Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online
CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-19 (2016) 096–101

A hierarchical framework to analyze shared control conflicts between human and machine

Makoto Itoh*. Frank Flemisch, David Abbink*****

**University of Tsukuba, Tsukuba, Ibaraki 305-8573 JAPAN*
(Tel: +81-29-853-5502; e-mail: itoh.makoto.ge@u.tsukuba.ac.jp).
*** RWTH Aachen University / Fraunhofer FKIE, Aachen / Bonn, Germany*
frank.flemisch@fkie.fraunhofer.de
**** Delft University of Technology, Delft Haptics Lab, Mekelweg 2, 2628 CD Delft, The Netherlands (e-mail: d.a.abbink@tudelft.nl)*

Abstract: Shared Control, where the machine and the human share tasks and control the situation together, and its extension cooperative automation are promising approaches to overcome automation-induced problems, such as lack of situation awareness and degradation of skill. However, the design of Shared Controllers and/or cooperative human-machine systems should be done in a very careful manner. One of the major issues is conflicts between the human and the machine: how to detect these conflicts, and how to resolve them, if necessary? A complicating factor is that when the human is right, conflicts are undesirable (resulting in nuisance, degraded performance, etc), but when the machine is right, conflicts are desirable (warning the operator, or proper assistance or overruling). Research has pointed out several types and causes of conflicts, but offers no coherent framework for design and evaluation guidelines. In this paper, we propose such a theoretical framework in order to structure and relate different types of conflicts. The framework is inspired by a hierarchical task analysis, and identifies five possible sources of conflicts: intent, information gathering, information processing, decision-making and action implementation. Examples of conflicts in several application domains such as automobile, telerobotics, and surgery are discussed to illustrate the applicability of this framework.

Keywords: Automation, Shared Control, Human-Machine Cooperation, Conflict, Safety, Automobile, Aviation, Teleoperation.

1. INTRODUCTION

Based on the rapid development of highly intelligent machines, automation or assistance systems are becoming real in not only professional use but also daily use. However, design of control logic and human-machine interface of the intelligent machines should be done with care. Research has clarified that poorly designed automation may cause serious troubles. One of the essential problems with "automation" is that the human operator's situation awareness may easily be degraded because the human operator is out of the control loop (see, e.g., Sarter and Woods, 1995).

In order to overcome the automation-induced problems, several approaches have been proposed. One starting point for this was the H(orse)- metaphor (Flemisch et al., 2003), where the relationship between the human and the machine is compared to the relationship of a rider and a horse or of a driver and a horse cart. The rider and the horse are coupled through the "rein" so that the rider can be in the control loop, and the H-metaphor applies this relationship to the one between a human and a machine. A related approach is *haptic shared control* (see, e.g., Abbink et al., 2012) where the human and machine exert forces on a control interface, of which its output position remains the direct input to the controlled system (Abbink and Mulder, 2010). Concrete designs of haptic shared control have been realized for the driving domain - car-following (Mulder et al, 2008), lane-keeping and curve-negotiation (Griffiths & Gillespie, 2005; Flemisch et al., 2008; Mulder et al., 2012)_and evasive maneuvers (Mulder et al, 2011) – as well as for tele-manipulation (Boessenkool et al., 2011; van Oosterhout et al, 2015). The notion of shared control is not limited to using haptics, for example, Carlson and Miller (2013) proposed shared control designs for brain-controlled wheelchairs, where the operator's intent is estimated from brain activity and blended with an automated controller to navigate the wheel chair through its environment. This shows that shared control can assume many forms. Based on general concepts of human-machine cooperation (e.g., Rasmusssen, 1983; Hollnagel and Woods, 1983; Sheridan, 2002; Hoc, 2000), shared control has been expanded towards cooperative control (Flemisch et al., 2003; Hoc et al., 2006; Biester, 2008; Holzmann, 2007; Flemisch et al., 2008a; Hakuli et al., 2009. If cooperation goes beyond the control level towards guidance of maneuvers, Flemisch et al. (2011, 2015) speak of *cooperative guidance and control* - as part of a cooperative automation. Pacaux-Lemoine and Itoh (2015) called this a 'vertical extension' of the shared control concept, and related this to their 'horizontal extension' along the information-

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.468

processing dimension (Parasraman et al., 2000). For more detailed discussion on the connection between shared control and cooperative automation, see Flemisch et al. (2016).

When sharing and cooperating in task execution, one of the challenges is conflicts between the human operator and the machine. If the machine is wrong, conflicts are undesirable and can cause not only annoyance or discomfort but also dangerous situations. On the other hands, if the machine is right in correcting the human, conflicts are desirable and can help to prevent critical human errors. How to determine which agent is correct – and in which situations? The answer might be impossible to achieve, and perhaps a measure of undesirable conflicts is simply the price to pay for an intelligent partner. The authors argue that we should accept that conflicts can arise, and methodically and experimentally investigate how conflicts can be categorized and - if needed - resolved satisfactorily.

The main contribution of the current paper is a proposal for a general framework on conflicts between a human and a machine, focussed on shared control applications.

2. WHAT ARE CONFLICTS AND WHAT TYPES EXIST?

2.1 *Definition of Conflict*

To the best of the author's knowledge, no clear definition of conflict in human-machine cooperation has been proposed. A human-centered definition of conflict is that *a conflict is defined as occurring as long as the human control input is not consistent with the expected control input of the (intelligent) machine*. A typical example of a conflict for an automotive shared control application is illustrated in Fig. 1. Suppose a human is driving a car that is equipped with haptic shared control. During curve negotiation, the intended trajectories might be different between the human and the machine (e.g., Boink et al, 2014). Regardless of the source (mismatch in perceived information, mismatch in trade-offs between speed, risk and/or lateral accelerations) the mismatch in intended trajectories gives rise to a conflict. The driver experiences this conflict through resistive torques, which may be annoying but not counteracted, or which may require additional effort is the driver persists in maintaining his/her own intended trajectory. This kind of conflict can be quantified and detected (Nishimura et al., 2015)

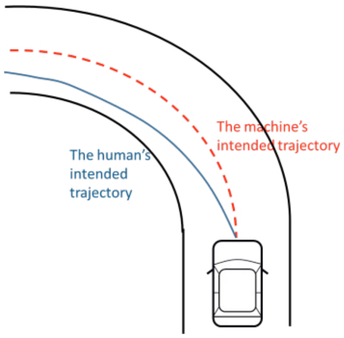

Fig. 1. An example of conflict in curve negotiation, where driver and machine select different intended trajectories.

2.2 *Types of Conflicts*

The conflict illustrated in Figure 1 is just one example of a conflict, but many more exist. In order to establish effective human-machine cooperation, all the possible conflicts should be identified. Here, we adopt the HAZOP strategy (Kletz, 2001) to identify conflicts in a systematic manner. In HAZOP, guide words are used to find possible deviations in the process to be analysed. The general guide words are: *no, more, less, as well as, part of, reverse, other than, early, late, before*, and *after*. In terms of control input from a machine, the following types can be listed.

- No input: A machine control input was expected by the human, but not given. A cause might be a sudden (sensor) failure of the machine.
- Too much/less input (quantitative): The control input from the machine is too strong (or too weak), or occurs when not expected at all. Additional manual input than expected is needed in order to achieve the original goal of the human. This might be a possible cause for the conflict depicted in Fig. 1.
- Slightly too early/too late: the control input from the machine is not well timed with the human's expectation. This might be another cause for the conflict in Fig 1. If the timing is very much different than expected, this might be perceived as a 'no input' or 'too much input'.
- Other additional input (qualitative): The control input of the machine acts on another mode/degree of freedom than expected. An example is that a driver assistance system applies brake as well as steering manoeuvre, but the human expected only the steering manoeuvre.
- Some input missing (qualitative): A part of the control input of the machine from other mode is missing. An example is that the human expected a driver assistance system to add some steering guidance torque to the steering wheel and to apply the brake, but only the former was done.
- Reverse input: The direction of the control input is opposite from the expected input. An example is that the steering assistance system avoids an object on the road by turning the steering wheel to the right when the driver expected it to turn left.
- Other input: The mode of control input is different. An example is that the machine turns the steering wheel, while the human expected it to brake.

Note here that reverse input may have two aspects. Let us think about avoidance of a pedestrian in the car driving domain (Fig. 2). First of all, the choice of the direction can be different from the human driver and the machine agent (Fig. 2(a)). This is the first aspect. Even if the direction is the same, the intended trajectory might be different (Fig. 2(b)). If the intended trajectory of the machine agent is dotted curve in Fig. 2(b), the human driver may feel scared because the machine adds some torque to the steering wheel which seems to go close to the pedestrian.