John R. Anderson and colleagues. CLARION [75] is a hybrid model integrating both symbolic and connectionist information processing. Belief-Desire-Intention (BDI) is based on theory of intentional systems [76] and human practical reasoning [77] theory.The most recent advances within cognitive architecture focus on supraarchitectural capability integration and incorporation of emotional processes [78, 79]

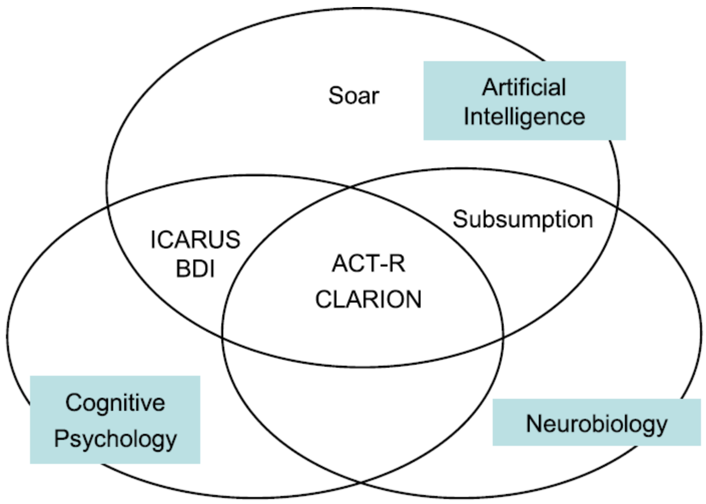

Figure 2-1: An overview of cognitive architectures [1]

Background on BDI architecture

Beliefs-desires-intentions (BDI) architecture [80] is described as a philosophical theory of the practical reasoning in [17], where the human reasoning is explained using: beliefs, desires and intentions.

> *"BDI architectures use Beliefs, Desires and Intentions to represent an agent's mental model of information, motivation and deliberation. Beliefs represent what the agent believes to be true, such as the location of an object and the distance to that location. Desires are what the agent aims to achieve, these are usually represented as a desired set of beliefs..."*
>
> Stocker [81]

The BDI model assumes that actions are deduced from a process which is called practical reasoning.

> *"Human practical reasoning appears to consist of at least two distinct activities. The first of these involves deciding what state of affairs we want to achieve; the second process involves deciding how we want to achieve these states of affairs. The former process - deciding what states of affairs to achieve - is known as deliberation. The latter process - deciding how to achieve these states of affairs we call means-ends reasoning."*
>
> Wooldridge [82]

In [83],[80], it explains why three components are essential: Belief is necessary for the usual reasons of the necessity for representation–as the "informative" component of a system state. A desire is thought of as a goal–the "motivational" state of the system. The system must represent the "currently chosen course of action" called system's intention: the "deliberative" component of the system state.

The three mental concepts that are part of the BDI model [17] are described as follows:

- Beliefs: They symbolize the knowledge of the world. They store all the sensors data and combine it with the agent's view of the world.

- Desires: They can also be called Goals. They save information about the objectives to be achieved and also about the priorities for each of them. Thus, they can be seen as the representation of the motivational state of the system.

- Intentions: They symbolize the current selected action plan. They can be seen as the deliberative component of the system. Therefore the agent chooses an intended action that will satisfy its desires given the current beliefs

Computational systems such as BDI agents "*provide the essential components necessary to cope with the real world*," [[83], p. 2].

The BDI architecture has its criticisms. As mentioned in [81]: *classical decision theorists and planning researchers question the necessity of all three mental attitudes while sociology and distributed artificial intelligence researchers question the adequacy of using only three.*

Wooldridge [82] summarizes the roles of intentions as follows.

- Intentions drive means-ends reasoning. They are pro-attitudes: if I have an intention, I will formulate plans to achieve it.

- Intentions persist. Intending implies committing for an extended period. However, an intention should be dropped by a rational agent if it is believed already achieved, believed to be unachievable or believed that its post-condition is no longer required. An agent will keep on trying to achieve an intention for as long as it is reasonable, even after failure to achieve it.

- Intentions constrain future deliberation. An agent should not entertain options (desires) that are inconsistent with current intentions.

- Intentions influence beliefs upon which future practical reasoning is based. Intentions must be consistent with beliefs.

Intentions can be formed, maintained and modified. Intentions are maintained by means of a commitment strategy.

> *As the agent has no direct control over its beliefs and desires, there is no way that it can adopt or effectively realize a commitment strategy over these attitudes [beliefs and desires]. However, an agent can choose what to do with its intentions.*
>
> Rao and Georgeff [[80], 315–316]

There are three commitment strategies [80], [83], [83]:

- A blindly committed agent keeps an intention until believed to be achieved. Agent keeps on trying to achieve its intentions, whether they are believed possible or not.
- A single-minded agent keeps an intention until believed achieved or until believed impossible, acceding to any beliefs that would indicate the impossibility of achieving the commitments, that is, allowing changes in beliefs to cause it to drop some commitments, but remaining unaffected by changes in desires
- An open-minded agent keeps an intention until believed achieved or impossible or until it is no longer a desire.

Agent's planning capability is represented by a plan function. According to Wooldridge [[84], p. 30] ...... *there is nothing in the definition of the plan function which requires an agent to engage in plan generation–constructing a plan from scratch. In most BDI systems, the plan function is implemented by giving the agent a plan library. [p. 30]*

Formal model of BDI

The BDI architecture includes seven main components[85]

- A set of beliefs (*Beliefs*), representing the agent's information about its current environment.
- belief revision function (brf), is a mapping from a belief set and percepts into a new belief set:

$$brf : \wp(Beliefs) \times P \rightarrow \wp(Beliefs)$$

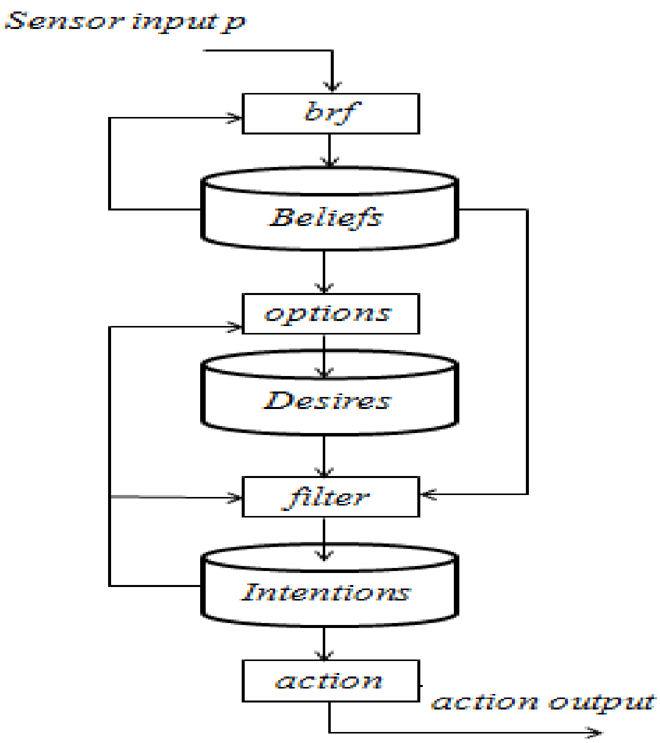

Figure 2-2: The BDI architecture

- A set of current desires (Desires), representing possible courses of actions available to the agent
- The options generation function (options) maps a set of beliefs and a set of intentions to a set of desires:

$$options : \wp(Beliefs) \times \wp((Intentions) \rightarrow \wp(Desires)$$

- A set of current intentions (Intentions), representing the agent's current focus.
- The deliberation process, i.e. deciding what to do, is represented by the *filter* function

$$\wp(Beliefs) \times \wp(Desires) \times \wp(Intentions) \rightarrow \wp(Intentions)$$

Which updates the agents intentions on the basis of its previously-held intentions and current beliefs and desires. It must drop intentions that are no longer achievable, retain intentions that are not yet achieved and it should adopt new intentions to achieve existing intentions or to exploit new opportunities. A constraint on filter is that it must satisfy current intentions which must be either previously held intentions or newly adopted ones.

- The function *execute* is used to select an executable intention, one that corresponds to a directly executable action:

$$execute : \wp(Intentions) \rightarrow A$$

The state of a BDI agent is at any moment a tuple $< \mathcal{B}, \mathcal{D}, \mathcal{I} >$, where $\mathcal{B} \subseteq Beliefs$ $\mathcal{D} \subseteq$ $Desiers$, $\mathcal{I}$ $\subseteq Intentions$.

2.2.3.2 Extended BDI architecture

A number of works on extension of BDI have been reported in the literature.

EBDI is an extension of BDI architecture supporting emotion [86]. EBDI is an extension of BDI architecture, which specifies a separate emotion mechanism within BDI agent. It implements practical reasoning techniques separately from the specific emotion mechanism. [87] extended the BDI model using the OCC (Ortony, Clore, Collins) model of emotion.

Myers and smith presented a BDI based framework for a cognitive agent that acts as an assistant to human user [88]. Busetta *et al.* [89] report an architecture, TOMAS (transaction Oriented Multi Agent System), that is extension of BDI framework with mobility. Coo-BDI [90] is a BDI extension for agent-agent cooperation. CooBDI, overcomes limitations of existing BDI agent by introducing co-operations among agents to retrieve external plans for achieving desires. Coo-BDI extends plans and extends intentions to take into account the mechanism for retrieving external plan instances. In [91], a human-aware belief-desire-intention (BDI) agents is presented and discussed. The framework monitors human user state, both physical and psychological to achieve an extensive human-context-awareness. The work of Padmanabhan [92] focuses on enhancing basic BDI logic with capabilities, opportunities and results. In his work, he describe a formal relationship between the Result, Opportunity, Belief, Desire and Intention modalities. Panzarasa *et al.* [93] modeled sociality in BDI framework. Norling [94] integrate a psychological decision making model to BDI framework. Norling examined the integration of Klein's Recognition-Primed Decision framework into BDI agent framework. The work focused on using reinforcement learning to enable agents to recognize the subtle cues that distinguish one situation from another. More recently, Singh *et al.*[95] proposed a framework that adds learning for improving plan in BDI agent.

There are reports specifically on extending BDI with a form of planning. Walczak *et al.* [96] investigated the requirements for the planner and for the coupling with a BDI system; introduced an approach where a BDI system takes responsibility for plan monitoring and re-planning and the planner for the creation of plans.

Walczak *et al.*'s work augment a BDI framework with a state-based planner for operational planning in domains where BDI is not well applicable. A fast state-based planner utilizing domain specific control knowledge retains the responsiveness of the system. In [97], it is noted that with BDI approach, an agent can reason over several goals, the agent lacks some flexibility by not being able to generate suitable plans on demand. The work describes the approach for combining BDI theory with a partially observable Markov Decision Processes (POMDPs) planner. Meneguzzi *et al.* [98] note BDI agents often rely on logic models that assume infinite computational power, while BDI architectures defined for run time efficiency have curtailed an agent's autonomy because of its pre-compiled plan library. But they also note the disadvantage of the inflexibility of pre-compiled plans compared to the relevance of real-time generated plans.

2.2.4 Section Summary: Context of Present Work

The work reported in this thesis is inspired by concept of the Shared Plan. Here the works of Lesh *et al.* [99] is worth mentioning.

> *" collaboration requires the participants to have substantial mutual understand- ing of their shared goals and the actions intended to achieve them (these are part of what Grosz and Sidner .. call the SharedPlan). One way to maintain this mutual understanding is through verbal communication.....However, it is often more efficient and natural to convey intentions by performing actions....".*
>
> Lesh *et al.* [99]

As pointed out by Lesh *et al.* [99], in this thesis the quote is interpreted as: agent maintains mutual understanding by performing actions.

This thesis views human-robot collaboration as an approach on "initiative" shifts between human and agent where by use of a control architecture, agent is allowed to have different levels of autonomy. Work on human centered automation inspired the ideas that a. human must be the ultimate authority of command and b. agent should have the knowledge of human intent.

The work of Alami and his group [100][46] [45] focused on one interesting aspect that collaboration between robot and humans can be greatly improved if the robots has the awareness of human. Our approach differs from their work in many aspects. Even though, their work inspires us, however our focus is that human centric behaviour is important for collaboration.

The reported work on extended BDI enabled us to have valuable understanding that BDI does not have the depth of planning capabilities required for collaboration. In our view, when the BDI agent is part of a human team, teamwork demands agent to perform the activity jointly. To include the mechanism necessary for collaborative planning, such as sharing plan with human partner, extension to the basic BDI architecture is required.

This work contribute to the idea that if a human and an agent approaches the same task using same strategy (for the agent if strategy is human centric strategy), then agent may more efficiently collaborate with human. This is attempted by extension to BDI architecture.
So, why BDI architecture other than Soar, Act-R?. This work find the following observation of Lee *et al.* [101] worth mentioning:

> *"..........Soar and Act-R concentrate on the actual mechanisms of the brain during information processing, including tasks such as reasoning, planning, problem solving, and learning. Consequently, these models become complex and difficult to understand. On the other hand, the core concepts of the BDI paradigm, originally based in folk psychology, allow use of a programming language to describe human reasoning and actions in everyday life............. ....Because of this straightforward representation, the BDI paradigm can easily map extracted human knowledge into its framework.*
>
> Lee *et al.* [101]

2.3 Human Robot Collaboration: The Robot

This thesis looks at how a system is to be designed for teaming with human who can act independently (of the system) and are often satisfied with a good solution (which may not be optimal), a phenomenon that is called *satisficing* [8]. It is worth noting that according to Burghart *et al.* [102], designing a system to teaming with human demands *cognitive stance*. Work of Burghart *et al.* also discussed on requirement of control architecture to perform a range of cognitive functions. It is Alan C. Schultz and the research team at the Navy Center for Applied Research in Artificial Intelligence (NCARAI) on cognitively enhanced control that has inspired this research [103] [104] [9].

There has been a lot of interesting work of Alan C. Schultz and NCARAI research team exploring how cognitively enhanced control enhance interaction with humans. Section 2.3.3 will describe some of their work.

2.3.1 Background on Robot Control Architectures

This subsection begins with definitions of what is a robot architecture. For Bonasso [105] architecture is "*the arrangement of control software for the robot*" [[105], p. 193].
For Gat [106] architecture in robotics means "*a set of constraints on the structure of a software system,*" [[106], p. 210]. Bekey [[107], p. 98] defines architecture as "*the practical structure of a robot's software ... its goal is to define the way in which sensing, reasoning, and action are represented, organized, and interconnected*"

Bekey [107] mentioned a connection between control and architecture as :

> *"An architecture provides a principled way of organizing a control system. However, in addition to providing structure, it imposes constraints on the way the control problem can be solved".*
>
> Bekey [[107] p. 99]

The common robot control architectures are divided into the following groups:

- Deliberative or hierarchical control architectures
- Reactive control architectures
- Hybrid control architecture

Deliberative or hierarchical control architectures

Deliberative or hierarchical control architectures are based on the *"sense-plan-act" (SPA)* principle. Deliberative or hierarchical control architectures require full environment information for their optimal functioning. The control is goal based.

This kind of architecture is structured in layers. Robot with this type of architecture first senses the environment, then plan and finally executes actions. Each step is performed at the corresponding layer. According to Arkin [108] deliberative robotic systems have the following common characteristics:

> *"They are hierarchical in structure with a clearly identifiable subdivision of functionality. Communication and control occurs in a predictable and predetermined manner, flowing up and down the hierarchy with little if any movement. Higher levels in the hierarchy provide sub-goals for lower subordinate levels. Planning scope changes during descent in the hierarchy. Time requirements are shorter and spatial considerations are more local at the lower levels. They rely heavily on symbolic representation of world models."*

Arkin[108]

Hierarchical control is well suited for structured and highly predictable environments, but is inappropriate for dynamic environments which require timely responses.

Reactive control architectures

Reactive control architectures use direct predefined reaction mapping from sensor to actuator. Reactive robotic systems consist of collection of rules that map specific situations to related actions. Reactive robotic systems are very fast in motion and computations. That is why for real world situation where the "*reaction time*" is a very important factor, a reactive robotic system is the answer. Reactive robotic systems have the following characteristics:

- These systems are based on the models of animal behaviour.
- The control is stimulus-response based.
- Each behavior in these systems consists of a sensor-motor pair.
- Avoid the use of abstract representational knowledge.
- Purely reactive architecture do not have the ability to learn.

The best known reactive control architecture is the Subsumption Architecture, introduced by Rodney Brooks [109].

Hybrid control architecture

As stated in Kortenkamp *et al.*:

> *We might call this approach P-SA; that is, the robot plans based on initial conditions and common knowledge (P) and then executes this plan using sense-act (SA) behaviors, re-planning only when the reactive behaviors run out of routine solutions. [[110], p. 12]*

The P-SA approach is the hybrid deliberative/reactive architecture.

Nowadays the robotic community believes that one way to mitigate the limitation and drawbacks found in Deliberative and Reactive architectures is to combine both architecture in a single Hybrid Deliberative/Reactive architectures.

The hybrid architecture uses deliberative planning to control the lower level of reactive components. As reviewed by Gat [106]:

> *" ....the three-layer architecture....has now become the de facto standard," [[106], p. 198].*

"By far the most popular hybrid architecture is the three-layer architecture, which consists of a reactive layer, an executive layer, and a deliberate layer," [[111], p. 933].

According to [106], the components of hybrid architectures are:

- Reactive layer: the bottom layer implements low level functionality that is used to deal with sensors and motors. Controller components have low computational complexity. Controller components are written as C modules [112].
- Executive layer or Sequencer layer: sit above the controller layer. This layer has to Select a primitive behavior at a given time under a situation. This layer manages multiple parallel interacting tasks. The most common method to implement Sequencer layer is to use a Conditional sequencer language [112].
- The Deliberate layer or Planner layer, relies on heaviest computational components. The planner can be implemented in standard programming languages [112].

The general approaches of communication and interaction within layers are:

- Deliberative layer–sequencing layer: the deliberative layer interacts with the sequencer in two basic ways:
  - the deliberative layer produce plans for the sequencer to execute
  - the deliberative layer act in response to request from the sequencer
- Sequencing layer–reactive layer: Interaction between these two layers involve (a) activation; (b) deactivation or termination; (c) passing of control parameters; and (d) monitoring of success, as well as conversion of non-symbolic to symbolic parameters and vice-versa [113].

2.3.2 Examples of selective collaborative architectures

The HRI/OS

HRI/OS [114] is a framework developed in NASA's peer-to-peer human-robot interaction project. Direct point-to-point communication is performed using ICE middleware. Tasks are implemented using the Task Description Language (TDL).

Architecture of HRI/OS is agent-based: it incorporates embodied agents (humans and robots) and software agents. A major difference compared to other architectures is that there is a human as a key member of the system.

The MII human-robot collaborative architecture

The MII human-robot collaborative architecture [59] is a mixed-initiative human–robot collaborative architecture. In MII affect–based sensing plays a significant role in initiative switching. Architecture is based on Riley's [60] model of mixed-initiative interaction. MII is organized as three blocks: HRI block-where the robot and the human interact with each other. The robot and the human dynamically modify the goals and constraints by interacting via the HRI. Planning block–responsible for outline and validation of the given mission with the human and the robot. The execution block executes plan.

DH↔DR collaborative architecture

DH↔DR collaborative architecture [115] has been developed for a human operator to interact with a remotely located robot. The architecture is based on the concept that :

> *“ what is Diffcult or Dangerous for the Human (DH) will be Done by the Robot (DR)" and what is Difficult and Devious for the Robot (DR) would be better Done by the Human (DH)”*
>
> Paolo [[115], p.117].

DH↔DR is organized in four basic modules: A Perception module for vision and sensory perception. A Shared decision making module to represent tasks for the robot as well as in charge of switching between an autonomous task and a supervised task; the Hardware abstraction module to convert high level task references in motor command and the Communication module responsible for establishing a collaborative dialog between the human operator and the robot .

The LAAS architecture

LAAS[1] architecture for autonomous systems is a three layer architecture [100]: Decisional layer: take charge of planning and supervision of actions. It includes: a procedural reasoning system, a high level task planner and a knowledge base.

[1] Local Area Augmentation System

An execution layer is responsible for action coordination. A functional level that is composed of set of modules to integrate all operational functions.

LAAS architecture is a representative example of hybrid control systems based on several paradigms. The different agents to sense (perception modules), model (mapping modules), plan (planners) and act (actuator modules) indicate the adherence of the SPA paradigm.

Decisional architecture for human–robot interaction

[116] devised a control architecture dedicated to robot decision and action in a human . The architecture has been developed as an instance of the generic LAAS architecture. The architecture consists of four components: 1. Task Agenda: a mechanism for robot higher level goal management. 2. Chronicle recognition system : An chronicle recognition system for modeling and recognizing scenarios. 3. Human aware task planner: task planning system to synthesize socially acceptable plans. 4. SHARY–constitutes decisional kernel.

2.3.3 Cognitively enhanced control

The studies of human robot collaboration at NCARAI mainly concern on "collaboration in a shared workspace". In his paper *Using Computational Cognitive Models to Build Better Human-Robot Interaction*[103], Alan C. Schultz, illustrated why a cognitively enhanced artificial system is important for human robot interaction. According to him, adding computational cognitive components to intelligent systems result in three benefits [103]:

- Giving the system cognitive models can enhance the human system interface by allowing more common ground in the form of cognitively plausible representations and qualitative reasoning.
- By incorporating cognitive models, we can develop systems whose behavior is more expected, natural and therefore compatible with the human team members.
- An intelligent system that is cognitively enhanced can be more directly compared to human-level performance

Alan C. Schultz and the research team at NCARAI addressed the importance of cognitive and computational theories. The studies on human robot collaboration at NCARAI is based on developing computational cognitive models (CCMs) of certain high-level cognitive skills humans possess and that are relevant for collaborative tasks. They use CCMs as reasoning mechanisms for agents. [117],

[118],[119], [120] and [9] are based on the belief that giving the robots similar representations and reasoning mechanisms to those used by humans, robots will that act in a way that is more compatible with humans. Trafton *et al.*'s [9] work is based on cognitive architecture ACT-R.

> *" ...We rely on the computational cognitive architecture ACT-R [74]. A cognitive architecture is a process-level theory about human cognition. For the purposes of HRI, a cognitive architecture can imbue a robot with a model of the mental state of a human teammate which can be exploited in any number of ways."*
>
> Trafton *et al.*[[9] , p.1]

In their related work, Trafton *et al.*[104] outlined a corollary for building agent:

> *" To perform collaborative tasks with humans in physical settings, a robot must be able to simulate and reason about the world from the perspective of vantage point of others"*
>
> Trafton *et al.* [104]

Trafton *et al.* [104] explored the issue of perspective taking as a crucial element in human–agent interaction in the context of a shared task and further addressed the importance of mapping between cognitively inspired and engineering-oriented internal models. Trafton and his colleagues [104] presented a conceptual guideline for effective human–robot interaction design as:

1. Robotic representation, reasoning and perception mechanisms should be as similar to those of humans as possible.

2. Cognitive systems for human–robot interaction should be based on integrated cognitive architectures.

3. The use of heuristics and principles in collaborative activities similar to those ordinarily employed by people is consistent with people's expectations, and so, is consistent with effective human–robot interaction design.

This thesis is strongly motivated by the guidelines for effective HRI as mentioned:.

> *" Humans, even highly trained ones, are unpredictable, error-prone, and are susceptible to mental states like fatigue....... ....In this light, we believe that it is important that a robot understands what human teammates are doing not only when they do something "right," but also when they do something wrong................. "*
>
> Trafton *et al.* [[9], p.1]

2.3.4 Section Summary: Context of Present Work

This thesis is influenced by Alan C. Schultz's concrete illustration on why an artificial system needs to be cognitively enhanced [10]. Such an enhancement is approached through cognitive architecture. In line with what is been being propounded by Schultz and his group [9],[10], this thesis provides an architecture for cognitively enhanced control in which knowledge of human strategy plays significant role. The term "cognitively enhanced" in the title of the thesis is motivated from the works of Alan Schultz and his group[103].

We are strongly motivated by Trafton *et al.*'s guideline for effective HRI design [104] and observation that is made in [9]. This thesis interprets the guideline and remark of Trafton *et al.* as: a. to collaborate with human, a system needs to have the strategy library (The strategy library stores a selection of possible strategy that a human in the environment may execute). And
b. As stated in [9]

> *"With knowledge of how people might perform in different situations, the robot can use that knowledge to improve overall performance."*
>
> Trafton *et al.* [[9], p.21]

A key concept is the systematic investigation of human behaviour in scenarios that the robot is expected to perform. Make it possible to mimic these so that behaviour of the robot is consistent with human expectations.

2.4 Human Machine Collaboration: The Intelligent Wheelchair and Human User

As mentioned, one of the goals within human-Robot teamwork research is systematic investigation of human behavior. This thesis looks at real-world domain where intelligent wheelchair provides navigational help to a human user. This allows us to focus on human navigation. This section will focus on notion of human navigation behaviour. Before going to details on navigation behaviour the section starts with a background on intelligent wheelchair.

2.4.1 Background on Intelligent Wheelchair

Intelligent Wheelchair (IW) is a solution for supporting individuals who have disabilities and are thus unable to perform their daily activities. Intelligent wheelchair as we understand in this thesis is best described by the following definition.

> *"intelligent wheelchair" as a robotic device .... provide with sensory system, actuators and processing capabilities".*
>
> Petry *et al.* [121]

Several prototypes of IW have been developed and many scientific works have been published. In this area what has not been addressed so far is identified from the following:

> *"First, we would suggest that the aim of creating autonomous wheelchairs is in itself not valid. The most intelligent agent involved with the powered wheelchair is the driver, not computer controlling sensors and motors, and so the most important design aim should be to develop systems which complement, maximize and augment the pilot's skills, not replace them".*
>
> Nisbet [122]

The above quote give a view that the user of a wheelchair is a more important factor in its design consideration.

Shared Control in Intelligent Wheelchair navigation

To express what is shared control in context of wheelchair navigation,we go with the view of Vanhooydonck [123] :

> *"For the case of a wheelchair, this cooperation between man and machine can be compared to the cooperation between a horse and its rider: the rider navigates and performs the high-level or global planning (the coarse control) by indicating to the wheelchair where (s)he wants to go; the horse/wheelchair avoids (small) dangerous obstacles and performs the low-level or local planning (the fine motion control). In other words: the intelligent controller should take over the autonomy that the user lacks (hence Shared Autonomy or Shared Control)"*
>
> Vanhooydonck *et al.* [123]

In shared control approach, user has an important role in the decisions for navigation. According to Vanhooydonck *et al.* [123], for a shared control system the vital questions in design is:"who gets control over the system when, and to which extent?"

Poncela *et al.* [5] view shared control as:

> *"It is important to note that wheelchair navigation strongly differs from other human-robot shared control approaches, in the sense that persons are actually riding the robot and, hence, have first-hand feedback on resulting actions, and also because a high collaborative profile is desirable to avoid loss of residual skills, whereas in other situations it is usually better to reduce the number of interactions with the robot to reduce the cognitive burden of driving the machine."*
>
> Poncela *et al.*[5]

Poncela *et al.*[5] identified following three existing shared control approaches in wheelchair navigation:

- Safeguard navigation
- Behaviour based shared control
- Deliberative shared control

In most of safeguard navigation approaches control is left mostly to its user, and automatic navigation is only triggered if particular situation is identified. Most of the works on safeguard navigation approaches used reactive algorithms.

MAID[124], NavChair [125],Omni [126],TinMan [127], Smartchair [128], Wheelesley [129], VAHM [130] Rolland III [131],[132], [131], [133] follow behaviour based shared control approach where some of the behaviour such as *avoidobstacle*, *followfall,* and *Passdoorway* are activated manually or automatic-triggering. Based on behaviour based approach, NavChair, selects its behaviour based on environment information- however it ignore user intention for its particular environment information. To avoid obstacles, NavChair uses a vector field histogram method.

SENARIO [134] follow deliberative shared control approach. The system uses a high level planner control local navigation layer. Rolland III [135] uses a dynamic window approach for local navigation and deliberative navigation relies on Voronoi diagram.

In literature, some works contribute to user intention estimation for collaborative control. Partially observable Markov decision processes (POMDPs) is used by Taha *et al.*[136] to model long–term goals and intention of of user . More recently Carlson [7] contributed to human factor analysis by defining intention prediction functions. SIAMO [137] models the environment by means of Visibility Graphs and local navigation is based on Potential Fields Approach. In [7], a collaborative mechanism was proposed to assist its users when help is required from them. The system uses a multiple–hypothesis method to predict the driver's intentions and, if necessary, adjusts the control signals to achieve the desired goal. [138], [139] reports on Bayesian approach in assistance for wheelchair. Parikh *et al*'s [140] work is based on the expansion of the most optimal solutions to meet the goal. Urdialest *et al.* [6], presents a shared control navigation approach as they called collaborative control [141]. Works of Poncela[5] and Urdialest *et al.*[6] is based on reactive control approach. More recently, Lopes and co-researcher presents an two-layer collaborative control approach.[142]. This is a context based approach of control. In their work, Vanhooydonck *et al.* [143] presented framework that is based on continuous estimation of its users intention. Recently, Vanhooydonck *et al.* [143] argued that

> *" existing approaches towards intelligent wheelchairs do not always grant the wheelchair user a sufficient degree of importance during the design of the shared control system".*
>
> Vanhooydonck *et al.*[143]

They identified a set of approach towards shared control for intelligent wheelchair:

- Assistance for daily manoeuvres.
- Safety and robustness.
- No modifications to the environment.
- Two-way communication.
- Intuitive control and comfort.
- Feeling in control.
- User intention.
- Adaptability.

This thesis accepted the view of Vanhooydonck *et al.* [143] and find" feeling in control" is one of the important factor in shared control of wheelchair navigation.

2.4.2 The Human User: The Wayfinder

An intelligent wheelchair assist the user to travel through the environment to reach a desired destination, safely and efficiently. Wheelchair navigation is different from conventional mobile robots [144]. Intelligent wheelchair navigation can be imagined as the robot and its human user sharing a space which is common and use various modalities to exchange information. While navigating, a wheelchair consider its user as a navigator; thus adapting human navigation strategies rather than following brute force navigation algorithms.

According to Montello [145], human navigation from one location to a given location consists of two components:

- First, wayfinding: high-level reasoning and interpretation processes.
- and second, locomotion: basic sensory-motoric processes.

In this work, we are only interested in wayfinding part of navigation.

Wayfinding may mean different things for different people! Wayfinding as we understand in this thesis is best described by the following definition.

> *"..... the process of determining and following a path or route between an origin and a destination. It is a purposive, directed, and motivated activity. It may be observed as a trace of sensorimotor actions through an environment."*
>
> Golledge [146]

2.4.2.1 Human wayfinding

To successfully way find, we use either environment information–"*knowledge in the world*" or our spatial knowledge about the environment– "*knowledge in the head*" [147].

Cognitive map

Representation of space has been subject of study for most navigation experiments. According to Weisman[148] we human must maintain a mental representation of spatial relationship among objects in the environment to find a way in real-world environment. This mental representation, also entitled as a cognitive map is exo-centric information for wayfinding. It was Tolman [149] who introduced the term cognitive map. In his work Tolman demonstrated the existence of map like internal (spatial) representations that reflect properties of the external world. However, it was soon realized that cognitive map is a collection of knowledge which is not complete and it is incoherent in type [3]. Hirtle and Jonides [150] and McNamara [151] works support hierarchies in spatial knowledge. Siegel and White [152] illustrated three stages of spatial knowledge development; assumed that these representations of spatial knowledge are acquired successively as the environment is experienced. Ishikawa [153] and Montello [154] criticized strict sequence as their individual empirical findings demonstrated that humans, children as well as adults acquire knowledge of all three types simultaneously. Furthermore many experimental study shows that human acquire exo-centric representation even after short exploration of an environment [155–157]. Tversky [158] illustrate two effects in acquisition of spatial knowledge: the alignment and the rotation effect. However, it still remains to be answered what constitutes a cognitive map? Lynch's [159] study gives an insight into the question. According to him cognitive map consists of five elements: 1) paths, 2) edges, 3) districts, 4) nodes and 5) landmarks. The understanding of a cognitive map within this thesis is very close to Lynch's above characterization.

Wayfinding aids

Whenever we have to find our way to a given unknown place we need wayfinding support. In this situation the wayfinder requires additional information to be able to plan a route and to reach the destination. If the place and a route are already known, the wayfinder can reason and plan with the information stored in the mind of the wayfinder, the so-called cognitive map or mental map. If the place or a route is not known to the wayfinder, the missing spatial information has to be completed with external information: the wayfinding assistance.

As long as origin, destination and a traversable route between them is known to way-finders, there is no need for assistance. In the absence of comprehensive or sufficient information about either the origin, the destination, or a complete route between the origin and the destination, we need assistance to find the way. External assistance is any form of given representation of the environment which is not retrieved from the mind of the wayfinder or perceived directly in the environment (such as following signage). Maps and route directions,verbal direction are examples for external assistance in form of representations of the environment and the route. In [160], they suggest verbal information to assist wayfinding activity. Thorndyke and Hayes-Roth [161] suggested exo-centric representations aid navigation and allow people to make accurate ego-centric judgments. They also make a remark on important differences between knowledge acquired from maps and from navigation.

2.4.2.2 Experiments for wayfinding strategy

Wayfinding through different experimental methodology has been addressed in the literature. Golledge on his wayfinding studies for The University of California Transportation Centre evaluated that wayfinder use various route selection norm in different environments and on different routes [162]. In one study, [163] investigated pedestrian shopping to determine shoppers behavior to minimize multi-stop shopping routes. To evaluate process of wayfinding in architectural spaces, [164] empirically evaluates the effects of complexity of plan on wayfinding. Through his wayfinding experiment in virtual urban environment, [165] concludes that wayfinder either select straightest route to the destination or most direct route on his route selection. To assess how people plan routes by using maps, [166] conducted four wayfinding experiments. Bailenson *et al.* [166] concluded that people use heuristics or strategies and this leads to asymmetries or systematic biases. To compare the complexity of buildings, Raubal and Egenhofer [167] has formulated a choice and clue model and verified their model by comparing two different airports. In another study, by taking an environment with and without landmark, [168] investigated the wayfinding behavior and spatial knowledge acquisition under a developmental perspective. Within a virtual reality simulation of a real city, [169] addressed cognitive processes of wayfinding. More recently, [170] introduce a new virtual environment call SQUARELAND to investigate cognitive processes in human wayfinding. In a simulation-based search and rescue scenario, Goldiez *et al.* [171] studied the effects of various augmented reality display settings on human wayfinding.

The existing experimental methodology in wayfinding are summarized in Table 2.1.

Table 2.1: Summary of experimental methodology in wayfinding

Experimental methodology	References
Map based	[162] [166] [172]
Virtual Environment	[170] [168] [169] [173] [165] [174]
Real-World Experiments	[171] [147] [167] [172]
Neurophysiology based Experiments	[169] [175] [176]

On wayfinding evaluation

Evaluation of wayfinding is based on empirical results. It was Weisman[148], who distinguishes categories of environmental variables that in some way affects wayfinding. These are:

1. visual access,
2. the degree of architectural differentiation,

3. the use of signs and room numbers to provide identification or directional information, and

4. plan configuration.

The works of [177], [164] confirmed Weisman's four categories of environmental variables. In addition to that, Raubal and Egenhofer [167] found that familiarity with the environment is another factor that influence wayfinding.

Metrics of Wayfinding

[178] categorized metrics of wayfinding. One of the category of wayfinding metric is to compare the resulting solution with a route which has been supplied by a human user for the same start and end points. The second category of wayfinding metric relates to the performance of a user navigating a route rather than the route itself. According to Ruddle and Lessels [179] wayfinding can be evaluated at three levels: task performance, physical behaviour and cognitive rationale.

Identifying strategy

There are varieties of ways to collect and identify strategies that humans employ when finding his ways.

Drawing of maps or plotting of routes is one of the most apparent ways of collecting wayfinding data. An individual's drawing of maps or plotting of routes produces a graphical indication of the environment and necessary movement.

Another method is through verbal or written descriptions. Ericsson and Simon [180], in their detailed study of the effectiveness of different categories of verbal protocols found that retrospective reports as well as thinking-aloud protocols provide true insights into mental processes.

Wayfinding Strategies

There is considerable research that tries to shed some light on the wayfinding strategies. In one study, T. Gärling and E. Gärling investigated pedestrian shopping to determine shoppers behavior to minimize multi-stop shopping routes [163]. Delton [165] concludes that wayfinder either select straightest route to the destination or most direct route on his route selection. Christenfeld [172] investigated human preferences of route choice in maps and in real world environments. In their experiment, it is found that human prefer routes with the longest initial straight segment. They called it as Initial Segment Strategy (ISS). To explain why people choose asymmetric routes, Bailenson [166] conducted four wayfinding experiments. According to Aginsky *et al.* [181] human follow either a *visually* *dominated* or a *spatially dominated strategy.* Interestingly, Murakoshi and Kawai [182] found that wayfinding strategies were less spatial in nature. Wiener and Mallot [183] concluded that wayfinders use fine space and coarse space connectivity. According to Wiener and Mallot human wayfinding is a fine-to-coarse process, which results in the hierarchical representation of spatial knowledge [183].

2.4.3 Section Summary: Context of Present Work

Even though shared control approach to wheelchair made some progress, this work agrees to Vanhooydonck *et al.* [143] and find "feeling in control" is the most appealing factor: the navigation controller supports what the user wants to do and help when required.

Collaboration in control for wheelchair navigation is largely motivated through Urdialest *et al.*'s observation

> *"In order to improve acceptance, we would need the robot to mimic the driver up to a point, so that users do not realize so easily that they are getting help."*
>
> Urdiales *et al.* [6, Page 190]

In our view, apart from above an IW helps human user with locomotion part of navigation while user is participating in wayfinding task. As stated by [184]:

> *".....wheelchairs are an excellent example of tight coupling between the desires of the operator and the robot. The primary challenge in techniques is to have the chair follow the desires of the operator while maintaining safety in navigation."*
>
> Zeng [184, Page 24]

This thesis does not strongly agree on what is being mentioned by [184]: chair should follow the desires of the operator while maintaining safety in navigation. In our view there must be a negotiation process: a means for settling on the command as suggested by operator and the command as suggested by the navigation controller of the wheelchair.

We propose to approach human-robot collaboration from the standpoint of mixed-initiative approach for collaboration, implying a sense of partnership that occurs when agents work "jointly with" human. To collaborate with user, wheelchair navigation controller must have the knowledge of human wayfinding. We propose to approach wheelchair navigation such that controller should not try to replace user abilities but, in contrast, should complement as team mate and use their available skills.

2.5 Chapter Summary

In this chapter, a back ground on the concept related to human robot collaboration is given. The chapter reviews the robot control architecture and the idea of the beliefs desires intentions paradigm was discussed. Towards the end of the chapter, intelligent wheelchair and human navigation are discussed. The focus of the thesis is also discussed in detail.